The Political Economy of Reform
LESSONS FROM PENSIONS, PRODUCT MARKETS AND LABOUR MARKETS IN TEN OECD COUNTRIES
The Political Economy of Reform

LESSONS FROM PENSIONS, PRODUCT MARKETS AND LABOUR MARKETS IN TEN OECD COUNTRIES

By William Tompson
In collaboration with Robert Price
The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation’s statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

Also available in French under the title:
L’économie politique de la réforme
SYSTÈMES DE RETRAITE ET MARCHÉS DU TRAVAIL ET DES PRODUITS DANS DIX PAYS DE L’OCDE

Corrigenda to OECD publications may be found on line at: www.oecd.org/publishing/corrigenda.

© OECD 2009

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d’exploitation du droit de copie (CFC) at contact@cfcopies.com.
Foreword

Much OECD analysis aims to identify economic policy settings that will generate high living standards. In many cases, such policy orientations have met with a considerable degree of consensus among economists. Yet, it has often been difficult in practice to put in place appropriate policies. This is what lies behind the interest in the political economy of both macroeconomic and structural policy reform. Basically, the aim of this line of research is to get a better understanding of the conditions which are conducive – or otherwise – to the implementation of better policies.

An early manifestation of this type of analysis was the 1988 publication *Why Economic Policies Change Course: Eleven Case Studies*, which looked at the factors behind changes in macroeconomic policies based on a number of case studies drawn from OECD countries. Subsequently, the focus of research shifted towards a better understanding of the factors driving structural policy reform. A number of econometric studies aimed to explain policy progress as measured by changes in various indicators of structural policy settings. A main advantage of this approach is the rigour that can be applied to testing whether some potential determinant is linked in a statistically significant way to an indicator of policy reform. A principal drawback is that the econometric approach only allows to trace the effects of determinants which can be quantified.

Recognising both the strengths and the shortcomings of the econometric approach, the OECD’s Economic Policy Committee in 2006 decided to embark on the project that lies behind the present publication. It uses a case-study approach to examine the determinants of structural reform, based on concrete reform “episodes” in ten OECD countries. The study of each episode involved not only consulting the relevant literature but also wide-ranging interviews with principal actors involved as either proponents or opponents of reform. To allow for greater comparability of the cases, they were restricted to attempts at reform of product markets, labour markets and pension systems.

The episodes differ in that some led to substantial and permanent changes in policies, whereas others led to no or only limited progress or even subsequent policy reversals. However, one conclusion that became clear early on in the project was that it is very hard to classify episodes into successes or failures. What may look initially like a failure to reform may provide the necessary impetus for a subsequent major policy change. And what looks like a successful reform may eventually be reversed. The study also provides a range of other conclusions on issues such as the role of electoral mandates, appropriate communication policies, the role of economic analysis, and the benefits of, and scope for, concertation with social partners.
The study was carried out in the Economics Department by William Tompson in collaboration with Robert Price, with statistical assistance from Thai-Thanh Dang and secretarial assistance from Susan Gascard, Veronica Humi, Sandra Raymond and Paula Simonin. Financial support for the study from the European Commission, Greece, Italy, the Netherlands and Switzerland is gratefully acknowledged.

Jørgen Elmeskov
Acting Head, OECD Economics Department
Table of contents

Executive summary .................................................................................................................................................. 9

Part I  Structural policy reform: Learning from the past

Chapter 1  Introduction: Why past can be prologue ................................................................. 15
Annex 1.A1  Statistical methods and data sources ................................................................. 27
Chapter 2  Key determinants of successful structural reform ........................................... 35

Part II  Case studies in pension-system reform

Chapter 3  France: The 2003 pension reform ................................................................. 81
Chapter 4  Italy: The 1994 pension reform proposals .................................................. 103
Chapter 5  Mexico: The reform of the federal employees’ pension system, 2001-07 .................... 123
Chapter 6  Poland: Reform of the general pension system, 1996-99 ........................................ 139
Chapter 7  Poland: Reform of the farmers’ social security system, 2003-05 ......................... 161
Chapter 8  United States: The 2005 Social Security reform debate .................................. 179

Part III  Case studies in labour-market reform

Chapter 9  France: The contrat d’insertion professionnelle, 1994 .................................. 203
Chapter 10 Germany: The Hartz reforms of the labour market, 2002-05 ..................... 223
Chapter 11 Italy: The Treu (1997) and Biagi (2002) reforms ............................................. 247
Chapter 12 Mexico: Labour law reform, 2001-03 .......................................................... 269
Chapter 13 The Netherlands: Reform of disability insurance, 2002-06 ............................ 289
Chapter 14 Spain: Labour market reform, 1994 and 1997 ............................................... 311
Chapter 15 Sweden: Sickness insurance, 1990-2002 ...................................................... 331
Chapter 16 United States: The 1996 welfare reform ....................................................... 351

Part IV  Case studies in product-market reform

Chapter 17 Australia: Power-sector reform, 1990-2004 .................................................... 375
Chapter 18 Australia: Water reform, 1994-2004 ............................................................. 397
Chapter 19 Germany: Retail opening hours, 1999-2003 .................................................. 421
Chapter 20 The Netherlands: Rent deregulation, 2004-07 ................................................ 439
Chapter 21 Spain: Retail entry and opening hours, 1995-2004 .......................................... 459
Chapter 22 Sweden: Postal liberalisation, 1992-99 .......................................................... 477

Boxes

1.1. More and less “successful” reform attempts ................................................................. 18
1.2. Screening the case-study hypotheses: the construction of the indicators ........... 21
3.1. The French pension system ..................................................................................... 85
4.1. Italy’s pre-reform pension system ................................................................. 107
4.2. The 1992 pension reform .............................................................................. 108
5.1. The Institute for Social Security and Social Services for State Employees (ISSSTE) ................................................................. 127
5.2. The ISSSTE reform plan ............................................................................... 130
6.1. The pension reform plan .............................................................................. 145
7.1. The Hausner plan .......................................................................................... 170
8.1. The Social Security system .......................................................................... 183
10.1. Unemployment and social benefits prior to the Hartz reforms ..................... 226
10.2. The “Hartz legislation” ................................................................................. 232
12.1. Key features of the labour law reform proposals ........................................ 278
13.1. Occupational disability insurance in the Netherlands prior to 2000 .......... 294
14.1. Collective bargaining under the Workers’ Statute ..................................... 315
14.2. Dismissal costs under the LET .................................................................... 316
15.1. Sickness insurance in Sweden .................................................................... 336
16.1. Federal social assistance programmes before PRWORA ......................... 355
17.1. Economic implications of the physical properties of electricity ............... 377
17.2. The National Competition Policy (NCP) ..................................................... 378
17.3. Electricity privatisation in South Australia ............................................... 385
18.1. Economic properties of water and the creation of water markets .......... 404
20.1. Rent regulation in the Netherlands ............................................................ 443

Tables

1.1. Political economy of structural reform case studies ...................................... 18
1.2. Spearman rank correlations ......................................................................... 22
1.A1.1. Spearman rank correlations ..................................................................... 33
3.1. Expected duration of retirement ..................................................................... 86
3.2. France’s pension problem: the “2040 choice” ............................................ 89
6.1. Expenditure on disability and retirement pensions ...................................... 141
6.2. Projected average replacement rates under the new system ....................... 146
7.1. Growth, deficits and public debt .................................................................... 163
7.2. Sources of farm income by farm size .......................................................... 164
7.3. Expenditures and revenues of the KRUS ..................................................... 166
10.1. Main recommendations of the Hartz Commission ................................... 229
15.1. Sickness absence of employees in eight countries, by age group ............ 334
15.2. Replacement-rate history, 1987-2003 .......................................................... 338
16.1. Economic and social indicators, 1989-96 ...................................................... 353
19.1. Shop opening hours in selected European countries, 1999 ....................... 425
20.1. Housing subsidisation in the Netherlands, 2001 ....................................... 445
20.2. Indicative targets for housing production by housing associations .......... 448
21.1. Regional barriers to retail competition, 1997-2004 ..................................... 468
## Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Reform scores across 20 cases</td>
<td>20</td>
</tr>
<tr>
<td>2.1</td>
<td>Reform mandates and reform scores</td>
<td>37</td>
</tr>
<tr>
<td>2.2</td>
<td>Reform scores and fiscal consolidation</td>
<td>41</td>
</tr>
<tr>
<td>2.3</td>
<td>Reform scores and speed of reform</td>
<td>46</td>
</tr>
<tr>
<td>2.4</td>
<td>Reform scores and reform history</td>
<td>47</td>
</tr>
<tr>
<td>2.5</td>
<td>Reform scores and the emergence of new actors</td>
<td>58</td>
</tr>
<tr>
<td>13.1</td>
<td>Disability benefit recipients, 1970-2004</td>
<td>292</td>
</tr>
<tr>
<td>13.2</td>
<td>Reform of disability benefit, 2000-04</td>
<td>296</td>
</tr>
<tr>
<td>15.1</td>
<td>Sickness cases and days lost, 1970-2002</td>
<td>335</td>
</tr>
<tr>
<td>20.1</td>
<td>Stages in the rent reform debate, 2003-07</td>
<td>448</td>
</tr>
<tr>
<td>22.1</td>
<td>The growth of CityMail</td>
<td>485</td>
</tr>
</tbody>
</table>
Executive Summary

The aim of this study is to identify political economy lessons that may be of use to policy makers seeking to design, adopt and implement structural reforms. Many of the political economy factors that facilitate or hinder economic reform have already been examined in econometric work by the OECD Economics Department. The present study builds on this work, adopting an inductive approach, based on twenty case studies of attempts to adopt and implement structural reforms in three domains, pensions, labour-markets and product-market regulation, in an effort to explore more deeply some of the econometric findings and their implications, and also to examine factors that may be extremely difficult to quantify or code for inclusion in regression analyses. The study employs a “focused comparison” approach to case-study research, which basically entails asking the same questions across a substantial number of cases in order to discern similarities among them that suggest possible generalisations. The comparative case-study approach does not permit the same the level of formal verification as may be achieved with econometric methods, but it facilitates more detailed study of the context-dependent nature of certain relationships among variables and makes it easier to trace the links between possible causes and observed outcomes in order to assess whether the causal relationships implied by a hypothesis are evident in the sequence of events as they unfold.

The study’s principal messages may be summarised as follows:

- **It is important to have an electoral mandate for reform.** This is one of the strongest findings to emerge from the study. Reform “by stealth” has severe limits, and major reforms for which governments have not previously sought public approval tend to succeed only when they generate visible benefits very rapidly, which major structural reforms generally do not. Crises do often create opportunities for such “reform surprises”, as governments are forced to reverse course in response to an emerging situation, but unless the pay-offs are very quick, the sustainability of such reforms tends to depend on the ability of reforming governments to make the case for the desirability of reforms on long-term, structural grounds. Reforms justified chiefly as responses to an immediate crisis may be difficult to sustain once the crisis has passed.

- **The importance of meaningful mandates makes effective communication all the more important.** Major reforms have usually been accompanied by consistent coordinated efforts to persuade voters and stakeholders of the need for reform and, in particular, to communicate the costs of non-reform. Where, as is often the case, the costs of the *status quo* are opportunity costs, they tend to be politically “invisible”, and the challenge is all the greater. Often, it is fairly clear who will pay the price for a reform – which firms are likely to come under pressure and which jobs may be at risk – whereas it is not at all obvious who is paying for the *status quo*: it is difficult to identify firms that have never entered the market, sectors that have not developed or the workers whom they would have employed.
• This points to the need for policy design to be underpinned by solid research and analysis. An evidence-based and analytically sound case for reform serves both to improve the quality of policy and to enhance prospects for reform adoption. The impact of economic analysis also depends on the source: research presented by an authoritative, non-partisan institution that commands trust across the political spectrum appears to have a far greater impact. Building such institutions can take time, as their effectiveness depends greatly on their reputation, but in countries that have them, their prior analysis appears to have enhanced the prospects for reform in particular areas.

• Partly for these reasons, successful structural reforms take time. The more successful reforms in the study generally took over two years to prepare and adopt – and this does not include the “pre-work” done in the many episodes in which problems and proposals had been debated and studied for years before the authorities set to work framing specific reforms. By contrast, many of the least successful reform attempts were undertaken in haste, often in response to immediate pressures: when it comes to policy reform, more haste can indeed make for less speed. Pension reforms, in particular, often have relatively long gestation times, involving a considerable amount of careful study and consultation. While governments should be ready to use political “windows of opportunity” when they open up, this may create problems if it leads to excessive haste.

• The cohesion of the government is also critical. If the government undertaking a reform initiative is not united around the policy, it will send out mixed messages, and opponents will exploit its divisions; defeat is usually the result. The case studies suggest that cohesion matters more than such factors as the strength or unity of opposition parties or the government’s parliamentary strength.

• Concertation may be helpful but is no substitute for government leadership. Reform progress may sometimes be facilitated by intensive tripartite discussions involving the government and the social partners in a highly formalised process. However, firmness of purpose on the part of the government also seems to be a critical element of success in such situations. A number of cases suggest that a “concertationist” approach is unlikely to succeed unless the government is in a position to reward co-operation by the social partners or can make a credible threat to proceed unilaterally if concertation fails. Where the government is too weak to lead or is unwilling to do so, the social partners have little incentive to make concessions. By contrast, corporatist arrangements can work well where the government is prepared to lead. Indeed, in cases where the social partners are involved in the implementation of reform, it may be useful to have them involved in reform design as well.

• The condition of the policy regime to be reformed appears to matter a great deal: some policy set-ups are more “ripe” for reform than others. Successful reforms of established policy regimes often appear to have been preceded by the “erosion” of the status quo through smaller piece-meal reforms or reform attempts; where the existing arrangements are well institutionalised and popular and there appears to be no danger of imminent breakdown, reform is far more difficult. The question of “reform ripeness” concerns not only the seriousness of the problem to be addressed but also the extent of awareness of the costs of the status quo and of
agreement on the need for change – it is a question of communication, as well as economic analysis, and is therefore to some extent endogenous to the reform process.

- Successful reform requires persistence. A further important implication of the finding concerning reform ripeness is that blocked, reversed or very limited early reforms need not be seen as failures: they may play a role in undermining the status quo and setting the stage for a more successful attempt later on. All of the more successful pension and labour-market reforms examined in this study followed earlier setbacks and many less successful reform attempts in all three domains can now be seen to have helped set the stage for subsequent, sometimes far-reaching reform initiatives.

The case studies also provide further evidence in support of some of the major findings identified by the Department’s earlier econometric work, particularly with respect to the impact of crises and the importance of sound public finances. Finally, the case studies cast some doubt on the oft-repeated claim that voters tend to punish reforming governments: the likelihood of subsequent re-election was little different between the more and less successful reform episodes, and the re-election rates for all governments in the study was close to the average for all governments at all national elections in the ten countries covered during 1992-2008 inclusive. While the number of cases involved is too small to permit generalisation, it is also worth noting that almost all the governments that successfully adopted and implemented reforms for which they had prior electoral mandates subsequently went on to win re-election.
Part I

Structural policy reform:
Learning from the past
Chapter 1

Introduction: Why past can be prologue

Introduction

The primary aim of this study is to identify political economy lessons that may be of use to policy makers seeking to design, adopt and implement structural reforms. Many of the political economy factors that facilitate or hinder economic reform have been examined in previous work by the OECD Economics Department\(^1\) and numerous other institutions and researchers.\(^2\) The present study seeks to build on this work, particularly previous work by the Economics Department, and also to contribute to the broader OECD project on “Making Reforms Happen”, which focuses on the most effective ways to realise structural reforms across a wider range of policy domains than is considered here. In contrast to most previous political economy work by the OECD, this study adopts an inductive approach, based on case studies of attempts to adopt and implement structural reforms.\(^3\) A relatively open approach, aimed at drawing general conclusions from the analysis of specific cases, would seem appropriate, given the lack of any well-established general model of the political economy of structural reform that could serve as the theoretical basis for a project covering such a wide range of reforms.

This study differs from most previous work on the political economy of reform in a number of respects. First, the academic literature on the political economy of policy reform, as well as much of that produced by international institutions, has tended until relatively recently to focus on developing countries and on instances of large-scale reorientation of economic policy, often in response to economic crises.\(^4\) While the present study includes some episodes linked to crises, its focus is on structural reforms in developed countries, most often in conditions of “normal” rather than “extraordinary” politics.\(^5\) Secondly, early research on the political economy of reform concentrated largely on macroeconomic and trade policy.\(^6\) Until recently, there has been rather less work on the political economy of structural reform, particularly the reform of product markets.\(^7\)

The study’s principal messages may be summarised as follows. First, it pays to have an electoral mandate for reform. This is one of the strongest findings to emerge from the study. Reform “by stealth” has severe limits, and reform surprises tend to succeed only when reform generates visible benefits very rapidly, which major structural reforms generally do not. The importance of meaningful mandates makes effective communication all the more important: major reforms should be accompanied by consistent co-ordinated efforts to persuade voters and stakeholders of the need for reform and, in particular, to communicate the costs of non-reform. Where, as is often the case, the costs of the status quo are opportunity costs, they tend to be politically invisible, and the challenge is all the greater. This communications challenge points to the need for policy design to be underpinned by solid research and analysis, which serves both to
improve the quality of policy and to enhance prospects for reform adoption. Partly for these reasons, the case studies suggest that successful structural reforms take time: the more successful reforms in the study generally took over two years to prepare and adopt, whereas many of the least successful reform attempts were undertaken in haste, often in response to immediate pressures. The cohesion of the government is also critical: if the government is not united around the policy, it will send out mixed messages, and opponents will exploit its divisions; defeat is usually the result. The case studies suggest that cohesion matters more than such factors as the state of the opposition or the government’s parliamentary strength. Finally, while much of the political economy literature focuses on agency and the interplay of interests, the condition of the policy regime to be reformed also matters: some are more “ripe” for reform than others. Successful reforms of established policy regimes are often preceded by the “erosion” of the status quo; where the existing arrangements are well institutionalised and popular and there appears to be no danger of imminent breakdown, reform is far more difficult.

Scope, method and structure

The method of “focused comparison”

The present study employs a “focused comparison” approach to case-study research. This method basically entails asking the same questions across a substantial number of cases in order to discern similarities among them that suggest possible generalisations. Findings generated in this way do not enjoy the level of formal verification that may be achieved via quantitative analyses of very large numbers of cases. However, the method of focused comparison offers significant advantages, chiefly by facilitating more detailed study of the context-dependent nature of certain relationships among variables. In particular, it permits a greater degree of “process-tracing” – i.e. tracing the links between possible causes and observed outcomes in order to assess whether the causal relationships implied by a hypothesis are evident in the sequence of events as they unfold. Because it examines specific cases in depth, rather than simply comparing data across cases, a focused case-study approach is better able to explore the policy process, to take account of institutional and political complexities and to explore more complex causal relationships, such as path dependence or the issues that arise when, for example, a given factor may favour adoption of a reform but hinder its implementation. A case-study approach also permits exploration of variables that can be extremely difficult to quantify or code for inclusion in regression analyses. Econometric approaches tend to set aside intervening processes and focus on correlations between the ex ante and ex post states. Hence, they can rarely give an answer as to why a correlation is observed between reform outcomes and particular conditioning factors. The present study is thus a natural follow-up to the earlier econometric work by the OECD, in providing a means to explore more deeply some of the econometric findings and their implications.

Case selection

The term “structural reform” is here defined fairly broadly as referring to changes in structural policy settings directed at improving static or dynamic resource allocation in the economy. When selecting cases for inclusion in the study, the approach has been to identify two cases for each country involved: one that was broadly successful and one that was judged to be less successful. For the purposes of the study, judgements about which cases were more and less successful are based on whether or not they were adopted and implemented, rather than on any ex post analysis of their economic impact. While
this approach helped avoid selection bias and ensured that the cases reflected a range of reform outcomes, the preparation of the case studies themselves underscored the limited validity of such labels as “success” and “failure” when applied to reform episodes (Box 1.1). Both positive and negative lessons emerge from both more and less successful cases. The cases included in the study were chosen in an attempt to satisfy a number of other criteria as well:

- **Geographic coverage.** The cases include a mix of OECD countries, in terms of size, location, income level and institutional configuration, in an effort to derive lessons which would be relevant to the broadest possible range of OECD members.

- **Coverage of structural policy domains.** Because different types of reform may entail different sorts of political economy problems, it is important to address structural reform attempts in a number of different policy domains. However, the need for depth, as well as breadth, of coverage implies that the range of such domains should be limited. The episodes selected therefore concern three broad policy areas: product-market reforms, labour-market reforms and retirement reforms. The three domains chosen are all important fields, in which there has been a good deal of reform activity in the OECD in recent years (including both successful and unsuccessful reform attempts). They also remain, to a greater or lesser extent, on the economic policy agendas of virtually all OECD countries.11

- **Economic importance.** While most of the proposed cases concern reforms expected to have a broad economic impact, some more narrowly focused reforms are also included, particularly in the field of product-market regulation. These can be limited in terms of impact, especially where they concern a single sector, but they exemplify the particular set of political economy problems that arise when governments try to open up markets to competition.

- **Balance of contrasting outcomes within policy domains.** There is a rough, though not exact, balance between more and less successful cases in each domain; the results could be skewed by focusing too much on successes in one sphere and frustrated reform efforts in another.

- **Rough contemporaneity.** The cases concern reform attempts undertaken since 1990, so that they to some extent share a common economic and political “background”, in terms of international trends and pressures, and present-day relevance to policymakers.12

With these criteria in mind, the terms of reference for twenty cases were agreed with the OECD members concerned (Table 1.1). The final list was also approved by the OECD’s Economic and Development Review Committee (EDRC) as a whole. Where possible, some indication of each reform’s impact is provided, but in some of the more recent cases, it is still too early for any definitive assessment. Nevertheless, the study is premised on the view that reforms selected for examination were potentially beneficial: some were clearly incomplete or imperfect, and a few changed so much in the course of the policy process that even some of their early supporters had doubts about the value of the “final product”, but all of the initiatives undertaken represented attempts to alter structural policy settings in ways that would improve the efficiency of resource allocation in the economies concerned.
Box 1.1. More and less “successful” reform attempts

Although this study speaks at times of more and less successful cases, these labels do not represent a stark dichotomy. Instances of total success and outright failure are both very rare, and many unsuccessful reform attempts may nevertheless yield some progress or help to pave the way for successful reforms later on. Even an unsuccessful reform initiative may therefore represent a significant step forward. Nevertheless, identifying both more and less successful cases is important in order to avoid the potential selection bias that could arise from focusing predominantly or exclusively on either positive or negative experiences of reform. “Success” is here defined not in terms of any ex post assessment of ultimate economic impact (which in many cases is as yet unclear) but as the adoption and implementation of a reform that would help a country achieve certain desirable goals, such as fiscal sustainability, enhanced competition or increased employment. An unsuccessful reform is understood not as a policy failure (i.e. an attempt to do the wrong thing) but as an attempt to undertake a desirable reform that either fails to be adopted or is adopted but not implemented.

Table 1.1. Political economy of structural reform case studies

<table>
<thead>
<tr>
<th>Retirement reforms</th>
<th>Labour market reforms</th>
<th>Product-market reforms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mexico: proposed reform of the labour law, advanced in 2002 but withdrawn in 2005.</td>
<td></td>
</tr>
</tbody>
</table>

Two points should be made at the outset in light of this approach to case selection.

- The selection of two case studies with contrasting outcomes for each country limits the scope for deriving hypotheses about the impact of institutional configurations such as electoral systems or separation of powers. By definition, there will be one more and one less successful case for each country with a given set of political institutions. This is not a problem, since the study focuses on factors that are more readily influenced by policy makers’ choices.

- The cases concern only reform initiatives that governments actually undertook; instances in which there were grounds for believing that reform was needed but no action was taken have not been considered, and this may relativise to some extent the conclusions that can be drawn from the study. Many of the most
serious political economy failures occur when resistance to change is so great that reform is not even attempted. This is important to bear in mind when comparing the findings presented here with other work.

In preparing each case study, an initial set of working hypotheses about what happened, and why, was formulated on the basis of available data and published primary and secondary sources. These preliminary ideas were then discussed with officials, experts, representatives of the social partners and other stakeholders in the country concerned, including, wherever possible, individuals directly involved in the policy process during the episode under study. The initial analysis was then revised in light of these discussions and written up following a common template to facilitate comparison. These draft case studies were then discussed by the EDRC, which provided an opportunity for additional feedback, particularly from the countries concerned; the cases were then revised again in light of this input. Although the case studies could not have been prepared without the help of national authorities in clarifying issues and facts, the interpretations and assessments are those of the OECD Economics Department.

**Screening the results: a statistical check**

Though the case-study evidence does not permit rigorous statistical assessment of potential linkages between reform outcomes and the factors analysed in the case studies, it is possible to cross-check the lessons identified in the qualitative analysis using a relatively simple set of correlations constructed on the basis of the cases (Box 1.2). To this end, a synthetic indicator of reform outcomes has been constructed across the 20 episodes, and Spearman rank correlations have been calculated to measure the strength and direction of the links between reform outcomes and individual conditioning factors. The rank correlation approach is well suited to use with small samples.

The indicator reform outcomes has been developed on the basis of the evidence presented in the case studies, scoring the reform initiatives examined in the case studies from 0 to 5 by summing their scores on three criteria:

- **Adoption** of the reform proposal is scored from 0 to 2, where 0 signifies a failure to adopt, 2 reflects the adoption of all or nearly all of the proposal, even if subject to minor modifications and concessions, and 1 stands for intermediate cases, where adoption was very partial.

- **Implementation** success is also scored from 0 to 2, again reflecting the range from implementation failure (or reversal of the reform) through partial implementation to full implementation of the measures adopted.

- The **follow-up** to the reform is scored, with 0 signifying no further reform progress as a result of the episode and a score of 1 applied in those instances where the evidence suggests that a reform proposal (whether adopted and implemented or not) contributed to subsequent reform progress in the same field.

This indicator is clearly rather crude, but it has the advantage of being coded on the basis of fairly clear, simple criteria. The construction of a more elaborate indicator would have required more complex and subtle judgments and would thus have introduced a greater degree of subjectivity into the coding. Figure 1.1 shows the reform outcome scores of the 20 cases.
In general, the signs of the correlation coefficients for the aggregate indicators are as expected and are statistically significant across the 20 cases. Many of the specific process/design variables are also significant at the 5 or 10% levels across the full set of cases, even though not all are significant for the individual policy domains. The most important results are summarised in Table 1.2, and their implications are considered throughout the discussion that follows. The full results are presented in Table A1.1 in the Annex. As Table A1.1 shows, some variables that are not discussed in the foregoing analysis were tested in order to ensure that the null hypothesis of no linkage to reform outcomes could not be rejected. It is important to emphasise that the data used in the correlations add no “new” information to the synthesis: the statistical exercise is intended not to extend the study but as a means of screening its major conclusions.

The very high correlations reported in Table 1.2 doubtless reflect in large part the fact that many of the individual factors explored are closely related to, and highly correlated with, one another. If the data were suitable for a regression analysis or other more sophisticated techniques, it would be possible to untangle these relationships and the apparent impact of the individual variables would probably be reduced in many cases. Since this is not possible, one can only take note of such cross-correlations among independent variables and observe that many of them clearly tap into the same underlying information.
The individual factors influencing reform outcomes and discussed in this study have been coded, using binary (0/1) or simple rank-order (0-2) values to score each episode with respect to the given factor. For example, with respect to electoral mandates, a score of 2 is assigned where the government could claim a clear mandate for reform; a score of 0 signifies a reform for which there was no mandate; and a score of 1 is applied to intermediate cases, where the government’s claim to a mandate was credible but still in some way qualified or contested. Although the case studies provide the evidence on which the codings are based, they do involve a degree of subjective judgment. The variables have therefore deliberately been kept as simple as possible (binary or 0-2), again in order minimise any potential bias that such judgments might introduce. The individual factors have all been coded in such a way that a higher score is expected to be correlated with a higher probability of success. Definitions and details can be found in the Annex.

Aggregate indicators have also been constructed in order to summarise the impact of a number of related factors by broad areas: political cycles, macroeconomic cycles and so on. These are calculated by simply summing the individual scores on the relevant individual factors operationalised as described in the previous paragraph. See the Annex for more detail.

Finally, broad indicators of policy reform (changes in overall structural policy settings) in each of the three domains were constructed in an effort to capture the larger reform “environment” in which specific reform initiatives were undertaken. These are based on the synthetic indicators of policy settings developed by the OECD Economics Department (see the Annex for details of their construction) and they are intended to reflect the intensity of reforms in each domain. They provide a basis for looking at interactions between reforms in different domains and also for assessing whether or not a given reform was undertaken in the context of a broader reform drive or as a one-off.

The data used for scoring the indicators come chiefly from the case studies. However, where variables are readily quantifiable (for example, data on growth, unemployment and fiscal balances), other OECD and World Bank data were also used. Data sources are discussed in greater detail in the Annex. Although these additional data were used, it is important to stress that all data employed concern only the episodes under study: the Spearman correlations are not a method for assessing whether the case-study findings are generalisable to other episodes but rather a screening and cross-check device to ensure that the case studies really do lend support to the hypotheses derived from them.

1. More complex ordinal rankings would require finer judgements. In the absence of clear, agreed criteria for making such finely grained judgements, this would increase the scope for subjective bias.
2. The policy indicators are coded in such a way as to capture the intensity of ongoing policy reforms in the three domains at the time any given reform was adopted. The indicators used are the following: i) a composite OECD product-market reform indicator covering deregulation in seven non-manufacturing sectors; ii) an average indicator for labour-market reform covering reforms of the unemployment benefit system and employment protection legislation; and iii) a measure of pension reform intensity based on changes in the implicit tax rate on older workers.
Table 1.2. **Spearman rank correlations**

<table>
<thead>
<tr>
<th></th>
<th>Labour-market reforms</th>
<th>Pension reforms</th>
<th>Product-market reforms</th>
<th>Reforms in all domains</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Political conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear mandate for reform</td>
<td>0.786 **</td>
<td>0.783 **</td>
<td>0.539</td>
<td>0.597 **</td>
</tr>
<tr>
<td>Government cohesion</td>
<td>0.675 *</td>
<td>0.906 **</td>
<td>0.420</td>
<td>0.680 **</td>
</tr>
<tr>
<td><strong>Economic cycles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Economic trigger (aggregate)</em></td>
<td>-0.644 **</td>
<td></td>
<td>0.487</td>
<td>-0.039</td>
</tr>
<tr>
<td>Ongoing fiscal consolidation</td>
<td>0.068</td>
<td>-0.533</td>
<td>-0.630</td>
<td>-0.379 *</td>
</tr>
<tr>
<td><strong>Policy interactions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labour-market reforms</td>
<td>0.060</td>
<td>-0.544</td>
<td>0.000</td>
<td>-0.114</td>
</tr>
<tr>
<td>Pension reforms</td>
<td>-0.834 **</td>
<td>0.889</td>
<td>0.258</td>
<td>0.284</td>
</tr>
<tr>
<td>Product-market reforms</td>
<td>-0.423</td>
<td>0.866</td>
<td>0.894 **</td>
<td>0.381</td>
</tr>
<tr>
<td><strong>Reform process</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Scope and timing (aggregate)</em></td>
<td>0.427</td>
<td>0.188</td>
<td>0.970 **</td>
<td>0.554 **</td>
</tr>
<tr>
<td>Time taken in preparation of reform</td>
<td>0.447</td>
<td>0.763</td>
<td>0.399</td>
<td>0.553 **</td>
</tr>
<tr>
<td>Exogenous event as trigger for reform</td>
<td>-0.410</td>
<td>-0.429</td>
<td>0.840 **</td>
<td>-0.032</td>
</tr>
<tr>
<td>&quot;Ripeness&quot; of policy regime for reform</td>
<td>0.900</td>
<td>0.853</td>
<td>0.105</td>
<td>0.599 **</td>
</tr>
<tr>
<td><em>Communication (aggregate)</em></td>
<td>0.842 **</td>
<td>0.429</td>
<td>0.493</td>
<td>0.592 **</td>
</tr>
<tr>
<td>Communication of costs of status quo</td>
<td>0.907 **</td>
<td>0.533</td>
<td>0.539</td>
<td>0.614 **</td>
</tr>
<tr>
<td>Government leadership</td>
<td>0.907 **</td>
<td>0.539</td>
<td>0.853 **</td>
<td>0.699 **</td>
</tr>
<tr>
<td>Clear &quot;ownership&quot; of the reform</td>
<td>0.761 *</td>
<td></td>
<td>0.320</td>
<td>0.379</td>
</tr>
<tr>
<td><em>Dealing with opponents (aggregate)</em></td>
<td>0.722 **</td>
<td>0.627</td>
<td>0.682</td>
<td>0.702 **</td>
</tr>
<tr>
<td>Emergence of new actors as force for reform</td>
<td>0.665</td>
<td>0.426</td>
<td>0.804</td>
<td>0.636 **</td>
</tr>
<tr>
<td>New role for opponents under reform</td>
<td>0.363</td>
<td>0.905</td>
<td>0.539</td>
<td>0.578 **</td>
</tr>
<tr>
<td>Activation of winners in support of reform</td>
<td>0.617</td>
<td>0.544</td>
<td>0.000</td>
<td>0.360 *</td>
</tr>
</tbody>
</table>

* significant at 10% level of confidence; ** significant at 5% level of confidence.

1. Current primary deficit down by at least 1% of GDP on previous year.

*Source*: OECD calculations, based on case studies, OECD and World Bank data.
Notes


2. For overviews of this literature see, in particular, Alesina and Perotti (1994); Williamson (1994); Rodrik (1996); Drazen (2000); and IMF (2004).

3. The Department has employed a case-study approach before; see OECD (1988).

4. This emphasis is apparent in the overviews provided in Alesina and Perotti (1994); Williamson (1994); and Rodrik (1996).

5. The distinction originates with Balcerowicz (1995), who sees periods of “extraordinary politics”, associated with such events as the fall of communism, as critical in creating conditions for radical, rapid and comprehensive reforms.


8. See George and Bennett (2005) for details. This is broadly the approach taken by the authors contributing to Williamson (1994) and Immergut et al. (2007).

9. For more on the benefits and limitations of this method, see George and Bennett (2005:17-32).

10. This follows Koromzay (2004).

11. A deliberate decision was taken to avoid fields such as healthcare, where there is little or no consensus regarding OECD “best practice”.

12. A study concentrating on, say, the 1960s or 1970s, would probably give greater attention to issues like financial and trade liberalisation. It would also confront a problem insofar as the political systems and policy processes of many members have changed fundamentally since then, raising questions about how relevant the study’s findings might be.
Bibliography


Annex 1.A1

Statistical methods and data sources

Introduction

Previous work on political economy by the Economics Department has employed econometric methods, relying on time-series data covering a large sample of OECD economies over an extended period of time. By contrast, the present study looks at twenty specific reform episodes in ten OECD countries at different points in time. Although the case studies are not suitable for econometric analysis, it is possible to use relatively simple statistical tests to ensure that the hypotheses derived from the case studies stand up to a more rigorous quantitative approach. To that end, a data set has been constructed on the basis of the case studies. These data add no new information to the synthesis: the statistical exercise is used not to extend the study but as a means of screening its major conclusions. It is also important to acknowledge that, since most of the data are coded on the basis of the case studies themselves rather than drawn from external sources, there is an element of subjective judgement in the coding. In order to minimise any potential bias arising from such judgements, the variables have been kept as simple as possible: most are binary or very simple ranked ordinal values, rather than continuous, and the criteria for coding have been kept as simple as possible. Given the nature of the data, the most appropriate way to explore the links or influences across variables is to use simple statistical screening methods, such as Spearman rank correlations. These measure the direction and strength of the relationship between pair-wise variables (the outcome of reforms and individual political factors).

This annex describes the methodological choices involved in screening the statistical correlations between reform outcomes and the political economy factors analysed in the case studies. The first section describes the approach used to code the qualitative information drawn from the case studies for use in the rank correlations, as well as the coding of economic and political economy variables drawn from other sources. The second section explains the method – Spearman rank correlations – used to screen the links across ranked ordinal variables.

Construction of the indicators and data sources

The following indicators have been developed to translate qualitative information into quantitative data based on ordinal values, either binary or rank-ordered. In some cases, continuous time-series data have been converted into ordinal values so as to identify specific episodes during the reforms’ adoption/implementation, while at the same time ensuring homogeneity across variables. The indicators have been constructed to reflect the political economy hypotheses discussed in the analysis.
Synthetic indicator of reform outcomes

A synthetic indicator of policy outcomes has been developed on the basis of the case studies, scoring the reform initiatives examined in the case studies from 0 to 5 by summing their scores along three dimensions:

- **Adoption** of the reform proposal is scored from 0 to 2, where 0 signifies a failure to adopt, 2 reflects the adoption of all or nearly all of the proposal, even if subject to minor modifications and concessions, and 1 stands for intermediate cases, where adoption was judged to be very partial.

- **Implementation** scores depend in part on the scores for reform adoption. For reforms with an adoption score of 2, implementation success is scored from 0 to 2, again reflecting the range from implementation failure (or reversal of the reform) through partial implementation to full implementation of the measures adopted. However, where reform adoption was partial, implementation is scored as 0 or 1.2

- The **follow-up** to the reform is scored, with 0 signifying no further reform progress as a result of the episode and a score of 1 applied in those instances where the evidence suggests that a reform proposal (whether adopted and implemented or not) contributed to subsequent reform progress in the same field.

Political cycle indicators

The following variables concern the political conditions in which the reform was initiated, including the strength and cohesion of the government, its mandate and the electoral cycle.

- **Government mandate (0-2)**. A score of 2 is assigned where the government has a clear electoral mandate for reform, 1 is assigned where its claim to a mandate is unclear and 0 applies in cases where no such mandate exists. The scoring is based on the case studies.

- **Stable political time horizon (0-4)**. The variable indicates the number of years left in the government’s current term. Only full years are counted. “0” is scored in election year and n-1 in the year after the election with n= length of term. Data are taken from World Bank (2007).

- **Government control of parliament (0-3)**. This indicator measures the strength of the government vis-à-vis parliament and opposition parties. This aggregate indicator is the sum of scores on three binary variables drawn from World Bank (2007), which have been coded as follows. A score of 1 is assigned if the government commands a majority (50%+1) of the seats in the lower house. A further score of 1 is assigned if the executive commands an absolute majority in all legislative chambers with law-making powers in respect of the reform. A third dimension is scored as 1 unless the main opposition holds a majority in one or more legislative chambers.3

- **Government cohesion (0/1)**. This variable is coded 0 if the case study shows clear evidence of government divisions over the reform and 1 otherwise.

- **Government composition (0-2)**. This indicator is based on different but exclusive cases. A single-party government is coded as 2. A “dominant-party coalition”, in which one party holds 40% of the seats in the lower house, is coded as 1. A more
fragmented coalition is scored as 0. These data have been drawn from World Bank (2007) but corrected in some cases by the Economics Department.  

**Macroeconomic cycle indicators**

This set of indicators includes those constructed in connection with the Department’s earlier econometric work, as well as new indicators developed to capture those elements of the economic cycle that could have influenced the adoption of reforms in the cases under study. Unless indicated otherwise, the sources for these variables are Høj et al. (2006) and OECD (2008a).

- **Economic crisis (0/1).** If the magnitude of the estimated output gap exceeds -3%, the case is coded as a crisis. Of course, the output gap may lag behind perceptions of a slowdown based on annual growth rates, since the output gap may take some time to turn negative after growth falters.

- **Economic downturn (0/1).** The downturn variable is coded as 1 where the year of the reform sees at least two consecutive quarters of negative economic growth. A lagged downturn variable is likewise coded as 1 if the previous year saw at least two quarters of contraction.

- **Economic upswing (0/1).** A value of 1 is recorded for both of the first two years of positive growth following a recession.

- **Economic trigger (aggregate).** This aggregate indicator is the sum of the economic crisis variable and the economic downturn variables for the current and previous years.

- **Large increase in unemployment rate (0/1).** This variable is set to 1 if the change in the unemployment rate is positive and greater than a threshold value, defined as the average change in unemployment rates from 1975 to 2007 within each country. The reference threshold is country-specific, because political resistance to, or support for, reforms is more likely to be sensitive to changes in the unemployment rate that are large by the standards of the country’s recent past, whether or not they would be considered large in other OECD economies.

- **High level of unemployment (0/1).** A value of 1 is assigned if the unemployment rate at the time of the reform is greater than 70% of the maximum value reached over the 1975-2007 period for a given country.

- **Falling unemployment (0/1).** This is indicator is set to 1 if the unemployment rate has fallen over the year prior to the reform.

- **High fiscal pressure (0/1).** Fiscal pressure is observed (score = 1) if the ratio of net government lending (including interest payments) to GDP is greater than 80% of the maximum deficit observed over the 1975-2007 period for the country in question.

- **Weak fiscal position (0/1).** A score of 1 applies if the primary balance is positive; otherwise, this is coded as 0, reflecting the limited scope for compensatory measures to soften opposition to reform.

- **Deterioration in primary deficit.** A score of 1 is assigned if the ratio of net government lending to GDP (excluding interest payments) is rising.
• **On-going fiscal consolidation.** If the primary deficit (excluding interest payments) as a percentage of GDP has fallen by more than 1 percentage point on the previous year, a score of 1 is assigned.

• **Fiscal recovery.** When the primary balance has been positive for two consecutive years, a score of 1 is assigned.

### Interactions with other policy reforms

These indicators seek to capture the broader reform “environment” in which the cases are set. As such, the indicators of policy reforms cover (where possible) different areas/sectors in each policy domain and are intended to reflect the degree of reform effort or the “intensity” of reforms in each domain. They use the synthetic indicators of policy settings developed by the OECD Economics Department. The intensity of reforms is coded as follows:

- When the reduction (improvement) in a reform indicator score is greater than 80% of the maximum variation observed in that country, a score of 3 is assigned, indicating a major reform effort.
- When the reduction (improvement) falls between 40% and 80% of the maximum variation observed, a score of 2 is assigned.
- A score of 1 is assigned when the variation is between 0% and 40%.
- A score of 0 applies when the variation is null or negative.

The indicators employed for this scoring are set out below.

#### Labour policy indicator

This policy indicator reflects reforms to current systems of unemployment insurance and employment protection legislation that have taken effect. The indicator is calculated as an average of:

- the gross unemployment benefit replacement rate for low-income workers (67% of average production worker earnings) in the first year of unemployment;
- the generosity of long-term unemployment benefit (in the fourth and fifth years of unemployment) relative to average spells of unemployment; and
- the summary indicator or the stringency of employment protection legislation for all types of contract (measured as the average of indefinite and fixed-term contracts).

The data for these indicators are taken from OECD (2008b).

#### Product market indicator

The broad policy reform indicator for product-market regulation is a summary indicator of regulatory impediments to product market competition in seven non-manufacturing sectors (gas, electricity, post, telecommunications, passenger air transport, railways and road freight). The indicator covers regulations on barriers to entry, public ownership, market structure, vertical integration and price controls. Data are taken from Conway and Nicoletti (2006).


Pension policy indicator

The indicator of pension policy reform reflects changes in both old-age pension systems and early retirement schemes. It is a simple average of:

- the implicit tax rate on continued work for workers aged 55-59 in early retirement schemes; and
- the implicit tax rate on continued work for workers aged 60-64 in both early retirement schemes and old-age pension systems.

The implicit tax rates as from 2000, estimated by the OECD Economics Department, are calculated so as to reflect reforms in policy parameters that have been legislated. This method excludes de facto cohort effects and phased-in effects; however, the data through 1999 reflect the actual implementation of legislated changes rather than their adoption, so reforms adopted with long transition periods are only reflected as they take effect. The break in the time series is non-negligible, since most pension reforms are phased in over an extended period. The original data on implicit tax rates on continued work for various age groups are described in detail in Duval (2004); the actual data are taken from that source for earlier years and from OECD (2008c) for the period from 2003.

Because the parameters of the main old-age pension system determine the value of this indicator to a great extent, a significant reform of that system will of itself drive a significant change in this indicator. The pension policy indicator is thus of limited use in defining a “reform context” for the specific pension reforms addressed in the case studies in the way that the other indicators do. However, it provides a useful indicator for assessing interactions among reform processes in different domains.

Process and decision-making indicators

The variables described in this section are all coded on the basis of the case studies. Aggregated indicators are constructed to summarise broad areas of policy process or design, such as the time frame, communications and negotiation with opponents. These are simply the sum of the specific factors underlying them.

Time and scope of the reform

- Preparation time (0-2). This variable measures the time from the point at which the reform process is initiated until the reform is either adopted or rejected: more than one year = 2; 3 months to one year = 1; less than 3 months = 0.

- Exogenous event as a trigger for reform (0-2). Exogenous event(s) forced the government to act = 2; exogenous event(s) relaxed the constraints on reform = 1; no exogenous trigger = 0.

- Reform “ripeness” (0/1). For simplicity, the notion of a policy regime’s “ripeness for reform” is coded purely on the basis of past attempts to reform it comprehensively or at the margins, whether these were thwarted, reversed or sustained. Where the reform under study is preceded by previous reform efforts, a score of 1 is assigned; otherwise, the value is 0.

- Breadth (0/1). Since some strands of the literature argue that broad reforms face stiffer opposition than narrowly targeted ones, a score of 1 is assigned if the
reform is targeted at a particular sector or group; reforms that are broad in scope (affecting the median voter significantly and directly) are assigned 0.

**Communication and consultation**

- *Awareness of the need for change (0/1).* A score of 1 is assigned if the evidence suggests a widespread understanding among relevant stakeholders of the costs of the status quo, whether or not there was a consensus on what should be done.
- *Government leadership (0/1).* A score of 1 is assigned to cases in which the government was strongly and publicly committed to reform; otherwise, this variable is coded as 0.
- *Consensus building institutions (0/1).* If the government relies on specialised commissions or other consensus-building institutions to win adoption of the reform, this indicator is scored as 1.
- *“Ownership” (0/1).* Clear “ownership” of the reform by an identifiable individual, group or institution is coded as 1; otherwise 0.

**Dealing with opponents**

- *Corporatist processes (0/1).* If formal corporatist negotiations involving unions and employers formed part of the policy process, this variable is coded as 1; otherwise, it is 0.
- *Acquired rights (0/1).* This variable is scored as 0 where acquired rights are violated and 1 where they are respected.
- *New actors (0/1).* A score of 1 indicates that reform progress was facilitated by the emergence of new actors and/or institutions as the reform process unfolded.
- *Opponents (0/1).* This variable is scored as 1 if the reform offered opponents a new role or new opportunities that partly offset the costs of reform to them.
- *Activation of winners (0/1).* If potential winners from a given reform actively lobbied in support of it, this variable is coded as 1.

**Spearman rank correlations**

**Definition**

The most appropriate method for measuring the degree of correlation between ordinal values is the Spearman rank correlation. In principle, the Spearman correlation (\(\rho\)) is simply a special case of the Pearson product-moment coefficient, in which two sets of data \((X_i, Y_i)\) are converted to rankings \(x_i\) and \(y_i\) before calculating the coefficient. In practice, however, a simpler procedure is normally used to calculate \(\rho\). The raw scores are converted to ranks, and the differences \(d_i\) between the ranks of each observation on the two variables are calculated.

If there are no tied ranks, i.e. if \(\exists_{i \neq j} (X_i = X_j \land Y_i = Y_j)\) then \(\rho\) is given by:

\[
\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}
\]
where:

\( d_i = x_i - y_i = \) the difference between the ranks of corresponding values \( X_i \) and \( Y_i \), and

\( n = \) the number of values in each data set (same for both sets).

**Test of significance**

To test the significance of the Spearman rank correlations, the estimated values have been tested against the null hypothesis (H0). The tests were run at both 10% and 5% confidence levels to ensure that the estimated correlations were significantly different from zero.

Table 1.A1.1. Spearman rank correlations

<table>
<thead>
<tr>
<th>Political cycles</th>
<th>Labour-market reforms</th>
<th>Pension reforms</th>
<th>Product-market reforms</th>
<th>All reforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear mandate for reform</td>
<td>0.7862 **</td>
<td>0.7833 *</td>
<td>0.5394</td>
<td>0.5972 **</td>
</tr>
<tr>
<td>Stable political horizon</td>
<td>-0.4521</td>
<td>0.3752</td>
<td>-0.0758</td>
<td>-0.1038</td>
</tr>
<tr>
<td>Government strength vis-à-vis parliament and opposition</td>
<td>0.3567</td>
<td>0.3947</td>
<td>-0.523</td>
<td>0.0781</td>
</tr>
<tr>
<td>Government cohesion</td>
<td>0.6751 *</td>
<td>0.9058 **</td>
<td>0.4201</td>
<td>0.6797 **</td>
</tr>
<tr>
<td>Government composition</td>
<td>-0.4826</td>
<td>0.3015</td>
<td>0.1112</td>
<td>-0.0257</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic cycles</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic factors (aggregate)</td>
<td>-0.644 **</td>
<td>-0.048 **</td>
<td>0.564</td>
<td>0.039</td>
</tr>
<tr>
<td>Major crisis</td>
<td>0.177</td>
<td></td>
<td>0.123</td>
<td></td>
</tr>
<tr>
<td>Economic downturn</td>
<td>-0.423</td>
<td></td>
<td>0.426</td>
<td>0.000</td>
</tr>
<tr>
<td>Economic upswing following adoption</td>
<td>-0.531</td>
<td>-0.135</td>
<td>0.539</td>
<td>-0.087</td>
</tr>
<tr>
<td>Unemployment factors (aggregate)</td>
<td>-0.303</td>
<td>-0.048</td>
<td>0.564</td>
<td>0.039</td>
</tr>
<tr>
<td>Falling unemployment</td>
<td>0.363</td>
<td>-0.213</td>
<td>-0.405</td>
<td>0.010</td>
</tr>
<tr>
<td>High level of unemployment</td>
<td>-0.531</td>
<td>-0.533</td>
<td>0.539</td>
<td>-0.245</td>
</tr>
<tr>
<td>Rising unemployment</td>
<td>0.000</td>
<td>0.426</td>
<td>0.320</td>
<td>0.223</td>
</tr>
<tr>
<td>Fiscal factors (aggregate)</td>
<td>-0.317</td>
<td>-0.091</td>
<td>0.379</td>
<td>-0.047</td>
</tr>
<tr>
<td>High level of fiscal pressures</td>
<td>-0.531</td>
<td>-0.539</td>
<td>0.426</td>
<td>-0.178</td>
</tr>
<tr>
<td>Rising deficit</td>
<td>-0.293</td>
<td>0.302</td>
<td>0.402</td>
<td>0.098</td>
</tr>
<tr>
<td>Weak fiscal position</td>
<td>0.060</td>
<td>0.302</td>
<td>0.426</td>
<td>0.254</td>
</tr>
<tr>
<td>Fiscal consolidation</td>
<td>0.068</td>
<td>-0.533</td>
<td>-0.533</td>
<td>-0.379 *</td>
</tr>
<tr>
<td>Fiscal recovery</td>
<td>0.270</td>
<td>-0.135</td>
<td>-0.405</td>
<td>0.045</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Policy interactions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour-market reforms</td>
<td>0.060</td>
<td>-0.544</td>
<td>0.000</td>
<td>-0.114</td>
</tr>
<tr>
<td>Pension reforms</td>
<td>-0.834 **</td>
<td>0.889 **</td>
<td>0.258</td>
<td>0.284</td>
</tr>
<tr>
<td>Product-market reforms</td>
<td>-0.423</td>
<td>0.868</td>
<td>0.894 **</td>
<td>0.381</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process and design of reforms</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope and timing (aggregate)</td>
<td>0.427</td>
<td>0.188</td>
<td>0.970</td>
<td>** 0.554 **</td>
</tr>
<tr>
<td>Time taken in preparation of reform</td>
<td>0.447</td>
<td>0.763</td>
<td>* 0.399</td>
<td>0.553 **</td>
</tr>
<tr>
<td>Exogenous event as trigger for reform</td>
<td>-0.410</td>
<td>-0.429</td>
<td>0.840</td>
<td>** -0.032</td>
</tr>
<tr>
<td>Ripeness’ of policy regime for reform</td>
<td>0.900 **</td>
<td>0.853 **</td>
<td>0.105</td>
<td>0.599 **</td>
</tr>
<tr>
<td>Communication (aggregate)</td>
<td>0.842 **</td>
<td>0.429</td>
<td>0.493</td>
<td>0.592 **</td>
</tr>
<tr>
<td>Communication of costs of status quo</td>
<td>0.907</td>
<td>0.533</td>
<td>0.539</td>
<td>0.614 **</td>
</tr>
<tr>
<td>Government leadership</td>
<td>0.907</td>
<td>** 0.539</td>
<td>0.853 **</td>
<td>0.699 **</td>
</tr>
<tr>
<td>Use of consensus-building institutions</td>
<td>0.423</td>
<td>0.302</td>
<td>-0.101</td>
<td>0.224</td>
</tr>
<tr>
<td>Clear “ownership” of reform</td>
<td>0.761</td>
<td></td>
<td>0.320</td>
<td>0.379</td>
</tr>
<tr>
<td>Dealing with opponents (aggregate)</td>
<td>0.722</td>
<td>0.627</td>
<td>0.682</td>
<td>0.702 **</td>
</tr>
<tr>
<td>Formal concertation with social partners</td>
<td>0.423</td>
<td>0.533</td>
<td>0.320</td>
<td>0.438</td>
</tr>
<tr>
<td>Respect for acquired rights</td>
<td>0.000</td>
<td>0.533</td>
<td>1.000</td>
<td>0.260</td>
</tr>
<tr>
<td>Emergence of new actors as force for reform</td>
<td>0.665</td>
<td>0.426</td>
<td>0.804</td>
<td>0.636 **</td>
</tr>
<tr>
<td>New role for opponents under reform</td>
<td>0.363</td>
<td>0.905 **</td>
<td>0.539</td>
<td>0.578 **</td>
</tr>
<tr>
<td>Activation of winners in support of reform</td>
<td>0.617</td>
<td>0.544</td>
<td>0.000</td>
<td>0.360 *</td>
</tr>
</tbody>
</table>

* Significant at 10% level of confidence; ** significant at 5% level of confidence.

Notes

1. See especially Duval and Elmeskov (2005); and Høj et al. (2006). IMF (2004) is based on a similar approach to this earlier work by the Secretariat.

2. The logic here is that a reform was not fully implemented if large elements of it failed to win adoption, even if those measures that were adopted were also implemented in full.

3. Clearly, a score of 1 on the second dimension precludes a score of 0 on the third; however, the third dimension distinguishes cases of parliamentary fragmentation from episodes in which the government faces at least one chamber effectively controlled by its opponents.

4. In the Swedish (1991) and Spanish (1994 and 1997) cases, the Bank’s database records single-party minority governments as majority coalitions, because they were sustained in office by minor party support in parliament. However, these were not coded as coalitions for the present study: no coalition programme was agreed, no other parties joined those governments and, in practice, the particular combination of non-government parties on which the cabinet relied to adopt legislation varied across time and from issue to issue. The literature generally treats these as one-party minority governments rather than majority coalitions. Mexico has been recoded as having a one-party government in both cases, because, although members of other parties worked with the Fox and Calderón administrations, there was no coalition between the PAN party and its rivals. Italy’s governments have been re-coded as coalitions for all cases, reflecting the continued importance of the distinct political parties that made up the two broad coalitions contesting power in the elections of 1994-2000; the World Bank database treats the “Olive Tree” coalition and the “House of Freedoms” coalition as one-party governments. Again, both popular discussions and the scholarly literature continue to regard them as multi-party coalitions. Finally, France is coded as a one-party government in 2003, reflecting the creation of the Union for a Popular Majority (originally the Union for a Presidential Majority) in 2002.

5. In previous econometric work by the Department, this threshold was defined as twice the standard deviation observed in 20 OECD countries.
Chapter 2

Key determinants of successful structural reform

Introduction

This chapter presents the main lessons that seem to emerge from 20 studies of structural reform episodes in member countries completed during 2007-08. While this exercise has not yielded any one-size-fits-all “toolkit” for reformers, nor even suggested that such a toolkit exists, it does point to a number of striking regularities in the way reform processes unfold. The chapter proceeds as follows. The next section considers the role of political and economic factors exogenous to the reform processes under study, though not necessarily to the political process as a whole. The discussion then turns to issues related to the scope and timing of reform, particularly speed and sequencing. The third major section deals with questions of consultation and public communication, and the fourth set of issues concerns the way reformers deal with actual or potential opponents of reform.

Exogenous factors

Many of the factors that affect the prospects for reform are clearly exogenous to the reform process itself. These include both political/institutional and economic variables. Some, such as demography, are largely beyond the influence of policy makers, except possibly over the very long run, while others, like government coalition agreements, are products of political choice but may still be understood as exogenous to the reform, since they cannot easily be altered in the short term and they therefore define the context in which reform proposals are advanced and debated. The evidence of the case studies examined here broadly supports the view that political-institutional and economic framework conditions are critical to the prospects for reform. The case studies underscore the importance of electoral mandates and sound public finances in providing a favourable context for structural reform. The importance of such conditioning variables implies that circumstances and timing are important, but also that, where framework conditions are not propitious, preparatory action may lay the groundwork for later reform. If, for example, fiscal weakness is an impediment to desirable product- or labour-market reforms, then the determination to pursue such reforms may represent yet another reason to pursue near-term fiscal consolidation.

The political context

Elections matter

One of the strongest findings to emerge from the case studies is that governments are likely to achieve more, at lower cost, when the architects of reform can credibly claim an
electoral mandate for reform and the issue is still salient in voters’ minds (France, Poland and Mexico: pensions; Germany and Spain: labour market; United States: welfare reform; Netherlands: disability insurance). Where newly elected or re-elected governments act without a mandate for reform, or on the basis of only vague commitments to address particular problems, voters and other actors are far less likely to accept what they propose (France: labour market; Italy: pensions and labour market; Poland: farmers’ social insurance; Sweden: sickness insurance; United States: pensions; Mexico: labour law; Netherlands: rent regulation). The claim that electoral mandates for reform are critical stands in contrast to some previous research, which suggests that they are not a necessity. Williamson and Haggard (1994), for example, point to many examples of successful reform programmes implemented by governments doing the opposite of what they had promised to do when running for office – usually governments elected on populist platforms which then pursued fiscal consolidation and market reforms. However, these cases generally concern reforms undertaken in response to crises, and the reforms in question often generated substantial benefits quickly, since they were usually aimed at macroeconomic stabilisation. Reforms that create large numbers of winners rapidly may win acceptance ex post even if they faced strong ex ante opposition, but since most structural reforms are unlikely to generate large pay-offs as quickly as a successful stabilisation, securing public acceptance of such reforms beforehand would seem to be all the more important.

The evidence of the case studies suggests that the importance of electoral mandates may vary across policy domains. The need for a mandate stands out most strongly in respect of pension and labour-market reforms, but the evidence for it is rather weaker in respect of product-market reforms, possibly because many of them had a sectoral focus and some were technically quite complex. Their direct impact on most households was not immediately apparent, since it was sometimes difficult to persuade the public of the benefits to consumers arising from reforms that would eliminate producers’ rents, and the conflicts of interest involved did not always break down along party-political lines. Regulation of retail market entry and opening hours in Spain and Germany, for example, does not appear to have been a salient issue in general election campaigns, nor was postal reform in Sweden. Power-sector reform did figure prominently in some Australian state elections, but electoral controversies were focused almost wholly on privatisation; other aspects of this very complex reform did not capture the electorate’s attention. The only product-market episode in which electoral factors loomed large was rent deregulation in the Netherlands, the initiative that had, potentially at least, the broadest direct impact on voters of any product-market reform in this study. The rank correlations confirm both the importance of clear mandates for reform and the differences in the relevance of this factor outlined above (Table 1.2 and Figure 2.1).

The case studies also suggest that the timing of a reform within the electoral cycle can be important. Governments generally do seem to have a window of opportunity to move on reforms early in their mandates: in the case studies, eight of the ten reforms that were both adopted and sustained were passed within the first two years of the government’s term of office (France, Mexico: pensions; Italy, Germany and Spain: labour market; United States: welfare reform; Netherlands: disability insurance; Sweden: postal reform). The fact that six of the eight cases just cited are also referenced in the previous paragraph suggests that this finding is closely linked to the issue of electoral mandates: it pays to act where those mandates are relatively fresh. The Mexican pension reform is particularly striking: the first attempt to adopt it failed largely because it was shortly before presidential and congressional elections, whereas a second attempt to pass the
measure, immediately after the elections, was successful. However, where the government is not seen to have a mandate for reform, prompt action may not help: in all three cases in which governments tried to adopt reforms within months of winning an election and failed, they had problems claiming a mandate for their proposals (France: labour market; Italy and United States: pensions). As noted above in connection with the Mexican pension reform, things can get difficult later on: approaching elections seem to have made legislators more risk-averse or undermined government cohesion in a number of other cases (Mexico: pensions; France: labour market; Poland: farmers’ social insurance; Sweden: sickness insurance, Netherlands: rent deregulation). While there were two cases in which forthcoming elections were actually a spur to reform (Poland: pensions; United States: welfare reform), these were both instances in which specific circumstances prevailed: the evidence suggests that in each of these episodes, the electorate strongly favoured reform, and the major parties were fairly close to one another on the issue in question.

It may be significant that the cases do not point to a strong relationship between pension reform success and the timing of elections, whereas the labour-market reform episodes provided the strongest support for the hypothesis that it pays to act swiftly following an election. While the number of cases is too small to generalise this finding with any confidence, these differences do make intuitive sense: pension reforms tend to focus on questions of long-term sustainability and to be adopted only after extensive consultation. They thus take longer to prepare and adopt, and they are in any case unlikely to generate significant pay-offs within a single election cycle. Moreover, although cross-party co-operation on pension reform proved to be the exception rather than the rule, the main axis of conflict in most cases was not party-political. Labour-market reforms, by contrast, are often highly contentious but can, if adopted early enough in the cycle, begin to generate real benefits in time for the next election, so prospects of success appear to be better when governments act in the immediate post-election period.

![Figure 2.1. Reform mandates and reform scores](image-url)
While all the cases examined here concern reforms that were at least attempted, the electoral cycle may also affect governments’ readiness to try to reform in the first place. They may simply undertake fewer reform initiatives when elections are approaching. However, recent econometric evidence concerning such pre-election “reform slowdowns” is mixed. While it is clear that governments facing close re-election battles may not wish to launch potentially unpopular reforms too close to polling day, it is also the case that governments up for re-election will wish to be seen to perform on their earlier election promises (United States: welfare). If facing probable defeat, they may push through last-minute reforms that limit to some extent their successors’ freedom of action (Poland: pensions).

Other political conditions matter but appear to be less important

Less important but still relevant is the state of the political opposition. In several cases, governments were able to pursue contentious reforms when the principal opposition parties were divided or otherwise in disarray (France and Mexico: pensions; Italy: labour market; Australia: electricity). In several of the labour market episodes, resistance to reform was weakened as a result of divisions among or problems within the major trade unions (Italy18, Germany and Spain: labour market). However, a weak or divided opposition was clearly not a sufficient condition for success, since some reform attempts were thwarted even when opposition parties had been weakened by recent electoral defeats and internal conflicts (France: labour market; Italy: pensions). This finding is consistent with previous research elsewhere. A look at the balance between government and opposition also suggests that government “strength”, as measured by its legislative majority or lack thereof, is not always important: the more successful cases were roughly evenly split between reforms adopted by minority governments and those commanding parliamentary majorities. Likewise, half the reforms that were blocked were proposed by governments commanding sometimes heavy parliamentary or congressional majorities. While reform can require much more negotiation under a minority government, it is also less likely to be reversed in cases where the opposition has cooperated in its adoption. The rank correlations tend to reinforce the sense that such factors as the government’s command of a majority in parliament, the presence of a single-party government or a coalition with one strongly dominant party, and the strength of opposition parties are far from decisive: the relevant indicators were not significantly associated with reform outcomes and sometimes generated the wrong sign (Table A1.1).

The cases considered here provide little support for the notion that the political complexion of the government is an important factor affecting prospects for reform success overall. As noted above, pension reform has not generally been a clear-cut left-right issue, and political orientation probably plays a small role for reforms in product markets, since the interest groups involved are often affiliated to a range of political parties. However, the cases themselves suggest that there may be a “Nixon to China effect” in respect of labour-market reforms: successful labour-market reforms were more often than not begun under centre-left governments, even if they were sometimes adopted with the co-operation of, or continued under, the centre-right (Italy, Germany and Spain: labour market; United States: welfare reform). At first glance, this might seem to contradict earlier econometric work suggesting that governments of the left undertake fewer reforms, particularly in respect of labour markets. However, since the episodes here only concern reform attempts undertaken by governments, it is possible that governments of the left are both less likely to pursue labour-market reforms and more likely to succeed when they do pursue them. First, it may be that centre-left governments
tend to pursue labour-market reforms only when the need for reform is most urgent. Secondly, because left-wing governments generally have closer relations with organised labour, they may find it easier to reach agreement with unions on reform proposals. Thirdly, where centre-left governments opt for reform, the trade unions may have few allies to support their opposition. In two cases, centre-left governments pursued labour-market reforms after growing impatient with union resistance and did so with the support, or at least acquiescence, of other parties (Spain and Germany: labour market). This left the unions isolated.

The explanation for the “Nixon to China effect” appears to lie in information asymmetries about the relationship between policy instruments and outcomes: faced with a government reform proposal, voters cannot always be certain whether it is motivated by a desire to increase social welfare or by the government’s wish to pursue partisan objectives and serve the particular interests of its own core constituents. Given the existence of such asymmetries, substantial policy change in a given direction may be more easily sustained if implemented by parties that would appear ex ante to be ideologically opposed to it. Mistrust of ideologically driven policies seems to have been a factor in a number of the cases under study, where opponents of reform were able to use ideologically and politically loaded terms to frame proposed reforms in a way that would strengthen resistance to them, in part by making them appear far more radical than they were (United States: pensions; Netherlands: rent deregulation; Australia: electricity).

The involvement of different levels of government may shape the reform

Six of the cases reviewed have a significant “federal” dimension (United States: welfare reform; Germany: retail and labour market; Australia: electricity and water; Spain: retail). The US and Australian cases suggest that the capacity for inter-governmental co-operation in policy making can be an important framework condition for reforms with a significant federal dimension; if this capacity is weak, then strengthening it may have to be a part of the reform agenda, as it was in Australia in the early 1990s. Otherwise, the case studies do not point to clear, simple lessons about the interaction between inter-governmental policy making and reform. In three cases (Germany: retail; United States: welfare reform; Australia: electricity), sub-national governments exerted an important influence in both advancing and shaping the reforms eventually adopted. These cases also highlighted the potential for policy innovation and experimentation by sub-national governments to influence national-level reform debates. However, in a fourth case (Germany: labour market), the federal dimension proved a complicating factor when it came to reform implementation, and resistance from sub-national governments emerged as an important impediment to reform in two others (Spain: retail; Australia: water). The capacity of provincial and municipal governments to influence reform implementation also made it harder for the government to pursue rent reform in the Netherlands. While the central government controlled most policies concerned with the demand side of the market (chiefly rent subsidies and tax treatment of owner-occupiers), constraints on supply were largely the result of local authorities’ actions (zoning and other land-use decisions).

In some respects, the Australian water reform is the most paradoxical episode: much of the impulse for reform came from the states, but it was conflict among the states and state reluctance to cede authority over water policy to national-level bodies that largely checked the reform’s progress. In the two retail regulation cases (Germany and Spain), the politics of reform at sub-national level were actually rather different from the national level – but in opposite ways: in Germany, it was the Länder that pushed for reform and
the central government that was reluctant, whereas in Spain it has been the regions (the Comunidades Autónomas) that have resisted opening up the market. The issue, then, is not whether decentralisation aids reform or not, but rather when and where it is likely to facilitate reform. As the German and Spanish retail episodes suggest, this may be highly context-dependent.\textsuperscript{30}

**The economic context**

*The business cycle may not matter so much for the fate of reform attempts*

While many reforms were initiated against a backdrop of poor performance, the cases do not point to a strong across-the-board link between current growth performance and governments’ ability to secure adoption of reforms. Again, this result must be considered in light of the focus of this study on reform initiatives that governments have actually undertaken.\textsuperscript{31} What the case studies suggest is that, granted a government is prepared to put forward a reform proposal, its prospects for securing adoption of that proposal do not appear to be strongly linked to the business cycle: reforms were roughly as likely to be adopted or blocked in cyclical upswings as in downturns. The rank correlations tend to reinforce this impression: none of the exogenous economic factors tested was strongly linked with reform outcomes, except for the aggregate variable for “economic triggers”, the negative sign on which suggests that labour-market reforms are especially difficult in the midst of a severe downturn or crisis, and the fiscal consolidation variable, which is discussed below. The weak coefficients for the economic variables overall are consistent with the conclusion that macroeconomic performance, though probably important in determining whether and when reform initiatives are undertaken, has limited impact on governments’ success in adopting and implementing those initiatives that they do attempt.

That said, the case studies do suggest some lines of analysis concerning the relationship between growth performance and reform in different domains that merit further exploration. First, whereas labour- and product-market reforms were typically initiated in periods of poor growth performance, if not crisis, pension reforms were undertaken by governments in economic conditions ranging from severe recession to strong growth. Since the costs and benefits of pension reform will be felt over the very long term, cyclical fluctuations would not be expected to shape the politics of pension reform much. By contrast, labour- and product-market reforms are aimed at improving current performance. Secondly, although the adoption of labour-market reforms does not appear to have been determined by the cycle, implementation appears to have been easier when it coincided with a recovery (Spain: labour market; United States: welfare reform; Netherlands: disability insurance). This observation is consistent with previous econometric work, which suggests that the most promising time to reform is immediately after a recession.\textsuperscript{32} Germany’s Hartz reforms are instructive: the positive effects on employment were substantial but did not materialise until sometime after the reforms were adopted; the Hartz measures remained unpopular and their role in improving the employment situation was not widely understood.\textsuperscript{33}

**Fiscal problems can create pressure for reform and yet make reform more difficult**

Pressure on public finances, particularly when linked to some fairly binding constraint, like entry criteria for European Economic and Monetary Union (EMU) or constitutional restrictions on public debt, can provide an impetus to pursue structural reforms (Italy, Poland and Mexico: pensions; Poland: farmers’ social insurance).
However, many structural reforms — particularly pension reforms — involve up-front fiscal costs, while the benefits are realised only later. A severe fiscal squeeze may therefore make reform harder to adopt and implement. Indeed, one of the most robust findings to emerge from recent econometric work on the political economy of structural reform is that sound public finances tend to be associated with more reform.34 A weak fiscal position also makes it harder to find the resources that may be needed to provide adequate compensation or transition arrangements for those who stand to lose from structural reforms, and governments with limited political capital may opt to “spend” it on fiscal consolidation rather than structural reform. The case studies reviewed here tend to be congruent with these findings, as is clear from the negative sign on the (weakly) significant rank correlation for the fiscal consolidation variable (Table 1.2 and Figure 2.2). The econometric literature suggests that fewer reforms are likely to be initiated by governments in weak fiscal positions, so there may be a twofold effect of fiscal strain — reform attempts are less likely to be initiated and less likely to succeed.

Figure 2.2. Reform scores and fiscal consolidation

![Figure 2.2](image)

While the case studies provide no basis for generalising about whether governments undertake more or less reform with a healthier fiscal balance, they do suggest that the fiscal position can indeed affect prospects for success if a government undertakes a major reform initiative:

- In several cases, concessions that facilitated reform adoption were possible on account of prior improvements in the state of public finances (Mexico and Poland: pensions;35 Spain: labour market). In others, difficulties in winning adoption were linked directly to a weak or weakening fiscal position (United States: pensions; Italy: labour market).

- The urgency of the fiscal situation matters. If immediate fiscal concerns rather than longer-term structural objectives are uppermost in policy-makers’ minds, the design of the reform may be distorted by the need for rapid savings (Italy: pensions; Poland: farmers’ social insurance; Sweden: sickness insurance; United States: welfare reform).
Reforms driven by immediate fiscal needs may also prove harder to sustain once the fiscal pressure has eased, not least because they are often presented as necessary responses to a financial squeeze rather than desirable changes in structural policy settings (Poland: farmers' social insurance; Sweden: sickness insurance). The academic literature on the political economy of policy reform suggests that this is often a problem with crisis-induced reforms, whether they are prompted by fiscal strains or other pressures.36

The contrast between the reform of sickness insurance (SI) in Sweden in the early 1990s and the Dutch disability insurance reform of the early 2000s is instructive: fiscal considerations were significant factors in both reforms, but their impact on the reform process was not the same. The need for immediate fiscal savings was paramount in Sweden when the first SI reforms were adopted in 1991, and the reform was presented as a fiscal necessity. As the fiscal situation improved, there was growing pressure to reverse it. In the Netherlands, by contrast, fiscal pressures were not the principal drivers and the reform was debated more in terms of activation and labour-market objectives than fiscal ones. The more recent Swedish reforms to SI, which so far appear to have been very successful, are also consistent with this hypothesis: because they were adopted at a time when the fiscal position was fairly healthy, they did not impose such heavy up-front costs on beneficiaries as had the reforms of 1991, and they were designed with a focus on activation and the social costs of inactivity rather than fiscal concerns.

The link between structural reform and the state of public finances can be particularly visible in respect of reform of employment protection legislation (EPL). EPL and unemployment insurance (UI) may be seen as alternative ways of protecting workers against unemployment risk. On the whole, there appears to be a trade-off between them, inasmuch as countries with strict EPL tend to have less generous unemployment insurance systems and vice versa.37 One obvious reform path for those seeking to lower EPL and reduce labour-market dualism would therefore be to combine relaxation of EPL with an increase in UI coverage. This possibility was indeed discussed in Italy throughout the period covered by the case study, but pressure on public finances limited the ability of successive governments to pursue such a compromise. A similar approach might also have facilitated reform in Spain, but there, too, the budgetary situation was tight. However, one reason labour-market reform in Spain was easier in 1997 than in 1993-94 was that public finances were healthier in the latter case. While fiscal discipline was still paramount, to ensure smooth entry into the first wave of EMU, the government could afford some tax and other concessions to encourage the use of the new permanent contracts and reduce reliance on fixed-term and other non-standard contracts.

Labour-market performance affects the nature of labour-market reforms

The political economy literature suggests that, although high levels of unemployment tend to increase the pressure for labour-market reform, a sharp rise in unemployment increases the likelihood of a “two-tier reform”.38 Since the value of employment protection rises with unemployment, regular workers, who tend to be the most organised segment of the labour force, have even greater reason to resist any weakening of EPL during a cyclical downturn. At the same time, because EPL tends to impede job destruction as well as job creation, governments may hesitate to relax EPL protections in the midst of a recession, for fear of even greater job losses. Consequently, reforms in such circumstances tend to focus on labour-market “outsiders” – new entrants, those on irregular contracts, the unemployed and others on benefit. The labour-market cases under
review are largely consistent with this hypothesis (France, Italy and Germany: labour market; Sweden: sickness insurance; Netherlands: disability insurance).39

The exception to this generalisation is Spain: it is the only case in the study in which a government relaxed conditions for dismissing labour-market insiders at a time of rising unemployment.40 In 1994, with unemployment still rising, the government moved to restructure severance procedures in an effort to reduce costs and to expand and clarify the grounds for “justified” individual and collective dismissals. Severance pay regulations remained unchanged, but even this limited reform marked an important watershed.41 Moreover, it helped pave the way for further steps in 1997, including the introduction of a new indefinite contract with lower severance pay. While the case study identifies many different factors at work shaping the reform process, the most striking feature of the Spanish situation is that by 1993-94, labour market dualism had gone so far that the unemployed and those on non-standard contracts outnumbered regular workers, so there were strong electoral incentives to help more outsiders into secure employment, even if this came at a cost to employed insiders.42 The Spanish case suggests that reform at the margins of the labour market can indeed weaken insider power and create an opening for deeper reforms to follow, but the experiences of the countries under study, as well as other OECD members, show that this process is not easy, automatic or quick.

External pressures tend to be felt most directly in product markets

It can be difficult to assess the role of external factors in shaping reform processes using a case-study approach, as many of the forces at work, such as trade liberalisation, the construction of the EU Single Market or policy shifts in other economies, are powerful but gradual. They shape the entire context within which economic policy debates take place. Their impact is therefore difficult to pinpoint in respect of any given reform debate, although recent econometric work consistently points to the importance of such international influences, particularly with respect to product-market liberalisation.43 The evidence emerging from the case studies is consistent with these findings and, significantly, offers support for some of the causal mechanisms thought to underlie these relationships. On the whole, the influence of international factors was most visible in conjunction with product-market reforms. These factors included trade liberalisation and international competition (Australia: electricity; Germany: retail44), as well as international regulatory change, technological development and EU policies (Sweden: postal reform). Research and policy innovation elsewhere also prompted policy makers to look at new ways of doing things (Australia: electricity; Sweden: postal reform). In two pension-reform episodes, pressure from international financial markets played a role in spurring reform initiatives (Poland and Italy: pensions), and governments in a number of cases linked reform proposals to EMU or EU entry (Italy: pensions and labour market; Spain: labour market; Poland: farmers’ social insurance). EU influences were most evident in connection with relatively “hard” policy instruments, like EMU, EU accession and European Court of Justice judgments. Softer forms of policy co-ordination, like the Lisbon process, tended to be used instrumentally – they provided tools for policy makers trying to legitimate changes in product- or labour-market policies that they believed were desirable.

Timing, scope and sequencing

The case studies suggest a number of issues to explore with respect to the timing and pace of reform, quite apart from the issues discussed above in connection with the
electoral cycle. A substantial literature focuses on the sequencing or “bundling” of reforms, particularly in developing or transition countries. Since bundling has tended to be more common in cases of systemic transformation, when “big bang” reform packages are adopted (e.g. the start of the post-communist transition), or in response to severe crises, most of the reforms covered by this study have not been adopted as part of a single large package. They tend instead to be sequenced, an issue on which several of the case studies cast light. Where bundling occurs, it usually involves multiple, closely related reform measures in a single domain rather than the combination of reforms across a number of distinct policy areas.

A “climate of reform” can facilitate progress on specific issues, particularly in product markets

The case studies suggest that individual reforms can be easier to pursue where they form part of a larger shift in structural policy settings; such reforms run with rather than against the grain of economic policy (Sweden: postal reform; United States: welfare reform; Australia: electricity and water; Netherlands: disability insurance; France, Mexico and Poland: pensions). Thus, when Swedish postal reform emerged on the political agenda, the country was already experiencing a broad shift towards opening up network sectors. Both power and water reforms in Australia were launched in the context of, and largely shaped by, a broader microeconomic reform agenda, which was ultimately given formal institutional shape in the National Competition Policy. The Netherlands tackled disability insurance reform at a time when a number of other labour-market and social-insurance reforms were on the agenda. The very high and significant score for the interaction between a given product-market reform and the overall “PMR reform environment”, as measured by the change in the OECD indicators of product-market regulation (PMR) for seven non-manufacturing sectors around the time of the reform, reinforces this impression: it suggests that individual product-market reforms are more likely to be accepted as part of a wide-ranging reform process.

It may be difficult to pursue simultaneous pension and labour-market reforms

Policy interactions across domains appear to be weak, with one exception: the correlation coefficient shows a significant and negative sign for the relationship between labour-market reform outcomes and the indicator of on-going changes in pension policy. In short, it would appear that major pension reforms may make successful labour-market reforms significantly harder. Given the labour-market implications of many retirement reforms, major initiatives in both domains at once might be perceived as hitting workers with a “double-whammy”: unions, with interests in both areas, would find it hard to accept major reforms in both, and governments might well accept that the political price of pursuing them together would be too high. This makes the Italian experience of the mid-1990s all the more remarkable, since the country did manage to pursue reforms in both domains, albeit only after making significant concessions to the unions in respect of both.

Events may open “windows of opportunity”...

Even in the absence of a recent electoral mandate, one-off events, such as crises or scandals, may expose weaknesses in the status quo and create an opportunity for reform. They may even result in a government having a clear mandate for reform in the absence of fresh elections. Crises or other “action-forcing” events can sharpen awareness of the need for change, shake up established interest coalitions or induce agents to accept
reforms despite high levels of uncertainty about the post-reform environment (Italy: pensions; Poland: farmers’ social insurance; Sweden: sickness insurance and postal reform; Spain: labour market). Alternatively, there may be developments that do not force the government to act but that reduce the impediments to reform by weakening its opponents or shaking up established interest coalitions (Italy and Germany: labour market; Mexico: pension reform and labour law; Australia: electricity). However, such reforms, if adopted, can prove difficult to sustain once the extraordinary circumstances pass, unless they are based on longer-lasting shifts in agents’ understanding of what is feasible or desirable (Germany: labour market; Poland: farmers’ social insurance). Significantly, exogenous events appear to have had least impact in respect of retirement reforms, which is perhaps what would be expected: given that they address the long-term sustainability of complex policy regimes that encompass most of the population, they tend to require long preparation and extensive consultation, making them less likely to be driven by current political or economic developments. Indeed, both the crisis-induced pension reforms in this study were thwarted (Italy: pensions; Poland: farmers’ social insurance). The correlation coefficients for the “exogenous triggers” variable provide some support for these impressions: the coefficient is both strong and significant for reforms in product markets (Table 1.2).

... but more haste can make for less speed

While it is important that policy makers seize reform opportunities when these arise, it is critical to bear in mind that successful reforms, particularly pension reforms, often have relatively long gestation times, involving a considerable amount of careful study and consultation (France, Poland and Mexico: pensions; Italy: labour market; United States: welfare reform; Netherlands: disability insurance; Australia: electricity). On average, the formulation and adoption of the reforms examined in the more successful cases took somewhat over two years – and this figure does not include the “pre-work” done in the many episodes in which problems and proposals had been debated and studied for years before the authorities set to work framing specific reforms. By contrast, many less successful reform attempts were launched without adequate preparation (Sweden: sickness insurance; France: labour market; Italy: pensions). While governments should be ready to use political “windows of opportunity” when they open up, this may create problems if it leads to excessive haste. This applies to implementation as well as adoption: overly ambitious implementation timetables can lead to otherwise avoidable problems (Poland: pensions; Germany: labour market). Slippage with respect to implementation schedules may thus prove necessary at times, in order to avoid potentially costly implementation failures (Mexico: pensions). If haste is often a harbinger of failure, however, deliberation is by no means a guarantee of success (Figure 2.3). A number of thwarted reforms covered in the episodes were the focus of long and careful study and consultation processes, which in the end produced little or no reform (Germany: retail hours; Netherlands: rent deregulation; Poland: farmers’ social insurance; Mexico: labour law). As expected, the rank correlation for preparation time across the 20 cases is highly significant; at the level of policy domains, it is also, albeit somewhat less, significant for pension reforms, but not for labour or product markets, reflecting long preparation and negotiation phases that precede adoption of most successful pension reforms.
Policy regimes may be more or less “ripe” for reform

Reforming early, though desirable in principle, may be very difficult in practice. Successful reforms of established policy regimes are often preceded by an “erosion” of the status quo, involving (1) a widely shared conception that the policies and institutions in place are failing and (2) a series of incremental decisions that have already begun to weaken the existing policy regime. While things need not reach crisis point, radical reforms often follow piecemeal changes to the old system that are found to be inadequate or even make things worse (Poland: pensions; Germany: retail hours and Hartz; United States: welfare reform; Netherlands: disability insurance; Italy and Spain: labour market). Reform is likely to be more difficult where the existing arrangements are well institutionalised and popular, and where there appears to be no danger of imminent breakdown if reforms are not implemented (Poland: farmers’ social insurance; Sweden: sickness insurance; United States: pensions; Mexico: labour law; Netherlands: rent deregulation). In such circumstances, opponents of reform may not even need to present any alternative proposals of their own; they can simply defend the status quo. The “reform ripeness” variable used in the rank correlations is thus defined in terms of previous reform attempts, even if these were partial, thwarted or reversed: successive attempts to reform a regime suggest a growing awareness of the need for change. The results suggest that reforms are indeed much more likely to advance where previous reforms have already been undertaken (Figure 2.4). “Ripeness” appears to be particularly important in respect of labour-market and pension reforms, whereas it has little impact on product-market reforms. This may reflect the fact that reforms of pension and labour-market institutions affect large parts of the electorate directly and alter arrangements which in some cases are seen as core welfare-state institutions. In the case of product-market reforms, it seems to be the broader policy context that matters, as was...
suggested by the policy interaction variables discussed above: a possible message may be that a government may well open up a sector at the first attempt, but that it is more likely to do so if it has been opening up other sectors as well.

![Figure 2.4. Reform scores and reform history](image)

The question of “reform ripeness” concerns not only the seriousness of the problem to be addressed but also the extent of awareness of the costs of the status quo and of agreement on the need for change – it is a question of communication (see below), as well as economic analysis, and is therefore to some extent endogenous to the reform process. The conclusion to be drawn, therefore, is not that reformers should leave “unripe” policy regimes alone, but that persuading stakeholders and the public of the costs of non-reform may be the first challenge in such cases. Indeed, the Spearman rank correlations show a statistically significant positive link between “reform ripeness” and “communication of the costs of the status quo”, apparently reflecting the degree to which even partial or unsuccessful reform attempts can focus attention on a problem. A further important implication of this argument concerning “ripeness for reform” is that blocked, reversed or very limited early reforms need not be seen as failures: they may play a role in undermining the status quo and setting the stage for a more successful attempt later on. All of the more successful pension and labour-market reforms examined in this study followed earlier setbacks and many less successful reform attempts in all three domains can now be seen to have helped set the stage for subsequent, sometimes far-reaching reform initiatives (Australia: water; Sweden: sickness insurance; Italy: pensions; Germany: retail).

*It often makes sense to “take what you can get”*…

A number of cases highlight the ways in which tackling the easiest issues first can generate a momentum for reform that facilitates further change. The initial steps may create pressure for further reform by changing agents’ expectations and/or activating new constituencies (France, Poland, Mexico: pensions; Germany: retail; Sweden: sickness insurance and postal reform; Australia: electricity; Spain: labour market; Netherlands: disability insurance). If early measures lead to an expectation of further reform, agents
begin to adjust their behaviour in anticipation, and this itself may reduce resistance to subsequent measures. Indeed, changed expectations and the activation of new interests are among the ways in which reform setbacks sometimes lay the basis for subsequent success (Italy: pensions; Sweden: sickness insurance; Netherlands: disability insurance). As the reform process unfolds, the privileges and rents retained by those untouched by the first reform measures may become ever more apparent and thus more difficult to defend. Hence, the typical pension reform sequence in many countries begins with (relatively less privileged) private-sector employees before moving to the main public sector scheme and then to the more specific schemes that exist in the most privileged parts of the public sector. Likewise, as noted above, labour-market reform often begins with reforms that target labour-market outsiders and leave insiders protected – a classic “take what you can get” strategy.

In product markets, the process is somewhat more complicated, but broadly speaking, sectors that have already been opened up to competition (whether via international opening or the deregulation of domestic product markets) often support efficiency-enhancing reforms in the sectors on which they rely for inputs and services (Australia: electricity; Sweden: postal reform). The literature suggests that this is a relatively common pattern in the diffusion of product-market reforms, with sectors under competitive pressure gaining from reforms “upstream”. In the present study, it is noteworthy that such pressures were lacking in respect of the product-market reforms that were thwarted (Germany and Spain: retail; Netherlands: rent deregulation; Australia: water). In three instances, the regulated sectors were oriented towards households, so there were no “downstream sectors” to mobilise. Households in these cases either found reform unattractive (rent) or simply saw too little immediate benefit from reform for it to be electorally salient (retail). In the case of Australian water, market-oriented reforms threatened the interests of an irrigated agricultural sector which had grown up largely in response to past water policies.

... but such a strategy also has risks

A strategy focused on starting with “low-hanging fruit” has its pitfalls. First, if very modest initial proposals are seen as harbingers of deeper reforms to come, some agents may adopt an intransigent position in opposition to even relatively minor measures in order to forestall the risk that the process will subsequently go further (Italy: labour market; United States: pensions; Netherlands: rent deregulation). Secondly, it can be difficult to ensure that the first wave of reform does indeed lead to further reform. The value of the rents enjoyed by exempt groups may increase as a result of reforms that target others, thus increasing the unreformed insiders’ incentives to protect their privileges. The problem of deepening labour-market dualism illustrates how hard it is to move from flexibility at the margin to a reform of the “core” of the labour market. This stands in contrast to the pension reform path described above. In all likelihood, the difference between the typical pension and labour-market reform sequences reflects the relative sizes of the unreformed groups. Pension reforms tend to begin with large but poorly organised and less privileged groups, before proceeding to address the various more privileged regimes. As reform proceeds, the unreformed minorities that remain are isolated and find it ever harder to defend their anomalous positions. In labour markets, the situation is often reversed: the initial outsider-oriented labour-market reforms typically target minorities, leaving the majority of workers untouched. The unreformed sector is thus larger than the reformed one.
Communication, consultation and leadership

There have long been debates about whether and when reformers should mask their intentions from the public (“reform by stealth”) or, alternatively, maximise the transparency of the policy process in an effort to win stakeholder and public support for reform. The case studies point to the desirability of the latter strategy, not least in view of the importance of electoral mandates discussed above. In addition to facilitating reform adoption, such an approach can improve the quality of the reform design and reduce the likelihood that the reform will be reversed if the political conjuncture changes. Effective communication, consultation and research can all contribute to a consensus for reforms; reform adoption may take longer when such an approach is employed, but the quality of policy is likely to be higher and subsequent reform reversals less common.

Effective communication and consultation efforts pay off

In the more successful episodes, governments made considerable efforts from an early stage to explain and sell the reforms to stakeholder interests and, in most cases, the public (France, Italy, Poland, Mexico: pensions; Germany and Spain: labour market; Netherlands: disability insurance). This process served not only to generate support for reform but also, in many instances, to enable the authorities to identify potential problems and, where necessary, to improve reform design in response to feedback from stakeholders. By contrast, communications strategies were weak in connection with some reforms that were blocked (France: labour market; Italy: pensions; Poland: farmers’ social insurance) or that provoked especially strong opposition (Germany: labour market). Effective communication was closely related to government cohesion in many case studies, since divided governments tended to send out mixed messages (France: labour market; Italy: pensions; United States: pensions; Netherlands: rent deregulation). Targeting the right audiences was also important: in the case of the 1993 labour-market reform in France, the adoption phase of the reform involved extensive discussions with the unions, but the real problem was a failure to explain the contrat d’insertion professionnelle to the young people who would be most directly affected by it. A similar problem arose in Germany when the benefit reforms that accompanied Hartz IV were being implemented. In cases where proposed reforms were likely to have a very broad impact across society (e.g. reform of pensions and entitlements, labour contracts or rent regulation), effective communication and consultation seems to have been essential: unilateralism has not proved a promising strategy for pension reform, in particular. By contrast, the case study evidence suggests that public communication efforts may be less important in respect of relatively technical or narrowly focused sectoral reforms (Germany and Spain: retail; Sweden: postal reform; Australia: electricity); in such cases, public understanding is unlikely to be decisive.

Communicating the goals of reform can be particularly important, since lack of clarity about objectives may increase uncertainty, allowing opponents of reform to (often misleadingly) define the likely endpoint of the process in the public mind (Netherlands: rent deregulation; Australia: water; Germany: labour market). Clarity about objectives may also be an aid to reform design: where the aims of reform are not clear and coherent, it is difficult to ensure that the policies chosen will meet them: in several cases, the complexity of the reforms’ objectives, which sometimes evolved in the course of the reform process, made it difficult to map the chosen instruments onto the desired outcomes (France: labour market; Australia: water; Netherlands: rent deregulation). Moreover, there is some evidence that better communications reduce the likelihood of reform.
reversals. Reforms imposed on the basis of force majeure conditions rather than intellectual persuasion may more easily be reversed once those conditions disappear.\footnote{76} The reversal of Swedish sickness insurance reforms in the 1990s illustrates this phenomenon, as does the blocking of Poland’s reform of the farmers’ social security system when fiscal pressures eased in 2004-05.

These conclusions stand in contrast to some of the arguments advanced in the literature on the political economy of reform, which hold that reformers should avoid declaring their intentions to the public \textit{ex ante}, since this will mobilise opposition to reforms that are expected to be painful; instead, they should aim for speed and stealth, seeking to advance reforms before reform opponents can react.\footnote{77} On this view, governments should not be afraid to get ahead of public opinion; rather than feeling constrained by the need to create a social consensus, they should offer leadership, confident that a consensus will form in support of a successful reform.\footnote{78} As noted above, however, such strategies typically prevail when reforms are adopted in crisis conditions and when they bear fruit quickly. The only such episode in the present study concerns the fiscal crisis in the Australian state of Victoria, which played a key role in pushing market reform of the electricity sector forward. Yet even in Victoria, the government’s early moves had to be ratified at a subsequent general election, since power-sector reform was too complex and time-consuming to impose rapidly and present to voters as a fait accompli.

\textit{Communicating the costs of non-reform can be critical}

Opposition to reform is often based on arguments about the balance – and distribution – of costs and benefits of policy change.\footnote{79} It can be difficult to make a case for reform to the public, given that structural reforms often involve substantial up-front costs, with the benefits coming later.\footnote{80} Even in the absence of any uncertainty or conflict about the allocation of costs and benefits, rational agents may question the need for reform, given that they will discount the future. Moreover, the link between a reform and its benefits can be far more difficult to establish, even \textit{ex post}, than the connection between a reform and its costs.\footnote{81} Distributional issues aside (these are addressed below), would-be reformers often face a challenge in simply persuading stakeholders and the public that the overall assessment of costs and benefits favours reform. The case study evidence suggests that this is largely because the costs of the status quo are often poorly understood. The depth of this challenge depends on the policy in question: since at least the early 1990s, it has become easier to convince the public of the need for – and, perhaps, the inevitability of – the reform of general pension systems (France, Italy, Poland, Mexico: pensions), even if the urgency of reform was sometimes questioned (United States: pensions). This reflects in no small measure the impact of public discussion of research on the implications of population ageing, in particular, for traditional public pension systems. Labour-market reform issues are less well understood – particularly the costs that high levels of EPL impose.\footnote{82} However, voters know when labour-market performance is poor, and the labour-market reform cases show evidence that reform progressed as the costs of the status quo came to be better understood (Netherlands: disability insurance; United States: welfare reform; Germany\footnote{83} and Spain: labour market). Reform was more difficult where these were not appreciated (France: labour market; Sweden: sickness insurance; Mexico: labour law). Indeed, the Spearman correlations confirm the importance of this factor, especially in respect of labour-market reforms (Table 1.2). By contrast, it was far more difficult to communicate the costs of the status quo in respect of product-market reforms (Netherlands: rent deregulation;
Australia: water; Spain and Germany: retail; Sweden: postal reform), in large part because these were often opportunity costs and therefore “politically invisible”.

The major exception to this rule – the Australian electricity reform – was characterised by a determined effort to quantify the costs of the status quo and the potential benefits of reform and to communicate these to stakeholders and the public. Comprehensive and transparent explanations of the short- and long-run costs and benefits of reform, underpinned by solid research, can be crucial elements of successful structural reforms. This is particularly important when the costs of the existing regime are opportunity costs. Often, it is fairly clear who will pay the price for a reform – which firms are likely to come under pressure and which jobs may be at risk – whereas it is not at all obvious who is paying for the status quo: it is difficult to identify firms that have never entered the market, sectors that have not developed or the workers whom they would have employed. Effective communication of the costs of non-reform can thus contribute to making a policy regime ripe for reform. It is also closely linked to a government’s attempt to win a popular mandate for reform: the rank correlations suggest that these two variables are closely and significantly related in the case studies. Effective communication of the implications of non-reform can also change the “reference case” that agents use when evaluating their options. One reason it can be difficult to identify potential “winners” from a reform is that people tend to evaluate their expected pay-offs under reform against the status quo. Where the existing policy is unsustainable (France, Italy, Poland, Mexico: pensions; Australia: water), it is critical that agents assess their position under a future reform against an unreformed future rather than the status quo ante.

Although good research is no “fix” for politics, it helps

The quality of the analysis underlying a reform can affect prospects for both adoption and implementation, as well as the quality of the policy itself. Of course, it is clear from both the case studies and the wider literature on the role of economic analysis in policy making that the political context will influence the reception of any particular line of analysis by the public or policy elites at any given moment: politically unwelcome research findings may at first be challenged or simply rejected by powerful institutions and interest groups. However, the evidence suggests that the influence of policy-oriented research, however diffuse and indirect, can be quite powerful over time, as it gradually reshapes the consensus concerning a policy regime. In the case studies, this is clearest in respect of reforms that were debated over a relatively long period (United States: welfare reform; Australia: water and electricity; Netherlands: disability insurance; Sweden: sickness insurance). The reception of economic analysis also depends on the source: research presented by an authoritative, non-partisan institution that commands trust across the political spectrum can have a far greater impact. Building such institutions can take time, as their effectiveness depends greatly on their reputation, which may take a considerable period to establish, but they can in the long run make a significant contribution to the quality of both policy making and public debate.

It is in this context that international institutions can have a role to play. The evidence of the case studies concerning the influence of the OECD, in particular, is generally consistent with the findings of other research in this area. There is little evidence of a direct impact of OECD advice on policy making in the case studies. This is not surprising, given that most OECD recommendations are neither binding on member states nor backed up by financial or other incentives. The evidence does suggest, however, that the case for reform is strengthened by the availability of internationally comparable data
and analysis. Often OECD influence was mediated via domestic institutions, such as the CPB Netherlands Bureau for Economic Policy Analysis, the Australian Productivity Commission or the various economic think-tanks in Germany. These bodies also feed their own research and analyses into policy debates and, indeed, into the work of international institutions. In the early 1990s, the OECD Jobs Strategy, for example, did become a point of reference in labour-market reform debates in some countries. While OECD recommendations and data were cited rather frequently in some debates, the clearest impact of OECD work was observed when countries were able to see their performance or policy regimes in comparative context: benchmarking often signals to electorates or elites that institutions or situations that they may have come to regard as normal may in fact be quite unusual by international standards. Finally, Armingeon (2004) highlights the role of the OECD in creating and sustaining “epistemic communities” – transnational networks of experts who then influence policy debates within their respective countries by framing both policy problems and potential solutions for decision-makers.

**Major reforms may benefit from institutions that foster consensus**

A number of the case studies point to the potential utility of credible “reform institutions”, like Australia’s Productivity Commission or the Dutch CPB. More specialised bodies like the French Conseil d’orientation des retraites or Poland’s Office of the Government Plenipotentiary for Social Security Reform (BPR) can play a key role in specific domains. Such institutions can serve as fora for study and negotiation and help to “de-politicise” sensitive reform issues (this is particularly common in the case of pension reform, because it directly affects virtually the whole population). Even if they do not actually design the reforms or resolve distributional conflicts, they can – if they are seen to be credible and reasonably impartial – make progress easier by fostering consensus on certain basic issues, including the costs and benefits of both the status quo and reform (France and Poland: pensions; Germany: labour market; Netherlands: disability insurance; Australia: electricity and water). They can also improve the quality of policy making, by providing research and analysis to inform the process, as well as a forum in which issues can be debated openly and research findings scrutinised. Although member countries may rely at different times on both permanent bodies and special commissions and committees, the case studies suggest that such institutions may carry more weight if they are permanent, independent public bodies rather than ad hoc commissions or working groups. The latter can have a significant impact but tend to disband soon after they report their findings. Moreover, permanent “reform institutions” with a fairly broad remit may be less susceptible to capture by specific interests than those that are specialised in very specific areas. Whether permanent or ad hoc, such bodies must be seen as impartial: if they are highly politicised or seen as an attempt to circumvent key actors in the policy process, they may actually intensify opposition to reform, since they will not be seen as impartial (United States: pensions; Mexico: labour law). Germany’s Hartz Commission was arguably something of an exception to this rule, but it was formed when the traditional participants in labour-market policy consultations had been weakened by a scandal at the Federal Labour Office. Moreover, after the Hartz report was issued, the relevant legislation went through “normal” consultative policy processes.

It is striking that in the Australian and Dutch cases, the key reform institutions were deeply involved in the more successful reforms (electricity and disability insurance, respectively) but played a limited role in the less successful episodes (water and rent...
deregulation). Since then, the Productivity Commission has done much more work on water, and the CPB and other bodies in the Netherlands on housing, as governments in those countries have sought to renew reform processes in those areas.95 Research in many instances initially tends to follow rather than lead early policy reform debates.96 Thus, early reforms to a regime are often undertaken with little basis in research, but they then trigger a large body of subsequent research that begins to shape the debate over subsequent steps (Sweden: sickness insurance; Australia: water; Netherlands: rent deregulation; United States: welfare reform).

Research capacity and reform institutions can help provide needed feedback and adjustment

A successful reform need not be perfect, particularly when governments are engaged in policy innovation and the reform is at or near the international frontier. As is clear from many of the case studies, reform often requires “learning by doing”, so it makes sense to design mechanisms for ex post assessment of outcomes; adequate feedback and adjustment can be critical.97 This can be done in a variety of ways, but strong analytical and research capacities generally need to be part of the process; new institutions created with a mandate to implement and develop a reform, like the Swedish postal regulator or the various institutions created to run the Australian electricity market, may also play an important role in reacting to implementation surprises and making needed adjustments as the reform unfolds.

Government cohesion is of crucial importance

Perhaps the strongest single finding to emerge from the case studies concerns the unity of the government. In almost all cases in which there was public conflict within the government or the governing party over a reform, the reform was ultimately thwarted (France: labour market; Italy: pensions; Germany: retail;98 Poland: farmers’ social insurance; United States: pensions;99 Sweden: sickness insurance; Netherlands: rent deregulation; Spain: retail regulation100). This held true across all three policy domains under study but it was especially strong in respect of labour-market and pension reforms (Table 1.2). There were only two exceptions to this rule among the 20 cases examined: in Italy, public differences within the governing coalition weakened the government’s position vis-à-vis opponents of the Biagi labour-market reform law but did not prevent its adoption, and the PRWORA welfare reform in the United States emerged as a compromise negotiated between a Republican congressional majority and a Democratic White House, over the objections of many of the administration’s own welfare policy makers. However, the US case is highly peculiar: both parties were committed to a reform reflecting certain basic principles, and the splits in the administration arose in the context of negotiation over how those principles were to be applied rather than a conflict between proponents and opponents of reform. Other studies also suggest that government cohesion comes close to being a necessary – though by no means sufficient – condition for successful reform.101

When governments are seen to be less than united, opponents of reform are quick to exploit the situation, whether that disunity reflects conflicts within the government unrelated to the reform – e.g. rivalry among parties in coalition – or differences within the government over the reform proposal itself. At times, lack of unity can also contribute to the incoherence of the measures proposed, as compromises adopted to appease various stakeholders within the government camp complicate the reform (France: labour market; Netherlands: rent deregulation). The rank correlations confirm that government cohesion
is not unrelated to the question of electoral mandates and the effectiveness of the government’s communication efforts. Logically, it should be easier to keep a government united in support of a reform it has already promised to the voters, and in the cases under review, none of the governments that were publicly divided over reform could credibly claim an electoral mandate for the measures they were proposing. Likewise, if the government is publicly split over policy, it is bound to send out mixed messages to voters and other stakeholders.

“Ownership” of reforms is also important

As the foregoing discussion of government cohesion and clarity implies, clear “ownership” of reform initiatives can be important. This is one of the hypotheses finding strongest support in the case studies. Many of the more successful reforms under study were driven by clearly identifiable individuals or institutions prepared to take ownership of them and having an interest in seeing them succeed (France, Italy, Poland, Mexico: pensions; Italy, Germany: labour market; Netherlands: disability insurance; Sweden: postal reform). By contrast, reforms were less likely to be pursued where such leadership was lacking (France: labour market; Mexico: labour law; Germany: retail hours; Spain: retail regulation). Reform “orphans” rarely succeed. This is true even after the initial adoption of a reform: sustaining reforms can be difficult if the policy entrepreneurs who pushed them forward leave the political arena before they are consolidated (Germany: labour market). There is, of course, a risk of ex post bias here: there will be plenty of actors ready to claim credit for a successful reform, while agents involved in the policy process will try to disown failed initiatives. In the cases cited above, however, the evidence is fairly strong that the presence or absence of clear ownership was apparent ex ante. This observation is very much in line with other research highlighting the importance of leadership from executives with a strong commitment to reform and a readiness to take risks. Yet the presence of such leadership is not simply a question of personality or good political luck: “policy entrepreneurs” do sometimes emerge independently of any government strategy or commitment, but they are more often “made” by governments, e.g. by the designation of ministers, officials or institutions with clear mandates to promote specific reforms. Even energetic entrepreneurs may find their initiatives blocked, however (Poland: farmers’ social insurance; United States: pensions; Netherlands: rent deregulation).

Dealing with actual or potential opponents of reform

Even if there is a clear consensus that a reform will generate net benefits for the economy as a whole, questions concerning the allocation of costs and benefits are often unavoidable, owing to differences in agents’ interests and endowments, the presence in many situations of potential “stranded assets”, and uncertainty about individual benefits. Often, the costs of reform are not only incurred up-front, they are also concentrated on specific groups, whereas the benefits not only materialise later but are also both more diffuse and less predictably allocated. The literature on structural reform devotes a great deal of attention to the question of when and how those who lose out as a result of reform might be compensated, whether by exempting them from the reform, at least for some period, or via some sort of alternative compensation. Failure to compensate may reinforce opposition to reform, but excessive compensation may be costly or may simply blunt the effects of the reform itself. In a repeated game it may also reinforce opposition to future reforms. The most common compensation strategies involve “grandfathering” rents and long transition periods, both of which are discussed
below. Concessions in the form of “side payments”, such as adopting policies in other domains that might offset the cost of reform for some groups, were rarely important, although the Spanish labour-market episode was an exception. In return for accepting some reduction of EPL for regular workers, the trade unions in 1997 secured changes to the system of collective bargaining that they wanted; while these changes had less impact than anticipated, the political exchange involved was an important part of the overall reform bargain.

Concertation can be useful but is no substitute for government leadership

Progress may sometimes be facilitated by intensive discussions involving the government and the social partners in a highly formalised process (Italy: pensions and labour market; France: pensions; Poland: pensions; Germany and Spain: labour market; Netherlands: disability insurance). However, firmness of purpose on the part of the government also seems to be a critical element of success in such situations. The rank correlations suggest that government leadership – as proxied by an assessment of evidence of the government’s readiness (or lack thereof) to act unilaterally or to sanction non-co-operation on the part of the social partners – is particularly important in respect of labour- and product-market reforms (Table 1.2). Consultation need not mean allowing stakeholders to be veto players, and several cases suggest that a “concertationist” approach is unlikely to succeed unless the government is in a position to reward (sanction) co-operation (non-co-operation) by the social partners or can make a credible threat to proceed unilaterally if concertation fails (Italy, Mexico, Germany and Spain: labour markets; Italy: pensions). Where the government is too weak to lead or is unwilling to do so, the social partners have little incentive to make concessions. By contrast, corporatist arrangements can work well where the government is prepared to lead. They can be particularly important when negotiating reforms to core welfare-state institutions, which may call into question explicit or implicit social pacts that are regarded by important segments of society as fundamental features of the social order.

In several of the cases under examination, governments paid a price for eschewing concertation in respect of reforms that the social partners believed should properly be the subject of social bargaining (Italy: pensions and labour market; Spain: labour market; Poland: farmers’ social security). Concertationist practices, though sometimes very useful, may lead the agents to conclude that sensitive reforms must be the subject of corporatist bargaining, so that any attempt at reform outside such a framework will meet resistance on procedural as well as substantive grounds: even fairly modest reforms may meet stiff opposition if they are perceived as being “imposed”. There is some reason to think that such “unilateral” reforms are more likely to be reversed than those negotiated with the social partners. This can have implications for implementation, since the perception that a reform is likely to be reversed reduces individuals’ incentives to adapt to the new conditions, and those threatened by a reform may opt to resist implementation in the hope that the status quo ante will soon be restored. Moreover, while a unilateral approach may enable the government to adopt bolder reforms than the social partners would accept, the social partners’ active co-operation may still be required for effective implementation. The degree to which stakeholder co-operation is needed for implementation of a reform – or to achieve progress on other reforms – may thus be a factor in determining whether and to what extent the government should accept the constraints of corporatist bargaining when trying to win adoption of reforms. The other factor that must be borne in mind is that concertation may distort the content of reforms, since the parties represented at the table with the government represent specific segments
Employers’ bodies represent incumbent firms (and, in many cases, they are more representative of some segments of the population of firms than others), while unions tend to be most representative of older cohorts of industrial workers. This leaves the burden of representing “outsiders” very much on the government.

**Involving potential opponents in the post-reform system may facilitate adoption of reforms – but at a price**

Where particular opponents of reform are positioned as *de facto* “veto players”, capable of blocking its adoption or impeding its implementation, it may be possible to win their assent by giving them a role in the new system. Successful pension reforms, in particular, frequently offer opportunities for trade unions in the reformed system, for example in the administration or running of pension funds (*France, Italy, Poland and Mexico: pensions*). It is less common in respect of labour-market reforms (Table 1.2), though here, too, it may play a role (*Italy, Germany and Spain: labour market*). However, this strategy has its risks: it may limit the scope of the reforms that can be adopted and it means that implementation of important elements of the reform may need to rely on the co-operation of parties who are not fully committed to it.

**Acquired rights are typically “grandfathered”**

Where acquired rights are concerned (particularly pension or benefit entitlements), large groups may need to be wholly or partially exempted from the reform in order to secure its adoption. Thus, current pensioners are rarely affected by pension reforms and older cohorts of workers usually experience only minimal change (*France, Italy, Poland, Mexico and United States: pensions*). In both Italy and France, the transition periods for pension reforms were long enough to ensure that the bulk of the voting age population when the reforms were adopted would escape their full effects. This reflects in part the fact that trade unions and political parties tend to be oriented towards the interests of older cohorts, but there are also good economic arguments for significant transition periods when it comes to pension reform. Sharp swings in pension policy can be costly to contributors making career/savings choices that will pay off only in the long term. The microeconomic benefits of pension reform depend in part on the clarity and stability of the link between contributions and benefits. The costs of pension reform thus tend to be borne chiefly by younger cohorts. Since it is they who will lose out if the system proves unsustainable over the long term and since they have longer to adjust their private arrangements for financing retirement, this tends to be politically acceptable. Indeed, the lack of overt inter-generational conflict is one of the most striking features of the politics of pension reform: there is no evidence of it in any of the case studies, and recent academic work on the political economy of pension reform in a large number of countries finds little evidence of it elsewhere.

In respect of labour-market reforms, too, large numbers of incumbent regular employees may also be effectively exempted, *e.g.*, by limiting reforms to new contracts or specific types of employment relationship (*Italy, Germany and Spain: labour market; Netherlands: disability insurance*). The problem of acquired rights is most straightforward where labour contracts are concerned. It is less clear when unemployment or other benefits are being reformed, since the question of whether existing contributors or beneficiaries in the system really have some sort of “acquired rights” that must be respected depends on how the system is organised and conceptualised. Thus, some benefit reforms directly affected the position of those already on benefit (*United States: welfare reform; Sweden: sickness insurance*), while in other episodes, benefit recipients
were shielded from at least some of the changes (Germany: labour market; Netherlands: disability insurance). This reflects not only the way social insurances are organised but also the organisational and other resources of the groups that stand to lose.

Product-market rents are usually phased out gradually

Compensation in the form of exemptions and “grandfather” clauses is likely to be more difficult in cases of market opening, since grandfathering rents (i.e. exempting incumbents indefinitely) may risk worsening rather than alleviating market distortions, if the result is that new entrants and incumbents are subject to different rules. Thus, the product-market reforms examined in this study typically applied to incumbents as well as new entrants but included transition arrangements designed to ensure that the producer or consumer interests affected did not experience abrupt changes in economic conditions (Spain: retail; Australia: electricity and water; Sweden: postal reform; Netherlands: rent deregulation). In some cases, such arrangements were insufficient to allay opposition to the reform. The most difficult problems tend to be those that arise when the rents resulting from anti-competitive regulatory policies are capitalised into the prices of assets. In the two retail cases, for example, investments undertaken on the basis of past policies would have earned lower returns, or even become unprofitable, in the event of reform. The Australian water episode highlighted a related but distinct problem – that of “stranded assets”: water reform threatened to make it impossible to earn an economic return on many investments in rural areas, particularly in irrigated agriculture. These investments had been made on the basis of the previous policy regime, which had strongly encouraged irrigation-based development and rural settlement. More generally, employer-employee relationships usually involve some element of relationship-specific investment that is lost when reform leads to a separation, so “stranded assets” problems can affect labour as well.

Long transitions pose their own risks

If the adoption of reform is made easier by introducing long transition periods and/or postponing resolution of some contentious issues or exempting certain groups, then mechanisms need to be put in place to ensure that the reform is completed. One way to do this is to introduce a degree of automaticity into the policy process, e.g., by providing for more or less automatic actuarial adjustments in pension systems (France and Italy: pensions) or by defining specific procedures and deadlines in the original legislation for handling reform-related issues still to be resolved (Poland: pensions). The government may also delegate authority to resolve some contentious issues to independent regulators or institutions with special expertise (Sweden: postal reform; Australia: electricity). Both these devices can have the effect of shifting responsibility for sensitive decisions from politicians to relatively apolitical bodies, a form of political pressure deflection that policy makers may find attractive. However, it is important that legislative provisions mandating future actions be credible and that independent regulators be strong enough and independent enough to pursue their mandates effectively; otherwise, issues deferred at the adoption stage may be postponed again and again, as lobbies mobilise in an effort to thwart full implementation of the reform (Poland: pensions; Italy: pensions and labour market). Sustaining the momentum of reform through a lengthy transition may be easier if the initial stages of the reform give rise to the creation of new institutions or the emergence of new actors who will have an interest in pressing for further reforms later on (Sweden: postal reform; Australia: electricity and water; Poland and Mexico: pensions;
Italy\textsuperscript{136}, Germany and Spain: labour market; Figure 2.5). It should also be easier where reforms are adopted on the basis of a fairly broad consensus.

![Figure 2.5. Reform scores and the emergence of new actors](image)

**Tactical concessions to potential losers need not compromise the reform**

Concessions are sometimes made to opponents of reform in order to secure their involvement or at least their acquiescence. Such concessions may contradict the logic of the reform, rendering it less coherent overall, usually by creating too many exemptions and exceptions to the reform (France and Germany\textsuperscript{137}: labour market; Poland: pensions\textsuperscript{138}). However, it is important to recognise that concessions to potential losers need not contradict the overall thrust of the reform. Some concessions may actually be coherent with it. These can take the form of offering potential opponents a role in the new regime or agreeing transition arrangements that protect vulnerable groups (France, Poland, Mexico: pensions; Spain: labour market). They may also take the form of explicit “sunset clauses” that provide for a review of the new arrangements at the end of a certain period (Spain: labour market). While commitment to such a review means that there may be a later opportunity for opponents to try to reverse the reform, it may be needed to make adoption possible and, if the reform is generating positive results, an evidence-based *ex post* review can help to consolidate it.

**Potential losers are more likely to mobilise than winners**

There is remarkably little evidence in the case studies of the active mobilisation of interest coalitions in support of reform, except in the case of some product-market reforms:

- Where pension reform is concerned, reform is usually about retrenchment: younger workers may in future have a lighter tax burden than without a reform, but they will also have less generous pension schemes. The system as a whole may be put on a sustainable footing, but individual workers are unlikely to feel like winners if they compare their post-reform pension rights with their pre-reform entitlements, rather than with the losses they would suffer if an
unsustainable status quo were not reformed in time to prevent a crisis. More striking is the lack of mobilisation by financial-sector interests in support of funded pension plans, except in Poland. However, it is not clear that activism on their part would have been helpful elsewhere (Mexico and United States: pensions), given opponents’ claims that reforms were intended to benefit financial interests at the expense of future pensioners.

- In the case of labour-market reforms, the potential winners are often labour-market outsiders – the unemployed or very precariously employed (or even those outside the labour force altogether) – whereas those threatened by reform tend to include workers on permanent contracts (Italy: labour market; Mexico: labour law). Outsiders are not as well organised or as influential as insiders. In other instances, the potential “beneficiaries” of reform may not have welcomed it: reforms aimed at activating those who would otherwise be on benefit might increase incomes for some in the end, but pressure on benefits would still be resented, particularly by those seeking a transition to early retirement (Germany: labour market; United States: pensions; Sweden: sickness insurance; Netherlands: disability insurance).

- Potential beneficiaries were much more active in support of product-market reforms (Germany and Spain: retail; Australia: electricity; Sweden: postal reform; Netherlands: rent deregulation). Water reform in Australia proved an exception to this generalisation; however, this seems to have been because there were far fewer obvious near-term winners than losers – the key question was how the costs of the transition to a sustainable policy regime would be allocated – and thus no strong private-sector lobby in favour.

The hypothesis that losers are more inclined to mobilise than winners sits well with much of the literature. This is precisely what would be expected if the costs of reform were up-front and relatively concentrated, while the benefits were longer-term and more diffuse – and, indeed, if the allocation of benefits were uncertain. Moreover, the losers will in many cases be more politically powerful than the winners. As beneficiaries of the status quo they will often, though not always, dispose of greater organisational, political and financial resources. Older cohorts are generally more influential than the young when it comes to pension reforms, and labour-market insiders are easier to mobilise than outsiders in most circumstances. Even in product markets, incumbents will more often than not have the edge on potential entrants. The literature on endowment effects and loss aversion suggests that, even given equal resources and similar positions, potential losers will be quicker to mobilise in order to resist losses than winners hoping to realise gains.

In most of the cases under study, organised labour viewed reform proposals with suspicion, at least initially. While trade unions did in the end play a role in designing and implementing a number of major reforms examined here, they more often constituted a significant – though by no means the only – source of active opposition to reform. This reflects at least two factors. First, two of the three policy domains under study – pensions and labour-market policies – are particularly sensitive to trade unions. In most OECD countries, unions tend to represent older cohorts of incumbent regular workers, and in many, they also represent large numbers of pensioners. In some of the product-market cases, moreover, unions were also engaged, because the rents generated by existing regulatory regimes were shared with labour. Secondly, owners of capital are, ceteris paribus, able to adapt to reform more easily than other agents, except perhaps in cases where reform risks leaving them with stranded investments. They can adjust their
portfolios in response to policy changes and thus reap at least some portion of the benefits of a growth-enhancing reform. They are also better able to cope with the costs of non-reform in many situations, having recourse to a wider range of adaptive strategies in response to sub-optimal policies and institutions. This puts them at an advantage in any conflict about how the costs of reform will be allocated. Moreover, this is likely to matter more where the distributional consequences of reform depend on the specific measures adopted and are relatively easy to anticipate – e.g., in connection with labour market reforms and pension reforms rather than, for example, financial or trade liberalisation. Thus, while unions’ opposition to reforms is sometimes presented by critics as ideologically motivated obstructionism, labour leaders often have reasonable grounds for worrying about the impact of reforms on their constituents.

In addition to casting light on the unions’ attitude to reforms, this last point suggests a further reason why business interests, whose role in most of the episodes was weak or ambiguous, rarely seemed to mobilise in support of reforms that, at first glance, would appear to have offered them benefits (Mexico and United States: pensions). In a number of the episodes, employers had already devised strategies for dealing with the status quo, which, though sometimes costly, reduced their interest in reform, particularly where support for structural reform might aggravate relations with the unions and thus make collective bargaining more difficult (Mexico: labour law; Spain: labour market; Sweden: sickness insurance).

**Relatively modest institutional changes can help strengthen the reform coalition**

As noted above, creating “early winners” or new actors and institutions in the course of a reform can help to sustain the momentum behind it. It is sometimes possible to foster the emergence of private-sector constituencies with an interest in sustaining the reform, such as the private pension fund industry in Poland, but it may well be easier to affect the balance of interests by creating new public-sector bodies, like the pensions regulator in Poland or the various regulators created in the Australian and Swedish product-market reforms, or by restructuring institutions that already exist. A number of cases highlight the potential for relatively limited institutional changes to affect prospects for structural reform. Where inter-ministerial conflict is an issue, reconfiguring the institutions involved in shaping a given policy may help break the impasse (Germany: retail and labour market; Poland: pensions). Others point to the degree to which fairly modest changes in administration and financing of social insurance schemes can alter the incentives facing the non-state actors involved (Sweden: sickness insurance; Netherlands: disability insurance). Such moves may not directly resolve the problems that give rise to the need for reform, but they may well make it easier to tackle that problem via subsequent, more significant changes.

**Can reforming governments win re-election?**

There is a widespread belief among policy makers that voters will punish bold reforms, and this has been cited as a major reason why progress in reforming labour markets, product markets and social insurances has sometimes been very slow. The empirical evidence casts doubt on this expectation: other things being equal, reforming governments do not appear to face particular difficulty in winning re-election. The present study, because it looks only at reform attempts actually undertaken, rather than at the propensity of governments to attempt reforms, cannot address this question directly. However, the likelihood of subsequent re-election was little different between the more
and less successful reform episodes, and the re-election rates for all governments in the study was close to the average for all governments at all national elections in the ten countries covered during 1992-2008 inclusive. While the number of cases involved is too small to permit generalisation, it is also worth noting that four of the five governments that successfully adopted and implemented reforms for which they had clear electoral mandates subsequently went on to win re-election. The cases studied here thus cast some doubt on the oft-repeated claim that good economics is necessarily bad politics.

Notes

1. For the sake of brevity, cases are referenced in the text with minimal explanation or elaboration. Since many of the episodes cover extended periods and discuss more than one policy process, individual cases may sometimes be referenced in ways that point to contrasting lessons. Where this is the case, endnotes are used to provide clarification.

2. Let alone the “manual for technopols” described by Williamson (1994).

3. Boeri et al. (2006) likewise emphasise the importance of such conditions.

4. It is important to note that this hypothesis concerns action following election or re-election: the case studies do not support the so-called “honeymoon hypothesis”, according to which new governments are more likely to advance reforms successfully. Williamson and Haggard (1994) find little support for this view, which is also contradicted by the results obtained by Høj et al. (2006). Governments often need to find a suitable opportunity to advance reforms, but such opportunities may not always occur early in the life of a new government.

5. The Biagi law.

6. The Mexican labour-law case is somewhat peculiar, in that the reform agenda endorsed by the president during the election campaign was not that which the administration pursued in office.

7. The one clear case of reform reversal following such a policy shift (a “defection” from populist promises to reformist policies) was also the only one in which there was not a widespread sense that the economy was in crisis at the time the reforms were initiated; see Williamson and Haggard (1994:584-86).

8. In Spain and Germany, for example, both major parties were divided over retail regulation; in Australia, the same was true of electricity reform. Boeri et al. (2006) suggest that conflicts over product-market reforms often involve interest configurations that do not reflect the structure of partisan competition.

9. The United States is a special case: with national elections every two years, virtually all reforms are undertaken soon after elections – but they are also inevitably adopted in “pre-election” periods. Separation of powers also complicates the picture: PRWORA was adopted towards the end of the Clinton Administration’s first term, but it was taken up by a new Republican majority in Congress almost immediately after it won control of the legislature. The Australian cases are also difficult to assess in terms of the electoral cycles, because both involve not only Commonwealth but
also state governments, whose electoral cycles are not synchronised with one another.

10. Since the present study includes one more successful and one less successful case per country, it is by definition impossible to draw conclusions about the impact of different voting systems. IMF (2004) finds that majoritarian systems are associated with more reform; Boeri et al. (2006) challenge this view, however, and Persson (2003) suggests that more proportional voting systems may favour more gradual reform than majoritarian systems but with a lower risk of reversal. Høj et al. (2006) examine the potential effects of proportional majoritarian systems and find them to be insignificant in both aggregate and individual policy regressions.

11. Divisions within the governing majority appear to have been related at least partly to manoeuvring ahead of the 1995 presidential election campaign.

12. In all of the more successful cases, reforms were undertaken shortly after an election and adopted relatively swiftly. The initial reform of Swedish sickness insurance, by contrast, was undertaken at the end of a parliament, and the long reform-design process in Mexico meant that two competing proposals reached Congress in the run-up to the 2003 elections. It should be noted, however, that France’s CIP was initiated and blocked within the first year of the parliamentary term.

13. The 1994 pension reform attempt in Italy and the US Social Security reform debate of 2005 were the exceptions to this rule. The 1995 Dini reform in Italy passed with the backing of the centre-left parties that supported the Dini government, but only the far left and far right actually voted against it; the main centre-right parties (Forza Italia and the Christian Democratic Centre) abstained.

14. IMF (2004:116) finds that the proximity of elections can indeed hamper reforms, but Duval and Elmeskov (2005) report that their election-year dummy variable is insignificant; lags and leads of the dummy were found to be insignificant in the same specification.

15. In the US case, the two parties were effectively competing to claim credit for what was expected to be a generally popular reform and to demonstrate that they had made good on past campaign promises.

16. For more on the incentive for outgoing incumbents to adopt reforms aimed at binding their successors, see Goodman (1991), Cukierman (1994) and Tompson (1998).

17. Specifically, the first wave of radical reform in Victoria, following the 1992 elections.

18. The Biagi law.

19. See Boeri et al. (2006); and Williamson and Haggard (1994).

20. These results are consistent with the conclusions drawn by Høj et al. (2006), who find the size of the government’s parliamentary majority to be insignificant in both aggregate and individual policy regressions, and by Boeri et al. (2006), who likewise find that seemingly “weak” governments do succeed in advancing reforms, though the strategies they adopt are naturally different.

21. The Australian product-market regulation episodes, of course, involved a number of state and Commonwealth governments of different partisan orientations.


25. It is possible that, in consequence, centre-left governments undertake more modest reforms than governments of the right. This would be consistent with the econometric evidence, but the cases presented here offer no basis on which to assess such a hypothesis.

26. It is nevertheless noteworthy that the 1994 reform in Spain and the Hartz reforms in Germany came at a time when relations between organised labour and centre-left governments were strained, not least owing to what the governments saw as union intransigence vis-à-vis needed reforms.

27. A final possible explanation for the apparent contrast between the case studies and the econometric results is that, in Europe at least, reforms of employment-protection legislation (EPL) seem to have been undertaken with roughly equal frequency by governments of the right and the left, but reform reversals in respect of EPL tend to be more common under the left. Saint-Paul (1996:284) concludes in respect of “two-tier” labour-market reforms that “there are about as many instances of the reform being achieved by a left-wing government as by a right-wing one.” See also the data presented in Ochel (2008).


29. Constitutionally, Spain is not a federation, but it is one of the most decentralised countries in Europe, and the relationship between the central government and the country’s 17 regions (Comunidades Autónomas) was critical to the debate over retail hours and entry. Mexico, by contrast, is a federal system, but both reforms included in the study concern federal policies, and the states played no significant role in these episodes.

30. The Spanish retail case study explores some of the reasons for the very different outcomes observed in the two countries.

31. Econometric studies are better for addressing this question than a case-study approach. These suggest that economic crises, protracted periods of slow growth and high unemployment are all important drivers of reform. See, in particular, IMF (2004), Duval and Elmeskov (2005) and Høj et al. (2006).

32. IMF (2004); OECD (2006).

33. See the Hartz reform case study for details.

34. See IMF (2004); Duval and Elmeskov (2005); Høj et al. (2006).

35. Poland is a special case: fiscal consolidation remained a priority during the period when the reform was first being prepared, but progress in shoring up public finances meant that the government could plan to use substantial future privatisation revenues to finance the one-off costs associated with structural pension reform rather than to plug holes in the current budget.


37. There is a micro-level link as well as a macro-level relationship: individuals who feel themselves protected by EPL are less willing to pay for unemployment insurance than those who do not. The United States, however, is an exception, with low EPL and relatively limited UI.

38. See, in particular, Saint-Paul (1996, 1998); Elmeskov et al. (1998); Duval and Elmeskov (2005); Boeri et al. (2006); and Ochel (2008).
39. The PRWORA reform in the United States also focused on labour-market outsiders, but US unemployment was already well below its cyclical peak and EPL in the US is in any case far lower than in much of OECD Europe.

40. In a study of EPL reforms in 16 European countries covering 1990-2003, Ochel (2008) reports a reduction in EPL for regular workers at a time of rising unemployment in only one other instance – Finland in 1990-91. See also the data in Høj et al. (2006:38), which shows Spain to be the OECD member recording by far the largest reduction in EPL for permanent workers during 1985-2003; indeed, the OECD average score on this indicator was roughly unchanged over the period.

41. On the unique position of Spain in this respect, see Ochel (2008): exceptions to the “two-tier reform” pattern remain extremely rare.

42. In other words, the median voter was arguably a labour market outsider; Bentolila et al. (2008).

43. IMF (2004); Duval and Elmeskov (2005); and Høj et al. (2006).

44. Cross-border shopping was a significant concern for many German Länder prior to the reform of the Ladenschlussgesetz.

45. The most significant exception is the Italian case, where many of the reforms of the early-to-mid 1990s stemmed from the social pacts agreed in 1992-93.

46. The OECD composite indicator for product-market regulation began falling fairly rapidly in 1990, and this rapid decline continued through 1996, with a somewhat gentler downward slope thereafter.

47. The coefficient for policy interactions within the pension reform domain is also significant, but, as explained in Annex 1.A1, the relationship between the pension policy indicator and pension reform outcomes should be viewed with caution. The product-market indicator measures changes in seven sectors around the time of the reform and is therefore unlikely to be driven solely by the reforms under study. It is thus a far more reliable indicator of the broader policy context.

48. In the case of postal reform, the action-forcing event was CityMail’s entry. After a long period of study and discussion, the government had to decide whether to uphold Postverket’s monopoly or not.

49. On the link between crisis and reform, see inter alia, Williamson (1994); Rodrik (1996); Elmeskov et al. (1998); Drazen (2000); IMF (2004); Duval and Elmeskov (2005); Høj et al. (2006); and Boeri et al. (2006).

50. While there is little doubt about the link between the crisis of 1992 and the labour-market reforms that followed later in the decade, as the July 1993 pact was implemented, the crisis did not force immediate labour-market reform. Rather, it set in train political changes that created a new opening for reform.

51. The end of PRI rule shook up many established interest coalitions in the early 2000s, lending a fluidity to government-union relations, in particular, that opened up opportunities for reform.

52. The fiscal crisis in Victoria was indeed an action-forcing event, but it did not compel the Victorian government of the day to opt for power-sector restructuring as part of its fiscal package; the government rather saw the crisis as an opportunity to press ahead with a reform that was considered desirable for other reasons.
53. There are good grounds for believing that the adjustment costs are higher when reforms are undertaken in response to crisis, though this depends to some degree on the nature of the crisis. See IMF (2004).

54. It is necessary to acknowledge, however, that the 1992 Amato pension reform in Italy was also born in the midst of a crisis. Though this reform stuck, some of the cuts it contained were restored when the implementation decrees were issued by the Ciampi government in 1993.

55. Such “pre-work” takes a variety of forms, from research and analysis by independent institutes or government bodies to inquiries by official commissions or committees of inquiry. What is crucial is that officials and policy makers seeking to design a reform have access to a pre-existing body of good-quality research on which to draw. See Cole (2007).

56. The initial 1991 reform package was adopted with little preparation in the midst of an intense crisis.

57. Some recent Slovak analyses of that country’s pension reform conclude that one of the main reasons for the need to revise the reform quite substantially within a few years of its adoption was that it was prepared and implemented too quickly and in the absence of a social consensus on reform; Lendacky (2008).

58. The blocked Italian pension reform of 1994 and the proposed US Social Security reform in 2005, both of which are covered in case studies, highlight the difficulties of trying to reform pensions rapidly, as does the defeat of the French government’s pension proposals in late 1995.

59. Ross (2007) refers to these two processes as de-legitimation (the policy regime comes to be seen as a contributory cause of the policy problem rather than a solution) and de-institutionalisation (incremental erosion of the programme over time as a result of piecemeal early reforms).

60. Analyses that focus on political agency and the interplay of interests, treating policy regimes only as a dependent variable, often miss the significance of “reform ripeness”, which figures more frequently in accounts emphasising the importance of institutions and path dependence.

61. The first attempt to reform Swedish sickness insurance not only represented a sharp reversal of course following a period of rapid expansion in the generosity of social insurances, it also cut against the prevailing approach to economic reform overall, which combined openness to competition with generous social protection.

62. This is particularly true in respect of proposed reforms of fundamental welfare-state institutions; see Ebbinghaus (2006:770).

63. This supports the view of Dewatripont and Roland (1994) that it may make sense to pursue reforms in sequence if they are complementary. See also Boeri et al. (2006:206-31).

64. For an example of this process at work in the pensions domain, see Palier and Bonoli (2000); and Palier (2007). Glazer (1992) goes further, in arguing that agents who have adjusted their behaviour in anticipation of a reform may even have put themselves in such a position that they would suffer losses if it were not implemented.

65. See James and Brooks (2001). In the present study, the French and Mexican cases, in particular, exemplify this trend. Poland does so in a modified form.
66. OECD (2007:175); Høj et al. (2006); and Boeri et al. (2006:216-19).
67. The Biagi law.
68. As noted above, only Spain, among the present set of cases, made any real progress in this respect, and the literature on labour-market reform highlights just how unusual the Spanish experience is. See, in particular, Saint-Paul (1996); Ochel (2008); and Bentolila et al. (2008).
69. On the “stealth reform vs transparency” debate, see Hirschmann (1968); Williamson and Haggard (1994); Pierson (1994, 1996); Rodrik (1996); and Arroyo (2008).
70. The 1995 Dini reform.
71. The initial presentation of the Hartz commission recommendations; as noted below, there were costly communications failures later on.
72. The 1994 pension reform proposals.
73. Specifically in respect of the implementation of the benefit reforms associated with Hartz IV.
74. In addition to the failed Italian pension reform of 1994 and the proposed US Social Security reform in 2005, both of which are covered in case studies, attempts at unilateral pension reform were stymied or reversed in France (1995) and Germany (1997). The Slovak reform was not reversed, but the adoption of a pension reform in the absence of a broad societal or political consensus was followed by a change of government and significant revision of the reform.
75. In fact, this particular point is less applicable to the Hartz reforms than to the Bundnis für Arbeit, which preceded them.
76. Williamson and Haggard (1994); see also Webb (1994) on the importance of the “ideological conversion” of policy-making and opinion-forming elites in sustaining crisis-induced reforms once the crisis has passed.
77. See Williamson and Haggard (1994); Rodrik (1996); and Arroyo (2008). Pierson (1994, 1996) argues that welfare-state retrenchment, in particular, can occur only via a policy of stealth. This view arguably finds support in the experiences of the many countries in which reforms have been successfully implemented with little prior consultation, often in the face of strong public opposition and at times in contravention of the declared policies of the governments undertaking them. See, for examples, Nelson (1990a, 1990b); Williamson and Haggard (1994) and Boeri et al. (2006).
78. This argument is expressed with considerable force in Sachs (1994). The corollary is that the existence of a prior social consensus will not prevent reformers from being punished for failed policies.
80. IMF (2004) confirms this frequently discussed “stylized fact” about structural reform: the short-term output consequences of structural reforms are often negative, although they can improve performance significantly over the longer run.
83. The Hartz case is an interesting example: the scandal at the Federal Employment Office in 2002 abruptly drew attention to the cost and ineffectiveness of prevailing policies.

84. The Productivity Commission and its predecessors, going back to the Tariff Board, have a long tradition of doing this, beginning with the first serious attempts to quantify the costs of Australia’s then protectionist policies and to determine who bore those costs.


86. See Armingeon (2004) for an overview, and the associated studies of the Netherlands (Binnema, 2004); Sweden (Carroll, 2004); Italy (Bertozzi and Graziano, 2004); France (Serré and Palier, 2004); Spain (Álvarez and Guillén, 2004) and Germany (Zohlnhöfer and Zutavern, 2004). See also Armingeon (2005) and Mahon and McBride (2008).


88. Armingeon (2005); and Mahon and McBride (2008). The Jobs Strategy and other OECD studies were frequently cited in French parliamentary debates on the CIP and related reforms.

89. On the role of benchmarking, see Armingeon (2004); Carroll (2004); and Mahon and McBride (2008).


91. In Finland, a similar role has been played by the Finnish Centre for Pensions (www.etk.fi/Default.aspx?Lang=2).

92. The Czech pension reform of the early 2000s proceeded in similar fashion: while five major reform proposals were put forward by the major political parties, the cross-party working group that introduced the proposals prepared and analysed them under a common agreed framework, including common economic and demographic assumptions.

93. While recourse to such institutions can slow the policy process, they may also help prevent mistakes that might be made if the government were free to act in haste; the cross-correlation between time spent preparing the reform and the use of such institutions is both high ($r=0.634$) and significant.

94. The Polish BPR represents an intermediate case: it was not permanent, but it operated for five years under governments of both left and right, and was thus able to establish its authority and credibility.

95. A look at the two institutions’ web sites reveals the marked increase in the salience of these issues in their respective work programmes after the first efforts at reform had run into difficulties.

96. Danziger’s (1999) assessment of welfare reform research in the United States could apply to a number of policy regimes examined here.


98. Differences of opinion between the Ministry of Economics and Technology and the Ministry of Labour and Social Affairs.
99. The conflict here concerns divisions among congressional Republicans over the package.

100. Mainly differences between the regional and national branches of the main political parties.

101. Williamson and Haggard (1994:578-9) find that every successful case among the 15 episodes in their study has a coherent reform team, arguing that “the more vital factor is the existence of a coherent and determined government with adequate political support”. Nelson (1990b:347) concludes, “The cases of clear failure [of reform] all traced collapse in large part to deeply divided economic teams.” See also Piñera (1991).

102. The rank correlation between government cohesion and electoral mandates is 0.586, and that between cohesion and communications is 0.638. Both are statistically significant at the 5% level.

103. In the case of the 1994 Italian pension reform proposals, one coalition party justified its decision to oppose the reform while remaining in government by arguing that the proposals had not figured in the coalition agreement or the government programme and that the coalition parties were therefore not bound to them.

104. The labour minister was instructed to explore possibilities for reform but given no clear mandate.

105. See Harberger (1993); Williamson and Haggard (1994) and Piñera (1991) advance this argument chiefly in terms of chief executives, which is probably correct where large reform packages and major shifts in economic policy paradigms are concerned.

106. The Swedish postal reform is a case in point.


108. When legislative or regulatory changes in market conditions leave companies or individuals stuck with investments that they would never have made if they had anticipated the change in the legal-regulatory regime, possibly rendering previously valuable assets worthless, the assets in question are said to be “stranded”. In such situations, it is common to compensate investors, provided that (1) the cost of the asset has not yet been recovered and (2) it cannot be recovered by sale or alternative use.

109. See, in particular, Fernández and Rodrik (1991), who show how, given uncertainty about who will win and lose, a majority of risk-neutral voters may reject a reform that is expected to (1) generate net benefit overall and (2) benefit most voters. Nelson (1990b) adds time horizons to uncertainty, since agents discount the future.

110. Williamson and Haggard (1994:531) suggest that policy reform “is like an investment that should ultimately benefit the majority by enough to make them happy they made it, but that in the short run will – like all investments – involve sacrifices. The distribution of these sacrifices over time and across groups is at the heart of the politics of economic reform.” See also Drazen (2000:432): “…agreement to enact a reform carries with it a distribution of the costs of reform.”

111. Some object to generous compensation in principle, seeing it as tantamount to buying out rent-seekers when the elimination of rents and the move away from a “rent-seeking society” is the essence of economic reform. For sharply differing perspectives on the desirability of such an approach, see Bates (1994), Williamson and Haggard (1994); Koromzay (2004); and Delpla and Wyplosz (2007).
112. The experience of Italy’s Treu reform was more typical: the side payments used to smooth its adoption were modest and aimed at small but strategically important parties whose co-operation was needed. They were not major elements of the reform bargain.


114. With the important qualification that corporatist processes worked after the BfA scandal and the Hartz report had changed the terms of the debate, strengthening the position of government reformers and putting the opponents of reform on the defensive.

115. In 1997; the 1994 reform was imposed by the government after corporatist consultations broke down; this experience strengthened the position of the government during the 1996-97 negotiations.

116. The coefficient for the pensions domain is not significant, but this appears to reflect lack of variation on the independent variable: in only one case (Poland: farmers’ social security) did the evidence suggest that the government, which was both weak and divided, was unable either to act unilaterally or to punish non-cooperation on the part of other agents.

117. The pre-Hartz experience of the “Alliance for Jobs”.

118. If the government is perceived to be divided over the policy, it will find it very hard to make any credible implicit or explicit threat to act unilaterally or punish non-cooperation; hence the high and significant cross-correlation between these two variables ($\rho=0.661$).


120. The Biagi law.

121. The issue does not appear to have been studied systematically in the political economy literature, but in the case studies examined here, partial or total reversals occurred only in respect of unilateral reforms. Negotiated reforms were sustained, even where significant criticism of, and opposition to, the agreed reforms subsequently emerged (Italy and Spain: labour market).


123. See James and Brooks (2001) on the use of this strategy in other countries.

124. In respect of the collective bargaining reforms adopted in 1994 and 1997 rather than the adjustments to EPL.


127. Bonoli and Palier (2007); Bonoli (2008). Indeed, in Italy, the very long transition agreed for the Dini reform meant that only a very small minority of those who had reached voting age by 1995 would feel anything like its full impact.

128. Uncertainty on this score may encourage older workers to retire earlier rather than later, as they seek to claim their pensions before any further retrenchment can occur.

129. If pension benefits are viewed as deferred wages, then it may make sense for agents to accept adjustments in pension entitlements – in effect, a reduction in the rate of
return on their pension savings – in return for greater confidence that their benefits will be paid. See Immergut et al. (2007).

130. See, in particular, Ebbinghaus (2006) and Immergut et al. (2007).

131. In Germany, for example, there has long been a tradition of treating status-securing social insurance entitlements, based on dedicated contributions, as a sort of property right, whereas in the United States this is not the case.


133. This is the so-called “taxicab medallion problem”; see Koromzay (2004).

134. Moreover, the stranded asset problem was not limited to the irrigation sector. If changes in water policy prompted a shift towards less intensive dry-land agriculture and a corresponding decline in the farming population, then many rural communities would experience second-round effects, with schools, hospitals, banks and other services closing.

135. This is the logic that underlies many arrangements for automatic indexing of pensions and other benefits. See Weaver (1986); and Anderson and Zanardi (2004).

136. The Treu law.

137. The extension of the duration of so-called “UB I” benefits for older workers under Hartz IV.

138. The preservation of sectoral early retirement rights.

139. The principal short-term winners were South Australian irrigators. Being situated at the end of the river system, they had much to gain from water reforms that liberalised trade, since they tend to receive less seasonal allocations than upstream users when water is scarce, and from measures that addressed overallocation upstream. The South Australian irrigators and associated farm lobby groups are a strong private sector lobby group in favour of water reform.

140. This is not to suggest that older cohorts are anti-reform or that ageing societies reform less: Høj et al. (2006) find that rising old-age dependency ratios are associated with more product-market reform and with greater reductions in the implicit tax on continuing to work after 55. However, the case studies suggest that pension and labour-market reforms typically shield older cohorts from most of their effects, while many product-market reforms may be particularly beneficial to pensioners, since the price effects of greater competition are likely to be especially important to those on fixed incomes.


142. See Visser (2007) for data on the age structure of trade union membership in 20 European countries.

143. Labán and Sturzenegger (1998) make this point with respect to the politics of stabilisation in Latin America, but it has a wider application.

144. Even when there is widespread agreement on the need for change, reform may be delayed by conflict over the allocation of costs; those with superior adaptive strategies – i.e., those better placed to prosper even in the absence of policy change – have a significant bargaining advantage in such conflicts. See Alesina and Drazen (1991); Labán and Stürzenegger (1998); and Drazen (2000). To take a relatively banal example: better-off households tend to be less reliant on the public pension system for retirement income than poorer households. They therefore have less to
lose if that system proves unsustainable and hence less incentive to make sacrifices to shore it up.

145. See, in particular, Alesina and Drazen (1991); Labán and Sturzenegger (1998); Drazen (2000).

146. Buti et al. (2008); Williamson and Haggard (1994).

147. The average re-election rate for the period was around 30%, not counting cases in which the dominant party in a government coalition remained in office after the election but with new partners. The Australian electricity case is omitted, owing to the fact that a number of governments were involved over an extended period, some of which won re-election and some of which did not; the Mexican pension reform is also omitted, since the Calderón administration is still in power.
Bibliography


IMF (2004), “Fostering Structural Reforms in Industrial Countries”, chapter III of *World Economic Outlook: Advancing Structural Reforms*, International Monetary Fund,


Part II

Case studies in pension-system reform
Chapter 3

France: The 2003 pension reform

Summary and main conclusions

In mid-2003, France adopted a major pension reform intended to reduce the inequalities that existed between private- and public-sector pension systems, to assure the financial sustainability of both, and to reduce the existing large incentives for early retirement. Despite strong opposition, the government succeeded in extending the provisions of the 1993 reform of private-sector pensions to the public sector. Incentives to extend working lives were strengthened via a phased increase in the minimum contribution period (MCP) for receiving a full pension, which is to rise by one year during 2009-12, with further gradual increases thereafter, as required to maintain a constant ratio of contribution years to years of retirement. In other words, the authorities have instituted a partial indexation of the MCP to life expectancy at retirement. The reform also introduced benefit accruals for persons who continue working past the statutory retirement age and beyond the MCP for a full pension. Prior to the reform, those who worked longer effectively lost significant net pension wealth. The reform generated fierce opposition, and its success owes much to the lessons drawn from earlier reform attempts.

Although the reform aroused intense opposition, particularly from public-sector unions, the government was able to secure its adoption with relatively few modifications. This success largely reflected the lessons learned from two previous reform attempts – the successful reform of private-sector pensions in 1993 and the frustrated attempt to reform public-sector pension schemes in 1995:

- **Confrontation and consensus-seeking.** The government’s strategy of “consultation without negotiation” enabled it to avoid the risks associated with the kind of unilateral action that had been attempted in 1995 without engaging in the kind of negotiations with social partners that might have led to a watering-down of the reform. While significant concessions were made at a key point in the process, the government demonstrated sufficient resolve in the face of protests to ensure that its principal aims were not compromised.

- **Consultation and communication.** In contrast to 1995, the government went to great lengths to prepare public opinion for the reform and to consult with a wide range of stakeholders, including the unions and employers’ organisations. While this did not generate a societal consensus in support of the reform measures themselves, which were fiercely contested, it did secure widespread acceptance of the need for reform and it defined the parameters of the pension reform debate. This success owed much to the work of the **Conseil d’Orientation des Retraités** established by the previous government.
• **A targeted reform.** Early on, the government took a deliberate decision to leave the special regimes covering employees of large public enterprises out of the reform. Unions representing these workers had spearheaded opposition to the 1995 reform proposals. Adoption of the reform, in turn, further increased the pressure for a reform of the special regimes, as their privileges became ever more anomalous and thus harder to defend. The question of supplementary funded schemes was also left to one side, in order to avoid confusing the debate over the basic PAYG system.

• **Time profile.** Both technically and politically, the reform was prepared with considerable care over the course of 2002, ensuring that the government as a whole – and, in particular, the ministries most directly concerned – remained united behind it.

### The context for reform

#### Framework conditions for reform

As noted in the case study of the CIP, France has a long tradition of state-led policy making. Prior to the 1990s, studies of the policy process in France tended towards the view that the state could largely define and implement its priorities without having to compromise openly with external interests.¹ This view, however, was always somewhat exaggerated, and a number of episodes in the 1990s, including a failed attempt at pension reform in 1995, underscored the dangers of such an approach. While the state remained strong and was indeed capable of managing the policy process with little input from other actors, unilateral government action that was perceived to threaten the interests of important social constituencies more than once led to large-scale demonstrations and, in the end, the withdrawal of proposed reforms. These lessons were not lost on the government of Prime Minister Jean-Pierre Raffarin as it prepared its pension reform.

The political conjuncture, at least, was especially favourable for reform in 2002-03: the centre-right had won both the parliamentary and presidential elections of 2002, so the government enjoyed the support of a newly re-elected president and commanded large majorities in the National Assembly and the Senate. Moreover, the government could claim a mandate for public-sector pension reform, which had been a prominent part of its 2002 election platform. The opposition was in disarray following its election defeats, and no other major elections were scheduled before 2004. Pension reform was not, in any case, an issue on which the partisan divide was really critical. Successive governments of left and right had identified essentially the same problems and even settled, in some cases, on similar solutions: many elements of the Raffarin government’s reform had been set out in a speech by Socialist Prime Minister Lionel Jospin in 2000.² While the left opposition was certainly critical of the government proposals in 2003, the real axis of conflict was between the government and the trade unions.³

Pension politics in France, as in many OECD members, are largely shaped by demographic trends. The country has in recent decades experienced relatively low fertility by historical standards (though the total fertility ratio remains well above the levels found in much of Western Europe), low immigration and rising longevity. The population is ageing fast, a process that reflects both the ageing of the “baby boom” cohorts, which began to retire in 2006, and the longer-term trend towards increased life expectancy.⁴ The implications of these two factors for pension reform are somewhat different: on its own, the ageing of the “baby boom” generation could be regarded as a one-off shock and addressed as such, but the longer-term increase in longevity presents a
structural challenge. The authorities in 2001 estimated that by 2030, the population over 65 would rise by 75%, while the working-age population (15-64) would be stable or even decline. The dependency ratio based on age groups alone (i.e. 65+ as against 15-64) was thus expected to rise from 1:2.4 at the beginning of the century to 1:1.5 in 2030 and 1.3:1 in 2050. Given the then-current retirement age of 60 and the incidence of early retirement, the effective dependency ratio was expected to rise from around 1:2 to roughly 1:1 by 2050.

Macroeconomic circumstances had little direct bearing on the politics of reform, the immediate impact of which was unlikely to affect, or be affected by, short-term fluctuations in economic performance. Nevertheless, poor economic performance probably contributed to the government’s gradual decline in the polls during early 2003, as the first half of the year showed a slight contraction in activity – real GDP fell by 0.2% year-on-year – rather than the modest acceleration that most had expected. Slower growth pushed up the general government deficit, which rose to around 4% of GDP, but this, too, had little impact on the pension debate, since the proposed reforms were aimed at the longer term and not geared to generating rapid savings.

Actors and interests

Apart from the government itself, the major actors in the pension debates of 2002-03 were France’s principal trade-union confederations: the Confédération Générale du Travail (CGT), Force Ouvrière (FO), the Confédération Française Démocratique du Travail (CFDT), the Confédération Française des Travailleurs Chrétiens (CFTC), and the white-collar Confédération Générale des Cadres (CGC). Overall, France’s trade unions are not particularly strong – unionisation has been below 10% of the labour force (and below 5% in the private sector) since the 1980s, and the three largest confederations often find it difficult to co-operate, owing to differences of approach. In general, the FO and the CGT, both of which are primarily public-sector confederations, tend to adopt a tougher line vis-à-vis the government and employers than does the more moderate CFDT. Moreover, the peak-level confederations are characterised by a low level of vertical co-ordination because of the independence of their constituent unions. Nevertheless, the blocking power of the unions in the field of pension reform is considerable, given their ability to mobilise public-sector strikes and their role in administering pensions, health insurance and unemployment insurance. Moreover, the ageing of the unions’ membership and the prevalence of early retirement mean that the power of the special pension sections of the major unions is considerable. So is their capacity to influence public opinion: polls during the spring of 2003 showed them consistently running close to the government on questions of public trust.

It would be a mistake, however, to see the unions as wholly opposed to pension reform. By 2002, all of them had accepted the need for some change, and most were moving towards acceptance of a 40-year MCP for both public- and private-sector workers as a condition for entitlement to a full pension. At an early stage in the 2002-03 reform process, the CFDT leadership indicated to the government that it wished to reach agreement if possible and that the CFDT would regard failure to do so as a defeat. While this was less than a promise of support, it clearly signalled a desire to shape the reform rather than to block it. Moreover, there were tensions within the CGT, which meant that its opposition was likely to be less energetic than in 1995, when it had mobilised against a previous pension-reform attempt. Indeed, the government’s aim vis-à-vis the CGT was not so much acceptance as “passive” rejection. From the beginning, therefore, the CFDT occupied a privileged place in the government’s strategy.
**Issue history**

**France’s pension problem**

The French pension system is characterised by a high degree of occupational fragmentation (Box 3.1). The régime général provides basic pension coverage for private-sector employees (65% of the working-age population). In addition there are separate schemes for farmers (3%), public-sector employees (20%) and the self-employed (12%). Pension coverage for the latter two groups is further fragmented according to employer or profession, with particularly generous provisions applying to the employees of public enterprises like the railways, the Paris metro, and the gas and electricity monopolies. In addition, complementary provision is compulsory and follows roughly the same occupational divisions.10

Operating on a defined benefit/pay-as-you-go (PAYG) basis and financed chiefly out of wage-based social contributions, the pre-reform pension system was highly vulnerable to demographic and economic pressures.11 By the mid-1980s, an ageing population (Table 3.1), high pension replacement rates and very low participation rates among the over-55s had triggered growing concern over the future of the PAYG system – despite continued reliance on early retirement to ease labour market pressures. A series of official or semi-official reports in the late 1980s and 1990s brought these problems to the attention of both policy makers and voters.12 All of these reports took a pessimistic view of the future, and it was clear from a relatively early stage that, in order for the system to remain in equilibrium in 2025, either benefits would have to be cut dramatically (by up to half) or contributions would have to rise sharply (by as much as 70% on some estimates).13 In the absence of reform, ageing alone was expected to increase the pension system’s funding shortfall by 8% of GDP by 2030. Even assuming substantial improvements in labour force participation and unemployment rates, rising pension expenditures were expected to increase the deficit by 4.7% of GDP.14 By commissioning such reports and launching public information campaigns, successive governments of both left and right sought to raise awareness of the problem and counter defence of the status quo.

Initially, however, the obvious reform options – longer careers, changes in the calculation of the reference salary and/or changes to indexation – were all less politically palatable than raising contributions. Employee contributions thus rose from 4.70 to 6.55% of the social security ceiling between 1985 and 1991. This was to be expected: the initial response of most governments running Bismarckian pension schemes that face financial constraints has been to raise contributions. However, it can never be more than a partial solution. Even if the electorate is prepared to countenance steadily rising payroll taxes, the rising tax wedge is likely to put downward pressure on employment. In France as elsewhere, therefore, the focus of attention shifted in the 1990s towards retrenchment and cost containment. The pursuit of such solutions, however, was greatly complicated by the fragmentation of the system, which meant that any would-be reformer was confronted with a wide range of interest groups mobilised to defend their specific arrangements.15 French reformers also faced a more general problem that confronts anyone trying to reform a system based on individualised, earnings-related benefits rather than universal flat-rate benefits: in paying social contributions out of earnings, a participant in such a system gains an entitlement – he/she is purchasing social rights and the contract is highly individualised. Any revision to the benefit scheme thus affects acquired contractual rights, which are hard to challenge retrospectively.16 Future pensioners are very sensitive to modification of such a system.
Box 3.1. The French pension system

Based almost exclusively on compulsory social insurance operating on a PAYG basis, the French pension system is funded by employer and employee contributions, managed by administrative councils composed of their representatives. On average, collective schemes in the early 2000s guaranteed a person with a full contribution history a pension with a replacement rate equal to 70-75% of wages. The fragmentation of the system and, in particular, the split between the general regime for private-sector employees in industry and commerce and the various special regimes for public-sector workers and for artisans, professionals and other self-employed people, reflect decisions taken when the general regime was created. Many groups opted to keep their own schemes, which either gave more generous benefits or involved lower contributions. Some professional groups have also created voluntary (funded or PAYG) schemes in recent years. However, the major split was between public and private sectors, as the scheme for the former has been consistently more generous with respect to both early retirement and the calculation of benefits: prior to the reforms of 2003, some public employees could retire as early as 50, with a pension benefit based on salary (less bonuses) over the last six months of work and replacement rates of up to 75%.

The basic benefit in 2002 amounted to a maximum 50% of a reference wage (for a full 40-year career, minus 1.25% per missing quarter), up to a ceiling of EUR 14 112. For private-sector workers, the reference wage is calculated over the 25 best contributing years (the 10 best prior to the 1993 Balladur reform); in the public sector, it covered (and still covers) the final six months. The basic benefit is topped up by a complementary pension, which is also compulsory and operated on a PAYG basis. Complementary regimes are organised by sector and job type and, although the régime général operates as a defined-benefit scheme with respect to the basic benefit, the complementary regimes are based on defined contributions, albeit with a ceiling on eligible earnings. Pension benefits in payment have been indexed to prices since 1987 (except in the case of the public sector). There is also a minimum pension (EUR 6 307.62 in 2002), contingent on a full contribution career, for those who have not made sufficient contributions or are on low incomes. The minimum pension is paid pro rata to those with shorter contribution histories.

1. There are in fact two general regimes administered by two different bodies – one for employees in industry and commerce, and another for agricultural workers – but they are now largely aligned with one another. These cover the basic pension benefit paid to most workers. However, mandatory complementary regimes are organised by broad sector.
2. Historically, this fragmentation has enabled numerous socio-occupational categories to defend their own specific interests; see Palier (2007).
3. Early retirement falls under a separate programme administered by the national employment fund and is possible from 57 or even 56 in some circumstances. The replacement rate is around 80%, and the individual switches to a normal public pension on reaching retirement age. Steps have recently been taken to curtail early retirement and reward longer careers; see OECD (2007:124) for details.
Table 3.1. Expected duration of retirement

<table>
<thead>
<tr>
<th>Year</th>
<th>Men</th>
<th>Women</th>
<th>Men</th>
<th>Women</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>15.4</td>
<td>18.4</td>
<td>82</td>
<td>89</td>
<td>12.6</td>
<td>16.4</td>
</tr>
<tr>
<td>1960</td>
<td>15.7</td>
<td>19.5</td>
<td>84</td>
<td>92</td>
<td>13.2</td>
<td>17.9</td>
</tr>
<tr>
<td>1970</td>
<td>16.2</td>
<td>20.8</td>
<td>85</td>
<td>93</td>
<td>13.9</td>
<td>19.4</td>
</tr>
<tr>
<td>1980</td>
<td>17.3</td>
<td>22.4</td>
<td>86</td>
<td>94</td>
<td>14.9</td>
<td>21.1</td>
</tr>
<tr>
<td>1990</td>
<td>19.0</td>
<td>24.2</td>
<td>88</td>
<td>95</td>
<td>16.8</td>
<td>23.0</td>
</tr>
<tr>
<td>2000</td>
<td>20.2</td>
<td>25.6</td>
<td>90</td>
<td>96</td>
<td>18.2</td>
<td>25.6</td>
</tr>
</tbody>
</table>


Pension reform in the 1990s

In 1993, the centre-right government of Edouard Balladur made the first serious attempt to tackle France’s mounting pension sustainability problems. The 1993 reform was limited to the private sector. While the official retirement age of 60 remained unchanged, the minimum contribution period (MCP) for a full pension was raised gradually from 37.5 years in 1993 to 40 in 2003, and the calculation of the reference wage on which the pension was based was gradually extended from the best ten years of one’s contributing career to the best 25. Babeau (1997) estimates that these changes reduced the de facto replacement rate from 50 to around 33% of the reference salary used under the un-reformed system. A number of factors contributed to the government’s ability to enact such a reform relatively smoothly:

- **A favourable political conjuncture.** The government commanded a solid and recently elected majority in parliament. In any case, the opposition Socialists, weakened by their defeat in the March 1993 parliamentary elections, could hardly exploit opposition to the reform, since it was based on a White Paper prepared under a Socialist government in 1991.

- **The targeting of the reform.** The reform affected only the private sector, where union penetration was lower and union activity tended to be less militant and less disruptive. Public-sector workers have been left out of the first stage of pension reform in many countries, as they tend to be older and better organised, and to have more generous pension schemes.

- **Concessions to the unions.** The Balladur reform involved significant concessions to the unions and exploited differences among them. First, the gradual phase-in of the reform meant that younger workers, who were less unionised and for whom pension issues were probably less salient, bore the brunt of the reform. Secondly, a special fund was created to finance benefits for those who had made insufficient contributions to qualify for a minimum pension. This separation of pensions based on social assistance from those based on social insurance met a long-standing union demand. The government’s confirmation of the unions’ role in managing the system also helped defuse potential opposition from the more moderate unions, such as the CFDT, which had publicly acknowledged the need for reform.

Although the CGT and FO called for a general strike in response to the proposals, there was little mass mobilisation, partly because the government put the reform through just as the season of summer vacations was beginning. This reduced the likelihood of large-scale strikes and made it easier for more moderate unions quietly to avoid taking
action. The CFDT and others had indicated that they would not support the reform but neither were they committed to active opposition, so it was important that the government leave them a discreet political exit. Thus, while the 1993 reform was unilateral, in the sense that there were no real negotiations with social partners over the shape of the reform, its adoption involved wide consultations and significant concessions to potential opponents.

In November 1995, the government of Alain Juppé attempted to tackle the question of public-sector pensions in the context of a broader social-reform package. The 1993 reforms were to be extended to the public sector, and the various special regimes in the public sector were to be eliminated. The proposals triggered a wave of strikes and protests that led the government to drop its pension reform package within a month. Given the importance that the Raffarin government attached to avoiding any repetition of the 1995 episode, it is important to highlight the reasons for the government’s defeat:

- The reform was a surprise. While it was well known that the Balladur reform had left much still to be done, pension reform had not loomed large in the 1995 presidential election campaign. Moreover, the government had prepared the reform plan in secret, fearing that preliminary negotiations with social partners would result in a watering-down of the package.  

- It threatened the unions’ management role, which had been confirmed in 1993. The new plan would have allowed parliament to vote on the social security budget, something it had never done before. For some union leaders, control of the social security system was a critical issue.

- On paper, the proposals looked much like those of 1993, but their immediate impact would have been far greater. In 1993, most private-sector workers were not really affected by the increase in the MCP: 70% of them reached retirement with over 40 years of contributions anyway. However, the vast majority of public-sector workers fell well short of the proposed new MCP.

The Socialist government of 1997-2002 studied the pensions issue at length but largely failed to act. The so-called “Charpin Report” in 1999 demonstrated quite clearly that the 1993 reforms were insufficient and suggested that the financial sustainability of the system could be at risk as early as 2010. Jean-Michel Charpin’s findings frightened many in the government, and both the cabinet and the Socialist Party were divided. A further report, prepared under the auspices of the Economic and Social Council, cast doubt on Charpin’s conclusions and strengthened opponents of reform. It stressed uncertain demographic projections, questioned the case for longer careers, and rejected the need to align special regimes with the general one. Its projections were much more reassuring than Charpin’s, although this was later shown to be due to a methodological error. The debate was further complicated by the appearance of several more reports, commissioned by the government or independent bodies, each offering its own assessment of the situation and arguing for a different approach to the problem. Many proposed solutions emphasising strong growth and higher employment rates, especially among older workers, rather than changes to the pension system. This muddying of the waters helped ensure that little was done to advance pension reform. However, in 2001, the Jospin government did establish a reserve fund, the Fonds de réserve des retraites, to accumulate EUR 150 billion by 2020 in order to smooth the sharp rise in contributions that would otherwise be needed to fund the system.
The COR and the search for solutions

The Socialist government of Lionel Jospin also created the Conseil d’Orientation des Retraites (COR), which has played a very important role in French pension reform debates since 2000. The COR brought together senior civil servants, parliamentarians and representatives of the trade unions and employers’ organisations, as well as the National Council of Retirees and Aged Persons and the National Union of Familial Associations. Its first major report, issued in the autumn of 2001, greatly influenced the debate over the pension reform of 2003. While the discussions within the COR did not lead to agreement on the precise measures needed, they did bring about a degree of consensus on some of the key issues:

- The COR came rapidly to accept the need for a serious pension reform. It was recognised that higher economic growth would not, of itself, solve the problem, as some on the left had argued.

- The range of politically and technically feasible options was narrowed. There was a strong commitment to preserving the PAYG character of the system – the pre-war history of funded pension schemes that had failed made the move to a funded system politically dangerous, and the COR experts were not in any case persuaded that it would be desirable. In fact, numerous opinion polls since the mid-1990s suggest that there is strong support for pension funds as a supplement to the existing system but that the electorate wants the public schemes run on a PAYG basis.36

- A notional defined-contribution system was also rejected relatively early on, partly owing to concerns about its technical feasibility, given the situation of the pre-reform system, but chiefly because the trade unions remained strongly committed to the retention of basic collective norms concerning such system parameters as the retirement age and the replacement rate. They feared that the flexibility inherent in a system such as that adopted by Sweden in 1998 would, in French conditions, be exercised by employers rather than employees - i.e. decentralised negotiations about retirement ages would lead to employers deciding when they wished to let older workers go, rather than workers deciding how they wished to trade off the attractions of early retirement against the loss of benefits it would entail.

The COR’s work meant that there was broad agreement that any reform would be parametric rather than structural and that the parameters subject to change included the MCP, the level of contributions, the replacement rate and the extent to which penalties and bonuses would be used to encourage longer careers. Thus, while tough distributional issues remained to be resolved, the outlines of the debate were broadly agreed, and this consensus was embodied in the COR’s 2001 report (Table 3.2).

Trends in public opinion

Public discussion of pension problems throughout the 1990s did much to prepare public opinion for the reform of 2003. A poll in 1999 found future pension levels to be respondents’ third major economic concern, after unemployment and fear for their children.37 By early 2003, some polls even showed a majority of public-sector workers ready to accept, however reluctantly, the realignment of their retirement regime with that of their private-sector colleagues. At the same time, efforts by successive governments of both left and right to encourage the development of supplementary private pension
provision contributed to a growing public conviction that pension reform was inevitable and would involve some retrenchment. Tax deductions for savings schemes for pension or general provident purposes were first introduced in 1994. While the Jospin government reversed its predecessors’ most ambitious initiative in this direction – the so-called loi Thomas (1997), which was repealed before it could be implemented – the Socialists adopted their own plan to encourage “voluntary wage savings partnership plans”, offering tax breaks for long-term savings.

Table 3.2. France’s pension problem: the “2040 choice”

<table>
<thead>
<tr>
<th>Effective retirement age</th>
<th>If the replacement rate is held constant, contribution rates will:</th>
<th>If contribution rates are held constant, the replacement rate will:</th>
<th>If the replacement rate falls to 64%, contribution rates will:</th>
<th>Change in the estimated duration of retirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unchanged</td>
<td>Rise 15 percentage points.</td>
<td>Fall to 43%.</td>
<td>Rise&lt;10 percentage points.</td>
<td>+5.5 years</td>
</tr>
<tr>
<td>+1 year</td>
<td>Rise 13 percentage points.</td>
<td>Fall to 46%.</td>
<td>Rise 8 percentage points.</td>
<td>+4.5 years</td>
</tr>
<tr>
<td>+2 years</td>
<td>Rise 11 percentage points.</td>
<td>Fall to 49%.</td>
<td>Rise 6+ percentage points.</td>
<td>+3.5 years</td>
</tr>
<tr>
<td>+4 years</td>
<td>Rise 8 percentage points.</td>
<td>Fall to 56%.</td>
<td>Rise &lt;3 percentage points.</td>
<td>+1.5 years</td>
</tr>
<tr>
<td>+6 years</td>
<td>Rise 5 percentage points.</td>
<td>Fall to 64%.</td>
<td>Remain constant.</td>
<td>-0.5 years</td>
</tr>
<tr>
<td>+9 years</td>
<td>Remain constant.</td>
<td>Remain unchanged at 78%.</td>
<td>N.A.</td>
<td>-3.5 years</td>
</tr>
</tbody>
</table>

Note: The estimates of duration of retirement incorporate expected increases in life expectancy to 2040.  

Policy process: formulation

The preparation of the loi Fillon (“Fillon law”) began even before the elections that brought the centre-right Raffarin government to power: at a political level, regular consultations involving leading politicians on the centre-right, as well as advisors to President Jacques Chirac, took place from the beginning of 2002. At a technical level, too, work was being done in the ministries concerned, in the belief that whatever government emerged after the elections would take action on pension reform, guided to a greater or lesser extent by the findings of the COR report. During the summer after the elections, work on pension policy proceeded largely within individual ministries, but from early September, there were regular (usually weekly) meetings of responsible officials from the finance, budget, social affairs, civil service and health ministries. The whole process was managed closely by the prime minister’s office. Working groups were formed, first at the level of ministers’ directeurs de cabinet and then at ministerial level. From October, the prime minister’s office also organised regular meetings of the officials responsible for communications in the five ministries, ensuring that they were informed about the reform and planning a communications strategy long before it was unveiled: there was an awareness of how hard it would be to win support for a reform that, in essence, involved working longer to receive less.

Co-ordination of the process by the prime minister’s office was designed to ensure both unity within the government and adequate preparation of both the political and technical aspects of the reform. Many of the participants were haunted by memories of 1995, and they also drew lessons from Balladur’s earlier success. Moreover, though determined to act, they were not in a hurry: soon after the elections, the government made a deliberate decision not to attempt a 100-day blitzkrieg on pensions, as some had advocated and as Balladur had done in 1993. Instead, President Chirac set a mid-2003
deadline for adoption of the legislation. The public phase of the reform was deliberately postponed until after the December elections to the Council of Industrial Tribunals, which would provide a test of support for France’s union confederations. In the event, the hardline Force Ouvrière fell back slightly and the more reformist unions did relatively well, a result that encouraged the government.

From a very early stage, the ministries involved in the reform shared a common understanding of the problem facing the pension system, the aims of the reform and the steps that they would like to take. Differences of opinion tended to concern what was politically feasible rather than what was desirable.

- The most urgent problem concerned the financial sustainability of the system: following publication of the COR report, there was strong consensus that the parameters of the system had to be adjusted.

- A second concern was equity, which was seen as an important goal in its own right but also as a political necessity if pension reform were to go forward: it was clear that further reform of private-sector pensions would be required, but a second reform of the private-sector schemes would have been difficult to justify as long as public-sector pensions remained so generous. More generally, there was an awareness that questions of equity were likely to be more sensitive politically during a period of retrenchment than they were when the system was being expanded.

- For all the talk of demographic shocks, France’s pension problem was partly rooted in its pattern of late labour-market entry and very early exit: the rate of activity among 15-24 year-olds was second lowest in the OECD, and France’s average effective retirement age of 58.7 was also extremely low by international standards, while life expectancy was comparatively high. A third goal of the reform was therefore to enhance the incentives for longer careers and to strengthen the link between contributions and benefits, which, in turn, required increased clarity, visibility and stability of rights within the system.

- Finally, given the obvious difficulty in securing adoption of major revisions to the pension system in response to changing economic and demographic circumstances, there was an awareness of the need to establish mechanisms that allowed a gradual, more or less automatic adjustment of the parameters of the system.

Like the COR, the Raffarin government quickly settled on a parametric reform. Although reforms in countries like Sweden and Italy were studied, there was never a serious discussion of shifting to a defined contribution system, largely for the reasons outlined above. Some of those shaping the loi Fillon have since indicated that they might have preferred the Italian-Swedish route if the reform had been undertaken earlier, but since they were still implementing the 1993 parametric reform of private-sector pensions, there was a strong sense that reforms would be easier to continue – both technically and politically – if they carried on along the same lines. In any case, memories of 1995 encouraged politicians to opt for low-risk strategies.

**The draft legislation**

The basic outlines of the loi Fillon reflected these considerations. The most important element involved aligning the public- and private-sector schemes, which meant raising
the MCP for public-sector workers to 40 years by 2008 and indexing public-sector pensions in payment to prices rather than wages, as had been done in the private sector since 1987. Thereafter, the public and private schemes were to evolve together, with the MCP rising to 41 in 2012 and almost 42 in 2020. Contribution rates were also to rise from 2008. The reform introduced a bonus benefit accrual \((\text{surcote})\) of 3\% per year to both public- and private-sector schemes for those who had reached the MCP and passed the age of 60, thereby rewarding longer careers. This also entailed changes in rules allowing enterprises effectively to dismiss employees once the latter are ready to claim a full pension: while the right to retire at 60 remained, the mandatory retirement age in the private sector was raised to 65, the level already prevailing in the public sector.\(^{43}\) Rules governing pensioners’ employment were also to be made more flexible. Individuals retiring before reaching the MCP were to be penalised not only via the prorating of benefits to years of service\(^{44}\) but via the introduction of an additional reduction \((\text{décote})\) for each “missing year”, up to a maximum of five years. The \(\text{décote}\) was to be phased in during 2004-14 in the private sector and 2006-20 in the public, reaching a maximum of 5\% per year in both cases. However, the right to retire on a full pension at 60 was preserved for those who had reached the MCP, as was the PAYG basis of the system.

These reforms were expected to reduce the expected private-sector system shortfall of EUR 15 billion in 2020 by EUR 5 billion. The remaining EUR 10 billion shortfall was to be covered by transferring part of the social insurance charges currently paid to the unemployment system to the pension system following a hoped-for fall in unemployment. Savings from planned changes to the public-sector system were expected to reach EUR 13 billion, as compared with a projected pre-reform deficit of EUR 28 billion in 2020.

Although there was broad agreement on the basic outlines of the package, the reformers had to resolve a number of politically difficult questions concerning scope and timing. It was decided at a very early stage that the special regimes of the large public enterprises (chiefly public transport, electricity and gas) would be left out of the reform, in order to avoid a re-run of 1995, when the unions representing these groups had played the leading role in blocking reform. Any reform of their regimes was to be negotiated between management and the unions at the level of the enterprise – something that would be necessitated in some sectors by EU policies opening them up to competition. The key question was whether further private-sector reform would be undertaken immediately or whether the government should focus only on civil servants and other public employees. There were obvious risks in trying to introduce further changes to private-sector pensions. However, there were also risks in targeting public employees only, and, given that the conditions for reform seemed favourable, there was a strong argument for seizing the opportunity to address both sectors. Moreover, MEDEF, the main employers’ organisation, indicated that it would not support a reform that affected only the public sector. In the end, the timing of the reform was altered in order to broaden its scope: the public-sector schemes were to be aligned with the private-sector ones over a five-year period, rather than ten years as initially envisaged.\(^{45}\) With the two schemes largely harmonised by 2008, it would be possible for them to evolve together: hence the decision to postpone further changes in contribution rates or the MCP until after 2007.

For a time, the alignment of public- and private-sector \(\text{décotes}\) was a controversial issue within the government. The civil service ministry feared that applying the 10\% annual \(\text{décote}\) introduced in the Balladur reform to public employees, to whom no such penalty had ever applied, would generate serious resistance. In the end, the two schemes converged at the above-mentioned 5\%. Public-sector employee interests prevailed in
respect of one rather crucial parameter of the reform: the initial draft legislation based the reference salary used for calculating the pension benefit on the final three years’ work, as against the final six months prior to the reform or the best 25 years in the private sector. In the end, however, the six-month rule remained unchanged – a significant victory for public servants, as it had long been common practice for them to be promoted in the last months of their careers in order to boost their final salaries and hence their pensions.

Policy process: adoption

The Raffarin government employed a combination of confrontation and consensus-seeking in securing adoption of the *loi Fillon*. Given its solid majorities in the National Assembly and the Senate, the government faced little difficulty with the parliamentary phase of the process; the key issue was to overcome trade union opposition, which had sunk the 1995 reform. The government’s strategy *vis-à-vis* the unions amounted to “consultation without negotiation”, and, with one notable exception, it followed this strategy fairly consistently. The President initiated the public phase of the reform process in early January 2003, with his new year’s message. Labour and Social Affairs Minister François Fillon outlined the broad contours of the reform in public for the first time in early February and spent much of the following month meeting representatives of the major trade unions and employers’ organisations. The unions made various demands (the CGT, for example, argued for counting years spent in education towards the MCP), but their principal argument was for higher employer contributions. The “main lines” of the reform were presented to the social partners in the form of a short memorandum in mid-April, and on 24 April, the labour and social affairs minister argued the case for reform on the television programme *100 minutes pour convaincre*. The draft legislation was presented on 7 May.

The unions did not wait for the draft legislation before mobilising. The first major demonstration, organised by the CGT, brought around 300,000 people onto the streets of Paris on 1 February. The protests that followed in March were smaller, but they resumed in earnest once the legislation was presented in early May. On 13 May, an estimated 1 million people took to the streets in opposition to the reform and public transport was almost completely paralysed by strikes. This, however, turned out to be the high-water mark of the protests. In negotiations on 14-15 May, the government reached agreement with a group of union leaders on a number of changes in the first version of the reform - 19 amendments in all. The CFDT, the CFE-CGC and the CFTC accepted the deal. After concluding this agreement, the government also made concessions to teachers protesting its proposed education reforms. While not directly linked to pension reform, this helped take further momentum out of the protests. With the unions thus divided, the strikes and demonstrations subsided, albeit slowly in some sectors.

This was roughly the outcome the government had desired and expected all along: it was convinced that agreement could be reached with the CFDT and other moderate unions, thereby isolating the CGT and the much smaller FO. Nevertheless, the price of agreement with the CFDT was not trivial. The three most important concessions involved a rise in the guaranteed minimum pension, from 75 to 85% of the statutory minimum wage; the partial inclusion of civil servants’ bonuses in the pension system; and early retirement for workers who reached the MCP and whose careers started exceptionally early (aged 14-16). This last concession was, of course, a transitional one – the rise in the school-leaving age means that almost no one starts a career before 16 – but it proved unexpectedly costly. In addition, the CFDT and its allies won an increase in employers’
old-age pension contributions and better pension benefits for particular categories of
workers, such as those who have disrupted their careers to act as care-givers and those
who have passed through multiple pension regimes during their careers and are
disadvantaged at retirement as a result.48 They also secured management of the new
régime surcomplementaire created to manage pension contributions from civil servants’
bonuses. As in 1993, then, the cabinet traded off concessions to the unions on
management issues in return for changes to the parameters of the system.49 These
concessions, though significant, were entirely coherent with the overall logic of the
reform: they softened its impact on particular groups without contradicting its overall
aims. With the labour movement thus divided by the concessions to the reformist
confederations, the government proceeded to confirm the revised text of the law, which
was soon passed by the National Assembly and the Senate.

The government’s communications strategy

From the beginning, a great deal of thought was given to the government’s
communications strategy. In contrast to 1995, there were to be no surprises. The question
of how to sell the reform to the public was high on the agenda from the first internal
meetings in June 2002, and the reformers very deliberately chose to emphasise two
themes: the need to save the PAYG system, while sustaining relatively high levels of
pension provision; and the need to align the public- and private-sector regimes for reasons
Advertisements were placed in print media, focusing largely on high-circulation regional
papers;50 explanatory leaflets were distributed; and an internet site was established to
explain the reform.51 Two letters from the prime minister, setting out the case for reform,
were addressed to French households, and an open letter by the labour minister,
specifically addressed to those who doubted the reform, was published in early May.52
Mr. Fillon travelled to a number of European countries to “study” their pension systems,
with representatives of the social partners in tow. This, too, was a part of the
government’s public relations strategy rather than an exercise in policy formulation: the
government had largely made up its mind about the reform, and the foreign travel was
intended chiefly to provide an opportunity to sell the plan to the public at large and the
social partners.

After much discussion, the government decided to present the reform with a time
horizon running to 2020, despite the fact that the finance ministry’s internal projections
ran to 2040. This, too, was part of the communications strategy. The 2020 horizon was
chosen partly to avoid unproductive debates over uncertainty about the projections – the
estimates running to 2020 could be presented with greater pertinence than those for
2021-40 – and partly to underscore the fact that pension reform would not be a once-and-
for-all exercise. The government was also careful to treat long-term savings and
supplementary funded pension plans separately and in a low key. Like its predecessors, it
wished to encourage private supplementary savings schemes, but it kept this issue out of
the debates over the loi Fillon. The reform was aimed at preserving the PAYG public
system, and the government did not wish to confuse matters by talking about funded
schemes. The provisions of the law concerning the extension of the MCP were also
drafted with public-relations considerations in mind: thus, the authorities were very clear
about the rise in the MCP to 41 years by 2012, but they enshrined in the law a procedure
for any revisions beyond that rather than extrapolating too far out and risking popular
rejection of the idea. The notion of stabilising the ratio of contributing years to retirement
years was seen as saleable and was a key theme. It also allowed for the creation of a mechanism that would adjust the system’s parameters more smoothly and automatically.

Despite occasional missteps in its presentation of the case for reform, the government on the whole made its case consistently and coherently, addressing itself to both the social partners and the public at large. Polling data from the first half of 2003 suggests that it did indeed win the contest for public opinion on the substance of the issue but that it lost the politics. They show public awareness of the need for reform and of the urgency of the issue to have been high and rising. They also reveal broad – and in some cases, growing – support for aligning the public-sector schemes with the general regime and acceptance of the need to extend the MCP in order to avoid a sharp rise in contribution rates. Heavy majorities were prepared to acknowledge the government’s courage in tackling the issue. Nevertheless, support for, and trust in, the Raffarin government was below 50% throughout the spring, and on questions of trust, the government actually lost ground vis-à-vis the unions over the period of the pension debate.53

Despite this very comprehensive communications effort, critics of the government argued that much of the reform was nevertheless pursued “by stealth”. Some trade union officials believed that the government’s decision to focus the discussion on public-sector workers meant that the media paid rather less attention to the far-reaching technical changes that were being introduced in the system for private-sector workers. Many of the changes, such as adjustments to the formulae used for calculating initial pension benefits, were not well understood, and the emphasis on the key features of the system that were not changing – its PAYG character and the right to retire at 60 – may have led some private-sector employees to underestimate the impact of the reform on them. While such an approach may have made it easier to reduce future pension liabilities, it might also have muted for a time the impact of the reform on future pensioners’ choices: the microeconomic benefits of pension reform depend in no small measure on individuals’ awareness of changes in incentives.54 Nevertheless, while there may be some merit to these criticisms, it is difficult to see the loi Fillon as a case of reform by stealth, given the pre-reform work of the COR, which was already well known by 2003, the government’s information-provision efforts and the trade-union mobilisation against the reform.

Policy process: implementation

Implementation of the loi Fillon has proceeded smoothly since 2003 and has helped to establish broad acceptance of the need for longer careers across the economy. In October 2007, moreover, the law’s “automatic adjustment” mechanism was put into action for the first time, as a special commission examined the data on life expectancy at 60, preparing the ground for a decision in 2008 on further adjustment of the MCP. Broadly speaking, they confirmed the estimates contained in the loi Fillon, although the 2007 demographic data suggested that the MCP might have to rise by a little less than anticipated during the period to 2020. It is against this backdrop that the government’s effort to reform the special regimes for public enterprises in late 2007 must be seen. These regimes were relatively small, but there was a real concern within the government that further adjustment for the private sector and the rest of the public sector would be politically unacceptable – and manifestly unfair – if the public enterprises remained untouched. The actuarial adjustments scheduled for 2008 were designed to be more or less automatic, as they were enshrined in law, but there was a concern that they might be resisted, or even de-railed, if the special regimes were not reformed. At the same time, the reforms of 1993 and 2003 made the privileged terms offered by the special regimes more
conspicuous and thus more difficult to defend: as the pension schemes covering the rest of the workforce were revised, the unreformed minority found itself increasingly isolated.

Conclusion

Overall, the Raffarin government’s approach to pension reform was similar to that of the Balladur government, albeit somewhat more confrontational. The strategy of “consultation without negotiation” – apart from its engagement with the CFDT – underscores the capacity of political leaders to secure acceptance of pension reforms in the absence of conventional forms of concertation. While unilateral action on the model of 1995 is rarely successful in the field of pension reform, the progress of reform in France has involved a combination of confrontation and consensus-seeking. It is not difficult to see why unilateralism is rarely tried and typically fails. Given long implementation timelines, a degree of consensus is required if a reform is to be sustained. This is especially important, given that most pension reforms are implemented over long periods and thus tend to involve several governments, and the fact that the predictability and reliability of pension policies are in and of themselves important values for contributors making career/savings choices. Sharp swings in pension policy can be very costly to the individual. Nevertheless, the French experience suggests that the options confronting politicians are far richer than the stark choice between “concertation” and “unilateral action” that is presented in some scholarly writing. Success may require a readiness to out-manoeuvre or even override some opposition, as well as the building of the broadest possible consensus in support of reform.

Also important was the painstakingly careful policy process by which the loi Fillon was prepared and adopted. In this, it stands in stark contrast to both the CIP reform and the aborted Juppé pension reform of 1995. The role of the COR in achieving widespread agreement on at least some of the key issues was critical in preparing the public at large and the trade unions for reform, as was the technical and political preparation of the reform during 2002. While some significant concessions were made at a key point in the process, the government was broadly successful in ensuring that its major aims were not compromised and in setting the stage for the further gradual evolution of the system from 2008. The limited aims of the 2003 reform mean that much remained to be done to ensure long-term sustainability. The COR estimated that the 2003 reform had reduced the implicit debt of the public pension systems from roughly 150 to 100% of GDP, equivalent to a permanent reduction in the annual structural deficit of 1% of GDP. The combined effect of the 1993 and 2003 reforms was expected to keep the system solvent until 2020. To maintain system solvency thereafter, the government envisages raising the pension contribution rate while lowering the rate of unemployment insurance contributions, so that the total social security contribution rate will remain unchanged. However, this solution depends on bringing unemployment down sufficiently to allow such a reduction in unemployment insurance contributions.

When the loi Fillon is seen against the backdrop of the Balladur and Juppé reform episodes, as well as the subsequent struggle over the special regimes in 2007, it becomes clear that the whole process of pension reform has proceeded in stages, beginning with the largest but most easily reformed segments of the system and working gradually towards those sectors where resistance is likely to be fiercest. As a result, acceptance of the necessity and, indeed, inevitability of reform has gradually spread, while support for those who continue to resist has shrunk. Public support for the strikes in opposition to pension reform fell significantly from 1995 to 2003, and fell even further in 2007. In part, this simply reflects the fact that the unreformed minority has grown smaller and its
privileges have thus become less acceptable in terms of equity, but it also reflects a change in public attitudes towards pension reform as the result of more than a decade of discussion and debate.

Notes

2. The Socialist government of 1997-2002 also continued the trend towards restricting access to early retirement and increasing incentives for private saving. The share of 55-60-year-olds in various state-financed early retirement schemes fell from 16% to 10% over 1996-2002 (OECD 2003:71-3).
3. There were also important differences of opinion within both the Socialist Party and the new centre-right Union for a Popular Majority.
5. The unions attach considerable importance to their co-management role in these schemes. See Rosanvallon (1995).
6. See Visser (2000). By the mid-1990s, around one-quarter of French union members were inactive and most of these were pensioners.
7. See, in particular, TNS Sofres’s Baromètre des retraites (http://www.tns-sofres.com/etudes/pol/170603_retraites_r.htm).
8. Some CGT leaders were more open to compromise than either the constituent unions or the rank and file, not least because they feared that, in the wake of the reform of the private-sector regime, too strong a defence of public-sector pensions would risk turning the CGT into a purely public-sector confederation.
9. For details, see Chanut et al. (2003a, 2003b).
10. It is important to note what this fragmentation implies with respect to the impact of demographic change: in some instances, the financial health of schemes depends on the demography of the profession rather than national demographic trends. Where employment is falling and/or workforces are ageing, these systems come under greater strain than a look at France’s overall demographic outlook might suggest.
11. Some clarification of terms is needed here. The usual French translation of PAYG financing is financement par répartition. This can be misleading, however. Some public pensions, it should be noted, are paid directly from the state budget rather than from payroll taxes. Thus, they do not fall under the French notion of répartition but they would nonetheless be considered PAYG in English, as they are not funded.
12. See, for example, Schöpflin, (1987); États généraux (1987); INSEE (1990); Commissariat général au Plan (1991); Brunhes (1992). These discussions were influenced in part by OECD work highlighting the need for reforms to ensure financial sustainability. See, for example, OECD (1992), Van den Noord and Herd (1993), Leibfritz et al. (1994, 1995), and Roseveare et al. (1996).

15. The harmonisation of a wide range of different regimes can be more difficult in the context of an essentially parametric reform; in Italy, one of the motivations for the shift to a defined-contribution system in 1994 was the awareness that harmonisation would be easier to achieve in the context of a structural reform that changed the essential nature of the system itself, and not merely some specific parameters.

16. Myles and Pierson (2001:320) note that by 2000, a number of OECD countries had adopted some form of selective targeting to reduce formerly universal flat-rate benefits for high-income retirees, but no country had means-tested earnings-related benefits that were the object of quasi-property rights.

17. The retirement age was lowered from 65 to 60 in 1983 in a misguided effort to combat rising unemployment.

18. See also Palier (1999, 2002).

19. Moreover, the unions themselves were extremely suspicious of the Socialists’ own pension reform proposals, which had been unveiled in January 1993, ahead of the elections. The government of Pierre Beregovoy rejected any extension of the MCP or increase in the retirement age, proposing instead to create a new pension fund. Cf. Le Monde, 15 January and 12 February 1993.

20. Union density was around 4% in the private sector but over 10% in the public sector (Labbé 1996, cited in Natali and Rhodes 2004). Around 35% of the total CGT and FO membership in 1993 was in the private sector.

21. As employees of the state, they also have less to fear from unemployment and lack of competitiveness arising as a result of very high payroll taxes. See James and Brooks (2001) and the case study of Mexican pension reform.

22. See, in particular, Bonoli (1997); and Da Conceição-Heldt (2007).


24. The Fonds de solidarité vieillesse.

25. James and Brooks (2001) find concessions to trade unions with respect to the management of pension systems are often used to reduce their suspicions of, or opposition to, pension reforms.

26. Bonoli (2001:257) quotes a French official involved in the 1993 negotiations as follows: “It was important for us to gain the approval of the CFDT because we knew that FO and the CGT would be hostile anyway… We needed at least the neutrality of the other confederations”.

27. Le Monde, 8 May 1993; the FO and CGT took the view that the reform would strip workers of acquired rights and was therefore unacceptable.

28. Levy (2000:336) refers to this as a throwback to the “insular, technocratic methods of the 1960s”.

29. This element of the Juppé package, which required a constitutional amendment, was implemented.

30. See Le Monde, 17 November 1995, for union reactions on this point. The leader of FO called the budget proposal “the biggest theft in the history of the French Republic”.

31. See Charpin (1999). Natali and Rhodes (2004:10) confirm that the unions were well aware that the impact of the Juppé plan on those currently working would have been far greater than that of the Balladur reform.

33. Teulade’s calculations came under fire on a number of grounds, but the most important was his failure to take account of the “noria effect”: even if pensions are indexed to prices rather than wages, the pension at retirement is salary-based, so each succeeding generation retires with a higher level of pensions than its predecessors. The Teulade report failed to integrate this fairly elementary point into its projections. See COR (2000); and Débats (2000:7484-85).

34. Perhaps the most influential of these was Taddéi (2000), which argued for a progressive system of transition to retirement and suggested that late entry into the labour market meant that many workers could not work 40 years before reaching retirement age. Others challenged the idea of a demographic shock and argued that higher contributions could be financed partly out of future productivity gains, and that ageing would reduce expenditure on children, youth and the unemployed – generating savings that could go towards pension expenditure.

35. The **Fonds de réserve des retraites** accumulate excess pension contributions, part of the income generated from state property, investment income and some income from one-off sources such as privatisations. As of 30 September 2007, these assets stood at EUR 33.8 billion.

36. See Palier and Bonoli (2000) and Palier (2007). Many French unions remain wary of funded schemes of any kind, but there are exceptions. In the late 1990s, CFDT Secretary General Nicole Notat argued that “European and French workers must ask themselves if they will continue to let Anglo-Saxon pension funds … continue to have the monopoly over intervention in the capital of French and European enterprises” (**Libération**, 14 September 1999). Former finance minister Michel Sapin (2000) argues for the creation of pension savings funds inside firms to reinforce workers’ power.

37. See also the polling data presented in Palier (2002:12) and Palier and Bonoli (2000), which point to a strong sense of the inevitability of reform, mainly for demographic reasons.

38. For details, see Babeau (1997) and Palier (2002, 2007).

39. The **loi Thomas** aimed to create a pension savings system for 14.5 million workers in the private and agricultural sectors, financed by member employees and eventually by employers who wished to contribute. The scheme would have been made attractive by exemption from taxes and social charges. However, implementation of the law was blocked by the Jospin government, which came to power just weeks after its adoption, and the law was repealed in 2001.


42. See Franco (2003).

43. There remain, however, derogations and exceptions in public and private sectors, *e.g.* for individuals in strenuous work or professions where agreement within the profession has been reached on an earlier age.

44. In other words, the benefit amounted to the replacement rate times the reference salary times a coefficient which was the ratio of actual contribution years to the MCP.

45. The Balladur reforms were also phased in over a decade.

46. “100 minutes to convince”.
47. The government believed it would cost about EUR 1.2 billion per year to implement, but it has in fact been running at more than EUR 2 billion. The estimates of the number of individuals with particular contribution histories on which both the government and the unions were relying simply turned out to have been wrong.

48. These so-called pluripensionnés tend to suffer losses relative to those with more stable careers. As fewer and fewer people make their entire careers in a single sector or institution, such poly-pensions are becoming increasingly common and remain a major item on the agenda for any future pension reform.

49. Strictly speaking, not all these concessions involved changes in the actual text of the loi Fillon: many involved changes in the secondary legislation adopted pursuant to the law itself.

50. The circulation of France’s major regional newspapers dwarfs that of the internationally better-known Paris-based “national” dailies like Le Monde and Le Figaro.


52. Many of the materials from the government’s press campaign, including the flyer, the letters from Raffarin and Fillon, and numerous documents addressing specific groups or issues, can still be found at the government’s web site on retirement reform, http://www.retraites.gouv.fr/.

53. When asked simply about extending the MCP, heavy majorities rejected the idea, but it was by far the preferred option when presented alongside other options, including higher contributions and lower benefits. For the most detailed data on the issue, see the Baromètre des retraites conducted by TNS Sofres (http://www.tns-sofres.com/etudes/pol/170603_retraites_r.htm), and the Baromètre social of CSA (http://www.csa-fr.com/dataset/data2003/opi20030619a.htm). See also the data in Chanut et al. (2003b).


55. Cf. the reversal of the pension reforms adopted by the CDU/CSU government in Germany after it lost the 1998 general election, or the repeal of the loi Thomas in France following a change of government just months after it was enacted.

56. By contrast, frequent changes in benefit regulations may have deleterious consequences but these are unlikely to be as damaging as instability in respect of pension provision.

57. See, for example, Myles and Pierson (2001:306).

Bibliography


INSEE (1990), *L’avenir des retraites*, published as a special issue of *Économie et statistique*, No. 233, June.


Chapter 4

Italy: The 1994 pension reform proposals

Summary and main conclusions

In the autumn of 1994, the centre-right government presented a package of pension reform proposals as a part of its budget package for the following year. The primary aim of the package was to help the government meet its deficit-reduction target for 1995 and to do so via structural changes that would add to the credibility of Italy’s medium-term fiscal strategy in a way that ad hoc cuts and postponement of expenditure would not. Though far less radical than the pension changes adopted just two years earlier, the proposals triggered two months of large-scale protests and divided the governing coalition. In the end, the reform was withdrawn and the government fell. Nonetheless, the lessons learnt probably helped pave the way for a major pension reform in 1995.

The defeat of the 1994 pension reform plan reflected a number of factors:

- **Coalition fragility.** The Berlusconi cabinet was the first “political” government to take office following a period of political and financial crisis; its members had little experience in government and its unity was fragile. The coalition was under considerable strain even before the budget cycle began, and the pension proposals aggravated matters considerably, since they would have hit the electorate of the Northern League (LN), one of the coalition parties, rather hard.

- **Unilateral approach.** Opponents of the reform objected not only to the substance of the proposals but also to the attempt to impose them as part of the budget law, with only minimal consultation with other actors. While some in government feared that a concertative approach would allow the unions to dilute the substance of the reform, the strategy the government adopted appears to have reinforced opposition to its proposals. The trade unions, in particular, were committed to defending the “concertative” approach to reform, involving negotiations between the government and the social partners, which had prevailed under the technocratic governments of 1992-93.

- **Uneven distribution of costs.** The burden of reform would have fallen overwhelmingly on dependent employees – the trade unions’ core constituency and, in Northern Italy, an important source of support for the LN. The government proposals hardly touched traders, farmers, craftsmen and other self-employed persons, groups that had tended to support the dominant coalition party, Forza Italia, in the 1994 elections. Moreover, whereas younger cohorts had borne almost all the costs of the 1992 pension reform – and would also bear the costs of the subsequent Dini reform – the 1994 reform proposals would have imposed significant losses on older cohorts. This reflected the dual purpose of the reform, which aimed not only to enhance long-term sustainability but also to...
generate substantial immediate fiscal savings, and it ultimately proved politically unacceptable.

- **Inadequate discussion and communication.** Having appointed an expert commission to study pension reform, the government ignored its proposals and presented a package that took most observers by surprise. Despite an awareness of the likely resistance to reform, moreover, the government had given little thought to how it could best win public support for the measures it proposed. Finally, the coalition’s internal divisions over pension reform became public during the discussion of the budget proposals, making it impossible to present a clear, unified message in support of the reform.

**The context for reform**

**Framework conditions for reform**

The Berlusconi government was the first “political” government since the corruption scandals and economic crisis of 1992 had effectively wrecked most of the political parties that had dominated Italy for the previous half-century, as well as the careers of many leading politicians. The country had been governed during 1992-93 by technocratic cabinets. Despite its victory at the polls in April 1994, however, the new government was politically vulnerable from the start. While the coalition commanded a majority in the Chamber of Deputies, it was two seats short in the Senate, where it had to rely on non-aligned votes. It was, moreover, an awkward coalition, bringing together the prime minister’s newly created *Forza Italia* (FI), the right-wing National Alliance (AN), the Lombardy-based Northern League (LN), and the Christian Democratic Centre (CCD), one of the parties that had emerged from the wreckage of the old Christian Democratic Party.

Although FI had contested the elections in alliance with its coalition partners, the centre-right had agreed no common programme and it took six weeks to form a government. Relations between the AN and the LN were very poor,¹ and the task of maintaining government unity was further complicated by the fact that most of the coalition parties, like most of the ministers themselves, were new to government. FI was from an early stage in public conflict with the LN, whose leader, Umberto Bossi, had refused to join the cabinet. Nevertheless, there were good reasons to expect that the coalition would hold: FI was riding high in the polls over the summer, with its national vote total rising from 21% in the March general election to 30.6% in the June elections to the European Parliament. Both the AN and the LN, by contrast, were losing ground. Given the widespread expectation that the fall of the coalition would trigger an election, it was generally believed that the AN and the League would stay in government rather than risk a snap election that would weaken them and strengthen FI.

The government took office just as the recovery from the recession of 1992-93 was gathering momentum: real GDP growth exceeded 3% in the first half of 1994. The government’s principal macroeconomic challenge was to put Italy’s public finances in order: in 1992-93, net borrowing by the general government exceeded 9.5% of GDP and the public debt/GDP ratio reached a record 125.1% in 1994. With the 1992 financial crisis still a recent memory and the Maastricht Treaty setting Europe on the path to Economic and Monetary Union (EMU), fiscal consolidation was bound to be the top priority for any Italian government, of right or left, in the mid-1990s. For 1995, the government’s aim was to bring net borrowing down to 8% of GDP, which would require
a primary surplus of 2%. While fiscal consolidation presented a very tough challenge, widespread support for European integration made the politics of fiscal consolidation rather less intractable than they would otherwise have been. Thus, while Maastricht limited the government’s freedom with respect to both fiscal and monetary policies, it also enlarged national decision makers’ room for political manoeuvre.

The government’s own early actions arguably served to strengthen the external constraint, adding to the pressure for rapid fiscal adjustment that lay behind the 1994 pension reform proposals. During the election campaign, FI’s major emphasis had been on stimulating job creation and economic growth, and it had promised both lower taxes and higher pensions. FI leaders had also expressed the view that the Delors plan for EMU was misconceived, and the new foreign minister was overtly Eurosceptic. The government’s perceived Euroscepticism, combined with its commitment to tax cuts, thus raised doubts about its readiness to continue the process of fiscal consolidation. When combined with monetary tightening abroad and signs that the coalition was unstable, these concerns contributed to a currency “mini-crisis” over the summer. Thus, by late August, the government was under more pressure than ever to demonstrate its commitment to fiscal adjustment.

**Actors and interests**

One consequence of the breakdown of the party system and the long period of government by non-partisan technocratic cabinets was the resurgence of concertation, involving the government and social partners, as a mode of policy making. Italy’s labour and business lobbies alike are dominated by several large national institutions – above all, the three big union confederations and the employers’ body Confindustria – and these had moved to fill the vacuum left by the collapse of the party system in 1992. Concertation was nothing new in Italy, but the discrediting of the old established political parties and the initial failure of any new ones to take their place gave it renewed importance. Major policy innovations, including critical elements of Italy’s post-crisis economic adjustment, were negotiated in more or less formal tripartite fora and set out in “social pacts”. Technocratic governments thus operated with relatively little parliamentary input but with the extensive involvement of the social partners, particularly the trade unions. They relied less on partisan support than on persuading the social partners that the costs of non-adjustment were vastly greater than the costs of reform and on negotiating with them over the distribution of those costs. Once the government and social partners reached agreement on new measures, parliamentary approval tended to follow relatively easily.

The advent of the Berlusconi cabinet, however, marked the resumption of party government, albeit on the basis of new parties. Having an electoral mandate, the government did not intend to rely on the social partners as its technocratic predecessors had done. This presented a challenge to the trade unions, which had grown attached to the concertation practices of the previous two years and were eager to defend their role in the policy process. It also presented a challenge to the coalition, as it sought to manage the tensions among its constituent parties. In the case of pension reform, these tensions were fairly acute, owing to their rather different constituencies. The AN, with its strength in the economically sluggish, low-employment south, was anxious to protect disability pensions and other measures that enabled the pension system to function as a mechanism for passive labour-market adjustment and regional redistribution. By contrast, the LN was mainly concerned about the impact of planned changes on northern industrial workers, who were an important part of its constituency. Finally, FI was keen to shelter its own
voters, particularly among the self-employed and the professions, from the impact of reform.

Although union density has been falling for some time in Italy, levels of unionisation in the early 1990s were still rather high by Continental European standards. In 1994, Italy’s three major union confederations represented around 39% of dependent employees, down from a peak of just over 50% in the late 1970s: the Confederazione Generale Italiana de Lavoro (CGIL), which represented around 17% of employees in 1994, the Confederazione Italiana dei Sindacati dei Lavoratori (CISL), representing roughly 12%, and the Unione Italiana del Lavoro (UIL), with about 5%. A feature of Italian unions – and one that may have strengthened their ability to participate in corporatist policy making – is their relatively democratic internal decision making: while very well able to mobilise protest, they can also “organise consent”. Thus, union referenda were used to ratify both the July 1993 pact and the 1995 pension reform. The confederations’ role was also entrenched in law to some extent: until 1995, the law gave them de facto power to represent all workers. However, in one critical respect, the unions by the 1990s were growing steadily less representative of the Italian workforce: a majority of CGIL (53.7% in 1995) were pensioners; and CISL had only a bare majority of active workers (54.1% in 1995). UIL had by far the most active membership (75.2% in 1995). Moreover, the pensioner share in total membership was rising, as the unions’ membership aged. Many unions had special pensioners’ sections. The ageing of the trade unions’ constituency has had an enormous impact on the politics of pension reform in Italy.

Though present at the table when tripartite talks took place, the various organisations representing employers and the self-employed played a much less important role in the pension reform debates of 1992-95. In general, they were less attached to concertation for its own sake and more in tune with government reform proposals, being largely concerned to support measures that would limit any future rise in contribution rates. However, they had only a limited impact on public or government opinion.

**Issue history**

**Italy’s pension problems**

By the end of the 1980s, Italy had perhaps the most generous pension system in the developed world (Box 4.1). It was also widely recognised as unsustainable over the long term: the effective retirement age was falling, even as life expectancy rose, and the population as a whole was ageing fast. These challenges to the system were compounded by sluggish growth and a secular increase in unemployment. In the absence of reform, the ratio of pension spending to GDP, already the highest in the OECD, was expected to hit 23% by 2035, with the old-age dependency ratio rising from just under 30% to almost 65% over the period to 2050. Even with contribution rates rising from 38 to 50%, the direct impact of this ageing process would, ceteris paribus, have entailed an 18-point drop in the average replacement rate, to 55.5%. The cost of the system was increased by the widespread use of early retirement and invalidity pensions to deal with labour-market adjustment, and the tendency for non-contributory pensions, which were a form of social assistance, to grow in both coverage and generosity of benefits. Even with payroll tax rates well above those found in France and Germany, large and growing transfers from general taxation were needed to cover system shortfalls by the early 1990s.
Box 4.1. Italy’s pre-reform pension system

At the beginning of the 1990s, Italy’s public pension system was a highly fragmented, occupational pay-as-you-go (PAYG) scheme financed by employers’ and workers’ contributions. The complexity of the system (there were 17 schemes for salaried workers, according to the Treasury, plus 3 for the self-employed and 11 for the professions) was largely the product of ad hoc lobbying and changes over time, as occupational groups scrambled to secure ever better deals for themselves. Early on in the reform debate, Castellino (1975) wrote of a “pensions labyrinth” – and the system actually grew more complex later on.

Costs and benefits were extremely unevenly distributed across generations and occupational categories, as both contribution rates and benefit formulae varied widely. Civil servants, railway workers, local authority workers and the armed forces, among others, enjoyed particularly generous treatment. Benefits were generally linked to earnings during the best period of one’s career. As a rule, the benefit amounted to 2% of the average salary over the last ten years of one’s career for each year of contributions (in some cases the assessment period was far shorter). This coefficient was very generous, even when compared with the levels then prevailing in Germany (1.50%) and France (1.86%). Moreover, benefits in payment were indexed to wages; in many cases, individuals’ pension indexation was linked directly to the salaries of current employees in the occupations in which the retirees had worked.

Perhaps the most striking peculiarity of the pre-reform system was the “seniority pension” (pensione di anzianità), which allowed workers to retire well before the official retirement age, provided that they had contributed for a certain number of years – the minimum contribution period (MCP) varied from less than 15 years for women with children working in the public sector (20 for men in the public sector) to 35 years for industrial workers. Not surprisingly, some 85% of civil servants retired ahead of the statutory retirement age. Largely as a result of the prevalence of seniority pensions, the average effective retirement age in Italy by 1992 was just 58. Retirement ages, moreover, had been falling for decades despite rising life expectancy. Participation rates for the 55-64 age group in the years leading up to the 1992 pension reform hovered around 32.5%, as against an OECD average of over 50%.*

* The gap between men and women was particularly large: while almost 52% of Italian men remained in the labour force between the ages of 55 and 64, only 15.1% of women did so. For the OECD as a whole, by contrast, the gap was just about 3 percentage points. See OECD (2007:32).

The unsustainability of Italy’s pension system had been recognised as early as the 1970s, but the 1980s saw only inconclusive discussions and tinkering with the system, some of which served to undermine, rather than to strengthen, its financial solidity.12 Numerous commissions were convened, official reports were issued and occasional pension reform proposals were floated by various governments. All of this served to build an awareness of the need for action, but “pension reform” in practice was limited largely to increases in contributions.13 By 1994, however, the gap between pension outlays and contributions projected for the following year was 84 trillion lire or about 5% of GDP, equivalent to somewhat over half the total projected general government deficit.14

The search for solutions

In response to the financial crisis of 1992, the Amato government succeeded in adopting a far-reaching package of adjustments to the parameters of the pension system (Box 4.2). Altogether, the reform reduced workers’ net pension wealth by an estimated 47.1%.15 Moreover, while most of its measures affected only future retirees and many touched only those with fewer than 15 years of contributions in 1992, the switch to price-based indexation of pensions hit current retirees. Fifteen years after retirement, the initial
pension was projected to lose over a third of its value relative to current wages.\textsuperscript{16} Even so, three-quarters of the cost of the reform fell on those aged 44 years or under (the age of the median voter in 1992). Employees under 45 lost 76.7\% of their aggregate pre-reform net pension wealth.\textsuperscript{17} Indeed, those under 30 were left with negative net pension wealth. By contrast, the 45+ cohorts lost 16.8\% (3.4\% for the over-60s).\textsuperscript{18}

Box 4.2. The 1992 pension reform

The major elements of the 1992 “Amato reform” included a phased increase in retirement ages;\textsuperscript{1} the assessment of the reference wage on the basis of the entire contributing career (rather than just the last ten years) for those with fewer than 15 years of contributions in 1992; the harmonisation of accrual coefficients across most schemes;\textsuperscript{2} an increase in the MCP for seniority pensions in the public sector from 20 to 35 years; an extension of the MCP for an old-age pension from 15 to 20 years; an increase in contribution rates;\textsuperscript{3} the indexation of benefits in payment to prices rather than wages; and a temporary suspension of the right to retire on seniority pensions. The government also passed legislation creating new options for private supplementary schemes.\textsuperscript{4} Although it made seniority pension thresholds for public and private employees more homogeneous and reduced some of the disparities in the benefit-calculation criteria of different schemes, the 1992 reform was essentially a parametric exercise that left most of the system’s inequities and technical incongruities untouched.\textsuperscript{5} The reform also concentrated on dependent workers in the public and private sectors, leaving the pension schemes covering the liberal professions and the self-employed, which had always run structural deficits, unchanged.

The underlying problem of seniority pensions was not addressed – even the relatively modest proposal to raise the MCP for a seniority pension from 35 to 36 years was blocked. This led to a perverse result: because the reform tightened up most system parameters, while leaving seniority pensions untouched, it increased the incentives to retire early. The Bank of Italy estimated that the net present value of pension payments received by a worker retiring at 65 after 43 years’ work to be 23\% below those of the same worker retiring at 57 after 35 years, all other things being equal.\textsuperscript{6} Despite these limitations, however, the Amato reform represented an important watershed in pensions policy and a crucial first step towards fiscal and pension sustainability.

1. From 55 to 60 for women and from 60 to 65 for men, by 2002.
2. Some powerful occupational groups secured exemption from this.
3. To 26.5\% for employees (employer contributions cover two-thirds) and 15\% for the self-employed.
4. These allowed for the possibility of using the future termination indemnity (Trattamento di fine rapporto, or Tfr), to finance them: the Tfr is essentially a deferred wage – money retained by the employer until retirement or severance and then paid to the departing employee, traditionally as a lump sum. It amounts to 7.4\% of annual earnings for each year of employment, revalued at a fixed interest rate of 1.5\% plus 75\% of the inflation rate.

While there was no formal concertation process, and no agreement with the trade unions and Confindustria was ever signed, passage of the reform was facilitated by extensive tripartite negotiations, which continued until a bill granting the government the power to reform pensions by decree was adopted in October 1992. The unions acquiesced, even though the cuts involved were more dramatic than had been anticipated just a few months earlier. Three factors contributed to this outcome. First, negotiations took place against the backdrop of the intense financial crisis of September 1992, which
drove home to all parties the potentially enormous costs of non-adjustment. The unions’ willingness to accept short-term cuts, in particular, increased as a result of the crisis. Secondly, certain highly visible features of the old system were left alone, including the 2% annual accrual rate and the 35-year MCP for seniority pensions. Thirdly, the cuts were conceived in such a way as to make themselves felt only over a relatively long period: as noted above, many changes affected only younger cohorts, and even the change to inflation indexation would have little immediate impact.

Crucial though it was, the Amato reform was only a first step. In 1993, a report issued by the National Institute for Social Protection (INPS) concluded that the 1992 reforms were insufficient to ensure the system’s sustainability, particularly in respect of the self-employed. Indeed, projections for their schemes were deteriorating, as a result of a 1990 decision to extend the earnings-related, defined-benefit system to these groups. The projections looked far better with respect to dependent employees, whose pensions were the focus of the Amato package. Yet even with respect to private-sector employees, there was a growing awareness that further reform would sooner or later be required.

Shifts in public opinion

By the early 1990s, the Italian welfare-state model was widely considered both unsustainable and unfair. Surveys of public opinion suggested a marked change in Italian attitudes towards the welfare state from the mid-1980s to the mid-1990s: an increasing share of respondents took the view that taxes were too high, and support for benefit levels declined. This clearly suggests that the climate of public opinion was open to reform in 1994, and pension and welfare reform were indeed much-debated topics during the election campaign that year. However, the government could not claim any real mandate for pension reform. First, neither Forza Italia nor its coalition partners had presented any clear plans on this point – the centre-right’s campaign was focused on tax cuts, job creation and growth, and was in any case short on specifics. Indeed, FI even promised higher pensions, albeit without explaining how they would be financed. Secondly, as noted above, the coalition parties had no agreed programme, a fact that LN leaders later cited in justifying their decision to oppose pension reform proposals while remaining in government.

Policy process: formulation

The preparation of the pension reform package

The outcome of the pension reform attempt of autumn 1994 owed much to the manner in which the reform was prepared. Initially, the Labour and Social Protection Ministry took charge of the pension reform portfolio, and in July, an 18-member ad hoc commission was constituted under its auspices. Chaired by Onorato Castellino, the commission issued its main report in August, recommending, inter alia: the introduction of an actuarial coefficient linking retirement age to the level of the final pension and a revision of the arrangements for seniority pensions, to discourage very early labour-force exit. One option was to adjust benefits using a formula based on the sum of contribution years and age at retirement. However, neither the commission nor its findings received much publicity, and the report was not published. The commission’s work coincided with an informal discussion of pension reform involving the unions and Confindustria, which continued through the summer and into September. This discussion made little progress, and some in the government began to unsettle the unions with talk of radical, unilateral
measures. At the time, many thought that such threats were intended to put pressure on the unions, but at the end of September, the government unveiled a package of far-reaching pension-reform proposals as part of its draft 1995 budget, as well as a bill giving itself authority to reform pensions by decree.

The government package bore little resemblance to the Castellino Commission’s recommendations. The proposals put forward were drawn up by the Treasury in consultation with the Prime Minister’s Office, with relatively little input from other parties (including the Ministry of Labour and Social Protection), and they were aimed chiefly at achieving immediate fiscal goals. This emphasis on fiscal consolidation as the motivation for the reform subsequently attracted much criticism: most pension reforms tend to bring fiscal benefits only over a long period, since any changes to pension parameters and entitlements that are drastic enough to generate rapid financial improvement are likely to impose unacceptable losses on current pensioners and older workers. The Treasury, however, saw little choice but to make pensions a key element in its fiscal package. First, pensions alone accounted for around 30% of general government expenditure, and the projected deficit of the INPS amounted to more than half the projected general government deficit for 1995. It would have been impossible to engineer a major reduction in the deficit without touching pensions. Secondly, permanent structural changes to the pension system would offer fiscal benefits not only in 1995 but for many years to come, and it was hoped that such changes would add to the credibility of Italy’s medium-term fiscal strategy in a way that ad hoc cuts and postponement of expenditure would not. The budget proposals therefore set a target of holding the rise in pension spending to 3.5% in 1995 and thus ensuring that reductions in pension expenditure generated 20% of planned fiscal savings for the year.

To achieve this end, the government’s budget team proposed the following. First, the MCP for a seniority pension was to be raised rapidly from 35 to 40 years. Secondly, the annual accrual factor was to fall from 2.0 to 1.75% for workers with more than 15 years of contributions. This would still have been higher than the level then prevailing in Germany (1.5%) and not far below that of France (1.86%), but it aroused particularly strong opposition, because it targeted older workers – precisely those cohorts that had been powerful enough to escape most of the effects of the 1992 reform. In addition, the indexation of pensions to prices was to be replaced by indexation to projected inflation; the increase in the retirement age legislated under Amato was to be accelerated; and the right to retire with seniority pensions was to be suspended in 1995, as it had been in 1993. The pensions package failed to address many of the concerns raised by the unions during the Castellino commission’s deliberations, including better coverage for women and other individuals with non-standard careers, and a clearer distinction between social insurance and social assistance. D’Amato and Galasso (2002) estimate that, if fully implemented, the proposals would have reduced the value of pension wealth by 9% in aggregate and that the employment categories managed by the Pension Fund for Employees would have suffered a 27.5% reduction.

**Short-term fiscal pressures and the design of the reform**

After more than a decade of serious study and public debate, there was wide agreement, at least among experts, on the major problems to be addressed: longer careers, a stronger link between contributions and benefits, and greater equity. The mounting burden of non-contributory pensions on the system also pointed to the need for a clearer distinction between social insurance and social assistance, something for which the unions had long lobbied. The most difficult issues were not technical but political:
pension reform was chiefly about the distribution of losses. In the case of the Amato reform, this had been possible largely because, in the midst of the political and economic crises of 1992, the choice facing agents seemed to be not so much between reform and the defence of the status quo as between reform and financial meltdown.\footnote{33}

While conditions in the autumn of 1994 were less extreme than in late 1992, pressure from financial markets over the summer greatly increased the sense of urgency about shoring up public finances. Indeed, a number of prominent economists from across the political spectrum signed a letter in support of the government proposals, describing them as the only “structural intervention” in the 1995 budget.\footnote{34} However, the need to address immediate fiscal pressures shaped the government’s reform plan in ways that made it much more difficult politically. In trying to make substantial savings in the near term, it adopted a package that spread the costs of reform more widely across different age cohorts: D’Amato and Galasso (2002) estimate that the 45-and-over age group would have borne around 53.5\% of the cost imposed on dependent employees by the reforms proposed in 1994, as against only 25\% in the case of Amato in 1992.\footnote{35} The 1994 proposals, however, would have largely exempted traders, craftsmen, farmers and other self-employed persons. The Amato reform, too, had largely left these groups alone, but the political significance of so doing was different in 1994, since these were all groups among whom FI had enjoyed disproportionate support in the March elections.\footnote{36} The government was very quickly accused of trying to solve the system’s financial problems at the expense of its opponents.

Policy process: adoption

\textit{Discussion of the reform}

At the end of September, the prime minister presented the cabinet’s budget proposals to trade union leaders. They reacted with hostility to the entire package, which they saw as hitting older employed workers – the most heavily unionised segment of the workforce – hardest of all. The meeting with the unions was followed by a cabinet meeting at which the proposal to raise the MCP for seniority pensions was dropped. In its place, the government proposed to penalise individuals taking seniority pensions before reaching the statutory retirement age in force for the Private Employees’ Pension Fund: such early retirees were to suffer a 3\% cut in benefits for each year of retirement ahead of the legal retirement age. This concession proved wholly unsatisfactory to the unions, who regarded it as tantamount to the abolition of seniority pensions. Moreover, a good deal of critical attention was focused on the government’s attempt to push such a far-reaching pension reform through on the back of the budget, which by law was subject to a very tight parliamentary timetable that would limit the scope for debating or revising the proposals. President Oscar Luigi Scalfaro was among those who objected to this method, and he declared that he would refuse to sign the budget until certain parts of the decree on pension reform were separated from it. The President insisted that elements of the reform required full parliamentary debate and should not be subject to the 31 December deadline that applied to the budget.\footnote{37} The government reluctantly complied, though it kept most of the pension reform measures bundled with the 1995 budget.

The confederations saw the budget proposals as a spectacular violation of the acquired rights of more mature cohorts of workers, who were the unions’ core constituency and who had been largely shielded from the Amato reform. The unions also took strong exception to the manner in which the reform was presented. By acting unilaterally, the government had broken the unwritten rules of the game that had been
emerging since the collapse of the first republic, according to which any major reform should be negotiated with the social partners \textit{ex ante}.\textsuperscript{38} The unions thus saw the conflict as a defence of concertation in its own right.\textsuperscript{39} The confederations thus began organising nation-wide protests. On 14 October, they called a general strike, during which an estimated 3 million people participated in protests in some 90 cities. On 12 November, a second general strike brought around one million people onto the streets of Rome, one of the largest demonstrations in a decade. Confindustria and other employers’ groups came out in support of the government, but their backing counted for little when set against the unions’ mass mobilisation.

The mounting protest exposed fractures in the governing coalition, a fact that merely encouraged the unions in their resistance. The LN was particularly nervous about the treatment of seniority pensions because industrial workers in the north of Italy – a key source of electoral support for the League – were far more likely than most workers to retire early. The LN’s advisors estimated that around three-quarters of retirements on seniority pensions took place in the north of the country, because southerners were far less likely to have long, uninterrupted contribution careers. The League had resisted the initial proposal for a 40-year MCP for seniority pensions, but as opposition to the reform escalated, it became clear that the proposed 3\% deduction was just as unpopular. By mid-October, the LN was publicly distancing itself from the pension reform plan, and in November, the LN Budget Minister publicly stated that it would be a mistake for the government to use confidence votes to push the reform through parliament. Moreover, the LN’s public position with respect to the coalition itself was increasingly ambiguous. Nor was the LN the only source of trouble for the government: some AN members began tabling amendments of their own, and the public statements of the Labour and Social Protection Minister, who represented the CCD, differed in tone from those of his Treasury counterpart, suggesting to some observers that the CCD was eager to compromise.\textsuperscript{40}

The government nevertheless remained determined to press ahead. The Treasury strongly rejected proposals to separate pension reform from consideration of the budget, fearing that such a move would secure passage of the budget at the price of leaving the pension reform blocked indefinitely. The Chamber of Deputies took up the budget, including the pension reform package, in mid-November, with the government calling three confidence motions in quick succession in order to force through the most controversial measures: the proposed amnesty for those guilty of tax evasion in the property sector; the accelerated increase in the statutory retirement age and the 3\% deduction rule applied to seniority pensions. On 17 November, however, the government failed to use the confidence vote to force through other parts of its pensions proposals. No explanation was given for this change of tactic, but without the discipline of a confidence motion, the LN defected, supporting opposition amendments to restore the 2\% annual accrual rate and full indexation of pensions to actual inflation.

For the trade unions, these victories were not enough, and they threatened a further general strike for 2 December. In the interim, the government, which was still trying to negotiate with the unions, suffered a significant setback in the first round of municipal elections. FI’s share of the vote plunged to just 8.4\% nation-wide, behind the Party of the Democratic Left, the AN and the Italian Popular Party, the main successor party to the former ruling Christian Democrats. While the LN slumped further, falling to 4.5\% nationally, the collapse of the FI vote meant that it could no longer expect to win a snap election if the government fell. This affected all parties’ calculations, as they considered whether and how to press for a change of government. As the strike approached,
therefore, FI became increasingly eager to reach agreement with the unions, and on 1 December, after 19 hours of negotiations, the government concluded a deal with them. The pension reform was separated from the 1995 budget; in return, the unions agreed to participate in the negotiation of a comprehensive pension reform in 1995. The projected savings from pension reform were therefore retained in the budget, on the understanding that the government would have the right to increase contribution rates by decree if no reform were adopted by end-June, and the right to retire on seniority pensions was suspended until a reform was adopted. The acceleration of the increase in the retirement age was also retained.

The negotiation of the promised reform, however, would not fall to the Berlusconi government. The conflict between FI and the LN continued to escalate, and in mid-December, three no-confidence motions were lodged in the Chamber of Deputies – one of them by the LN. Four of the government’s five LN ministers then resigned, prompting the prime minister to do likewise, thus pre-empting the confidence votes. In late January, Lamberto Dini, the Treasury Minister in the outgoing cabinet, took office at the head of a technocratic government with a mandate to pursue a limited number of specific objectives, including the pension reform included in the 1 December agreement.

The Dini government, supported mainly by the centre-left and the LN, soon reached agreement with the trade unions on a major overhaul of the pension system. Rather than pursuing an essentially parametric reform, the Dini government inaugurated a phased transition from the existing defined-benefit system to one based on notional defined contributions.

The government’s communications strategy

The government’s communications strategy, vis-à-vis both stakeholders and the public, was confused. The failure to publish the Castellino commission’s report deprived the government of one instrument that could have been used to shape public opinion. The views of a commission that included both independent experts and representatives of the social partners would probably have carried weight. Yet having opted to ignore the commission’s recommendations, the government could hardly rely on its analysis to make the case for reform. Moreover, despite an awareness of the likely trade-union reaction to the proposals, the government had given little thought to how the reform was to be presented to the public. Indeed, its one major media blitz backfired: commercials aired on state television declared that, without reform, “the state would no longer be in a position to pay out pensions.” These triggered a wave of criticism, because the advertisements, though publicly financed, were felt by many to cross the line between public information and propaganda. The opposition called on the regulator to take action, and the President scolded the government for its failure to ensure that state-controlled media presented a balanced picture of the issues. The coalition’s failure to maintain even a façade of unity as the conflict escalated confused matters further. Some ministers were clearly less willing to compromise than others, and the leadership of the LN seemed increasingly inclined to use the situation to engineer a politically advantageous exit from the government. Such public evidence of internal divisions undermined not only the government’s public image but also its ability to negotiate with the unions. Matters only got worse as a result of the LN’s increasingly public attacks on the prime minister and the latter’s mounting legal troubles.

One thing that was conspicuously missing from the government’s public relations effort was, oddly enough, European integration, which had been – and would continue to be – a powerful argument in support of reforms in Italy throughout the 1990s. Given the
importance of EMU as a source of pressure for fiscal consolidation, this is ironic, but it is not wholly surprising. Although the government remained formally committed to Maastricht, its line on EMU was hesitant and its overall outlook somewhat Eurosceptic. As a result, the government was unable or unwilling to make the Maastricht criteria an important element of its case for reform.

Implementation: the Dini reform

The reform proposals were withdrawn and the government fell just three weeks later. It may seem odd, therefore, to include a discussion of implementation in the present case study. However, the principal pension reform commitment contained in the 1 December 1994 settlement between the government and the unions was implemented. This was the agreement to adopt a pension reform in the course of 1995. The strengths and weaknesses of the pension reform adopted in mid-1995 reflect the lessons that key actors drew from the pension battles of 1994.

The critical element of the Dini reform was the gradual shift in the assessment of benefits from the old earnings-related system to a contributions-related formula for public and private employees and the self-employed. Benefits were also linked to such factors as retirement age, which was flexible (57-65 years), economic trends and demographic changes. The Dini package further introduced increases in contribution rates, new arrangements for workers with “atypical” contracts and for periods of child-rearing, and more generous tax treatment of supplementary private pensions. It was expected that private funded schemes would increasingly be developed to supplement contributions-based public pensions; these were in many cases expected to be based on the future termination indemnities (Trattamento di fine rapporto) accumulated by firms for their employees. In response to pressure to separate social insurance from social assistance, the pensione sociale, which was paid to people who had never made social security contributions and were without other income, and the low-pension supplement were replaced with a new, budget-financed, means-tested benefit. Finally, a gradual increase in the MCP for seniority pensions was initiated: this was to rise from 35 to 40 years over the period to 2008.

The big surprise of the reform was the shift to a defined-contribution system – something that experts had discussed for some time but that had not really been on the political agenda in 1994. There were many reasons why various experts were attracted to the idea of a defined-contribution system, but the key political advantage concerned its role in bringing about a harmonisation of Italy’s bewildering array of different schemes: harmonisation via parametric reform made the distributional consequences of change painfully obvious to all. The shift to an entirely new principle for calculating benefits – a principle which, moreover, was relatively easy to explain and to defend on equity grounds – seems to have presented itself as an attractive way to smooth the realignment of the various pension schemes.

The remarkably smooth adoption of the Dini reform reflected a number of factors. First, the unions were better disposed towards the Dini government, simply because it was supported by the left in parliament and not FI or the AN. Secondly, many union leaders had been shaken by the Berlusconi government’s attempt at pension reform and had become convinced of the need to work towards a reform their members could accept, rather than simply fighting defensive actions against further retrenchment. Thirdly, the policy process in 1995 was marked by a return to concertation: the unions were not merely consulted, they played a significant role in the design of the reform. They also
secured its “ratification” in a nationwide referendum of union members. Employers’ organisations, too, were involved in the negotiations, though Confindustria eventually refused to sign off on the resulting proposals, arguing that they did not go far enough or fast enough on account of the concessions made by the government to the trade unions. Fourthly, the acquired rights of older workers were protected: in return for trade-union support, the government conceded exceptionally long transition periods with respect to the reform’s most important elements. It is unlikely that the government could have secured adoption of the reform had it not accepted such very long lead times.

The Dini reform was in any case more generous than the Berlusconi government proposals had been: one contemporary analysis estimated that the Dini reform would reduce the replacement rate for a 35-year contributor to 65% at age 65 and 50% at 57; the corresponding figures for the Berlusconi package were 56 and 43%. Finally, financial pressures in early 1995 helped ensure that all the social partners were somewhat more cooperative. During the first quarter of the year, the spread between Italian and German government bonds more than doubled, the lira plummeted and the probability assigned by the markets to Italian participation in EMU fell sharply. The major players “learned that failure to reform would no longer mean maintaining the distributive status quo but would mean suffering hitherto unexpected and unavoidable new losses.”

The Dini reform went some way to reducing inter-sectoral inequities – when it is fully implemented, public- and private-sector workers and the self-employed will have essentially the same rules for benefit calculations, seniority pensions and retirement age. However, it substantially increased inter-generational inequities. While transition periods are often needed in connection with pension reform, the transition enshrined in the 1995 reform was exceptionally long and led to very large differences in the treatment of adjacent cohorts. In addition to the very gradual increase in the qualifying period for seniority pensions, the reform retained defined-benefit formulae for workers with 18 years of contributions in 1995 (i.e. cohorts that had accumulated 15 years of contributions in 1992). For workers with fewer than 18 years of contributions in 1995, pensions would be calculated on the old defined-benefit formulae for years worked up to that time, and on a defined-contribution basis for subsequent years. This implied a sharp break in the benefits available to two adjacent cohorts in the system: as a result of the Amato and Dini reforms, two otherwise identical individuals with a one-year difference in contribution history in 1992 could retire with a 15 percentage point difference in their respective pension replacement rates. Only workers entering the system from 1995 were entirely subject to the new contributions-based formulae. As noted above, the reform implied no loss of pension wealth for older employees, though it did affect older cohorts among professionals and the self-employed. By contrast, Beltrametti (1995) found the estimated losses inflicted on employees under 40 to be well in excess of their pre-reform pension wealth.

The long transition period built into the Dini reform has also raised questions about implementation. Some elements of the reform were not automatic but depended on subsequent political decisions. Thus, the reform provided for an adjustment of the conversion coefficients used to link pension entitlements to life expectancy every ten years. The first adjustment was to take place in 2005, but it was not automatic and in the end it was postponed. Similarly, the rise in the minimum age for seniority pensions was negotiated at length in the early 2000s (the so-called Maroni-Tremonti reform) only to be postponed. As Franco (2003) argues, implementation of the Dini reform has thus involved a process of continuous negotiation, which tends to generate uncertainty. Yet the microeconomic benefits of a defined-contribution system depend largely on the clarity
and stability of the link between contributions and eventual benefits. If the relationship is unclear, or if citizens anticipate further changes, workers may be more likely to opt for early retirement, seeking to exit the labour market before possible further retrenchment can occur.49

Conclusion

It has been argued that the 1994 pension reform defeat supports the view that, in Italy at least, “commitment to a concertation process leading to a substantial consensus among political and social actors – especially between the government and the unions – seems to be a fundamental prerequisite for successful pension reform.”50 This overstates the case and points to an approach to reform that risks making the trade unions and/or other lobby groups de facto veto-players. While it is certainly correct to say that unilateralism is not a promising strategy for pursuing pension reforms – unilateral pension reform attempts have been thwarted or reversed in a number of countries51 – it does not follow that concertation, in the sense of a real negotiation of the reform with social partners, is the only path to success. The experience of France, as well as other Italian reform episodes,52 would point to a strategy that combines consensus-seeking with a readiness for confrontation.

In the case of the 1994 pension reform attempt, the critical factor in the end seems to have been the fragility of the coalition – if the LN had not chosen to leave the coalition, the pension package might well have passed into law – and the radical nature of the proposals, driven by the need for near-term fiscal consolidation. The government was simply not prepared for the confrontation that its proposals triggered. A more popular and cohesive government might have held the line against union protests, something that trade-union leaders themselves seem to have recognised. Indeed, it was in part this awareness that prompted them to cooperate with Dini in 1995: many senior union officials concluded that if they did not take action on pensions themselves, a future government might try to address the issue without them again – and with more success.53

The September proposals also appear to have been put forward in response to immediate financial pressures and with too little thought given to either their distributional consequences or their political implications. They hit older cohorts much harder than the Amato reform had done, thereby antagonising both the unions and the LN. Somewhat tougher treatment of older workers may have been justifiable in terms of equity, given the extent to which younger cohorts had already borne the cost of the Amato reform, but the apparent protection of certain constituencies of the centre-right looked to many like favouritism, and the government offered no quid pro quo to the unions. There was little ex ante consultation to prepare elite and public opinion for the pension reform plan, nor could the government claim any mandate for it on the basis of the March elections. Its attempt to push through the reform as part of the budget process, with its strict deadlines, merely aggravated the opposition further. Moreover, having presented its proposals with little advance preparation, the government had no clearly defined strategy for selling them to the public. The near-total absence of Europe from the government’s pension reform discourse is particularly notable, given the significance of Europe in Italian political debates at the time and the importance of European integration as a source of pressure for fiscal consolidation.

As noted above, a number of factors facilitated the relative speedy adoption of the Dini reform in 1995, but the key point seems to have been the government’s acceptance of a very long transition period before the reform was fully implemented. Without that
concession to the unions, the reform would probably have been blocked: the speed with which the 1994 proposals would have taken effect was one reason for the fierceness of the unions’ opposition. Yet the very long transition enshrined in the 1995 reform has been criticised by some for going too far in the other direction. To be sure, a certain degree of gradualism is unavoidable in matters of pension reform – sharp swings in pension policy can be very costly to contributors making career/savings choices that will pay off only in the long term. However, very long lead times not only imply slower progress towards financial health, they can also raise political economy risks of their own, particularly if important but contentious decisions are postponed to the implementation phase, as has been the case in respect of the Dini reform. The reform has not been reversed, but continuing uncertainty about how and at what pace it will be fully implemented has imposed costs of its own.

Notes

1. So poor, indeed, that the FI had had to conclude two separate electoral pacts, the “Pole of Good Government” (FI, AN) in the south and the “Pole of Freedoms” (FI, CCD, LN) in the north.

2. See OECD (1995), especially Table 16. In the end, the Dini government achieved a far better outcome for 1995 – net borrowing equivalent to 6.9% of GDP and a 2.9% primary surplus, with the discrepancy reflecting the underestimation of debt-service costs in the original budget.


4. While nothing like as severe as the crisis of 1992, the lira’s slide took it near a record low against the DM and was serious enough by mid-August to prompt the prime minister to cut short his summer holiday.

5. See, for example, the July 1992 Pact on the Termination of the Cost of Living Index Mechanism (the so-called scala mobile) and the tripartite agreement of 22 July 1993 on incomes policy, collective bargaining and labour-market reform.

6. For details, see Ferrera and Gualmini (2000) and Pérez (2000).

7. Smaller independent confederations and unions are estimated to account for a further 10-20% of union membership.


9. A 1995 referendum abolished Article 19 of Law 300/1970, which regulated the selection of representative bodies on the basis of the most representative trade unions.

10. These include, in addition to the main business lobby, Confindustria; Confcommercio (merchants), Confartigianto (artisans) and Confagricoltura, and many minor associations.


12. In 1990, the defined benefit method was extended to the self-employed, one of the most expensive changes made to the system. See Baccaro (2002:416).


15. So radical were the implications of the Amato package, that many of the cuts were offset when the implementation decrees were issued by the Ciampi government: individuals were able to exclude their worst earnings years, up to 25% of the total, from the calculation of benefits, and the rate of revaluation of earnings, previously set at 1% per year, was raised to about 1.5%. Pizzuti (1998:47) estimates that this made for about 20% recovery of pension values for the final system. See Decreto Legislativo 11 agosto 1993 N 373.


17. That is, the present value of the promises made under legislation then in force to retirees and workers, net of the contributions that the latter would have to pay later on.

18. For detailed estimates of the cost of reform by age group, see D’Amato and Galasso (2002:197).

19. When the proposals were first unveiled in July, the unions organised protests against the measures with immediate impact, while signalling their openness to negotiating longer-term measures; after the lira’s ejection from the ERM, they accepted the need for drastic action in the near term.

20. Pizzuti (1998:47); for a good overview of the politics of the Amato reform, see Ferrera and Jessoula (2007:431-34). On the “seniority bias” in union representatives’ behaviour (which is not limited to pension issues), see Brugiavini et al. (2001).

21. See the poll results presented in Natali (2004:1086); and Ferrera (1997).


23. That is, an individual retiring at 58 with 40 years of contributions would be roughly equal to one retiring at 65 with 33 years.


25. This was not without recent precedent. Act 421 of 23 October 1992, adopted in the wake of the floating of the lira, gave the Amato government the “delegated authority” to change the pension system within general guidelines set down by parliament and with no requirement to negotiate with social partners.


27. In addition, the power to adjust this coefficient in future was to be delegated to the Ministry of Labour and Social Protection, in consultation with the Treasury.

28. Opponents were quick to point out that this would simply have transferred the power over cost-of-living adjustments to the Treasury.

29. The retirement age was to be raised to 62 for men and 57 for women as of July 1995 and to rise by one year every 18 months thereafter, rather than by one year every two years. This would have reached the level envisaged by Amato (65 and 60, respectively) two years earlier than planned.

30. The right to retire on seniority pensions has often been suspended during pension reforms, as the prospect of reform constitutes an incentive for those in a position to take seniority pensions to exit the labour market quickly, before any reform can impact on their pension entitlements.
31. See Artoni and Zanardi (1997); and Natali and Rhodes (2004).

32. See also Beltrametti (1995).


34. Cf. *La Repubblica*, 20 December 1994; signatories included, in addition to Nobel laureate Franco Modigliani, both Romano Prodi, a future centre-left prime minister, and Mario Baldassarri, who would later enter government as a deputy minister from the right-wing AN.

35. They estimate that the Dini reform of 1995 cost these cohorts nothing: it had no significant impact on the pension wealth of employees 40 or older.


38. Indeed, Regini and Regalia (1997) go so far as to argue that the unions and the government could probably have reached agreement on substance had it not been for the latter’s decision to abandon concertation.


42. Under Dini, in 2008, workers could retire at any age with 40 years of contributions or at 57 with 35 years (the latter has since been raised to 40). Public sector retirees retained, for the time being, the right to retire after 20 years, but with pensions reduced by 1-35% according to the shortfall of contribution years *vis-à-vis* the 35-year norm.

43. On this point, see Ferrera and Gualmini (2004).

44. The Dini reform was overwhelmingly approved by pensioners, whom it barely affected, and was also approved by a majority of active workers. Strikingly, it was rejected only in a few sectors and regions where the issue of seniority pensions loomed particularly large. Although Dini’s revision of the rules for seniority pensions was extremely gradual, and regarded by many observers as one of the weakest elements of the reform, it dominated debate within the unions and generated more opposition than any other measure. See Baccaro (2002:420-23).


46. See Ferrera and Jessoula (2007:436-7) for details.


48. Billari and Galasso (2008). This may not be the only such breakpoint: Franco (2002) points out that the revision of the conversion coefficients at ten-year intervals may also generate large differences in the treatment of adjacent generations of pensioners.

49. Franco (2003). This is one reason why suspension of the right to retire on seniority pensions has often accompanied reform initiatives: when pension reform is on the agenda, those eligible to retire on such pensions tend to do so in larger numbers than otherwise.

50. Ferrera and Jessoula (2007:435); see also the conclusions of Natali and Rhodes (2004); and D’Amato and Galasso (2002).

51. Other examples include the French government’s attempt at pension reform in 1995 and the changes adopted in Germany by the last Kohl government, which were reversed after the centre-left came to power following the 1998 general election.
52. Such as the Amato reform discussed above or the Maroni reform of 2004.

53. Cf. Baccaro (2002:419): “...given the government’s attempt to impose pension reform unilaterally in 1994, unions could easily have supposed that if they failed to reform the system on terms they were comfortable with, a future government might try again – with perhaps more success – to marginalise them.”

54. See Boeri et al. (2007:203-04).

55. See especially Franco (2002) and the references therein.
CHAPTER 4. ITALY: THE 1994 PENSION REFORM PROPOSALS – 121

Bibliography


Chapter 5

Mexico: Reform of the federal employees’ pension system, 2001-07

Summary and main conclusions

A reform of the pension system for federal employees administered by the federal Institute for Social Security and Social Services for State Workers (ISSSTE) was adopted in March 2007. Like the reformed general pension scheme for private-sector employees, the new federal employees’ pension system is based on defined contributions. Under the reformed system, pension contributions are to be channelled to individual accounts, allowing portability of pensions between the ISSSTE scheme and the general scheme. The transition generation has the option of remaining with the old, defined-benefit system, subject to some modifications, or receiving a bond recognising accrued pension rights as debt that will mature when the worker reaches retirement age. This will allow inter-temporal distribution of the impact of the transition on public finances. The monetisation of acquired rights in this manner is expected to cost the taxpayer around 8% of GDP in net present value (NPV) terms. The reform reduced the NPV of contingent liabilities arising from the ISSSTE pension system from an estimated 57% of GDP to around 35%.1

The reform was finally adopted with the backing of Mexico’s main public-sector unions, who had initially been strongly opposed to reform. Among the factors contributing to this outcome were the following:

- The formation of a consensus on the need for reform. It was clear by the early 2000s that the ballooning deficits in the ISSSTE system could not be sustained over the long term. While the terms and timing of the reform were hotly debated, the government had little difficulty persuading all the major players of the need for change. Moreover, the fact that the pension scheme for private-sector workers had been reformed a decade earlier made it still more difficult for federal employees to defend their very privileged pension regime.

- Acknowledgement of acquired rights. Those who entered federal employment prior to the reform can remain with the old system, subject to some modifications, or switch to the new. Only workers entering federal employment from 1 April 2007 must automatically join the new regime. Members of the transition generation who choose to join the new system receive a bond recognising the net present value of their accrued pension rights as debt that will mature when they reach retirement age. The government’s low overall debt position meant that it was in a good position to issue the bonds needed to finance this monetisation of acquired rights.

- Extensive consultation, explanation and promotion. The reform plan aroused considerable concern among public sector unions. Finance ministry officials
therefore consulted trade union leaders when crafting their initial proposals, in some cases amending them to address union concerns. Then, during 2003-05, the administration devoted considerable energy to selling the reform to both union leaders and the rank-and-file. In the end, almost all the public-sector unions accepted the reform.

- **Presidential leadership.** The first attempt to pass the reform legislation through Congress failed, largely because it was submitted very close to the 2006 presidential elections and because key actors involved doubted the strength of the outgoing administration’s commitment to the reform. After the elections, however, the incoming president quickly signalled his personal commitment to the reform and, by acting at the start of his term, raised the potential cost of opposing him.

- **A cash injection into the ISSSTE system.** Pension reform typically involves significant up-front costs, while the benefits of reform are not realised for decades. In order to ensure that reform would bring some immediate and visible benefits, the pensions overhaul was accompanied by a substantial injection of new resources into ISSSTE, a move made possible by Mexico’s relatively healthy fiscal position. ISSSTE, which had been chronically under-funded for years, is responsible not only for pensions but also for healthcare and some other social services and benefits for federal employees. The additional funds thus helped not only to modernise its pensions administration but also to improve healthcare and other facilities in the system.

The context for reform

**Framework conditions for reform**

The reform of the ISSSTE pension system was undertaken during the first years following Mexico’s transition to real multiparty politics following 70 years of rule by the Institutional Revolutionary Party (PRI). This created a challenging environment for reform. Following the 2000 elections, the newly elected Fox administration faced a Congress which was already well on its way to becoming an independent and active legislature and in which the president’s National Action Party (PAN) did not command a majority in either chamber. Moreover, the sweeping powers of the presidency during the decades of PRI rule had reflected not only the PRI’s permanent congressional majority but also its penetration of virtually every important state or social institution. The PAN presidents elected in 2000 and 2006 did not inherit this “metaconstitutional” authority. Indeed, the position of Mr. Fox’s successor, Felipe Calderón, initially looked even weaker, owing to the very close election result in 2006 and the refusal of the main opposition candidate to acknowledge the legitimacy of the results.

Economic circumstances also created challenges for the Fox administration, which took office just as Mexico’s recovery from the “Tequila crisis” was losing momentum and the world economy was slowing. Real GDP growth during its first three years (2001-03) averaged just over 0.6% a year, as compared with 4.5% during 1996-2000. While the slowdown did not have a direct impact on the politics of pension reform, sluggish growth did rapidly undermine public support for the administration: the PAN lost around one-quarter of its congressional seats at the 2003 mid-term elections, leaving the president weakened for the remainder of his six-year term.
Nevertheless, electoral defeat undermined the internal discipline of the PRI itself, as well as its hold over the major trade unions and other institutions. This created openings for coalition-building involving elements of the PRI and some traditionally PRI-affiliated trade unions. Moreover, while damaged by the slowdown in 2001-03, the PAN administration did reap some of the benefits of the recovery of the late 1990s: it inherited a relatively healthy federal budget and, in particular, a low and declining ratio of public debt to GDP. The public debt/GDP ratio when the Fox administration was elected stood at about 26% on the traditional definition and was falling fast. By the time of the reform, it stood at about 20%, leaving plenty of room for the government to issue the debt needed to finance the “recognition bonds” awarded to federal employees entering the reformed pension system.

**Actors and interests**

Within the Fox administration, the driving force for ISSSTE reform was the Ministry of Finance and Public Credit (SHCP), which saw federal employees’ pensions as the greatest public expenditure challenge then facing Mexico and which also believed that reform could contribute to greater labour-market flexibility. The leadership of ISSSTE itself was also involved, both in convincing the public sector unions to accept the reform and, later on, in implementing it. Initially, it approached pension reform rather cautiously, fearful of a backlash from the trade unions, which played an important role in ISSSTE’s governance. However, ISSSTE’s managers could hardly oppose reform, given the catastrophic state of its finances, and it stood to gain as an institution from the injection of budgetary funds that eventually accompanied the reform plan. The installation in late 2006 of a new director-general, with a clear mandate to pursue reform, also helped galvanise ISSSTE’s backing for pension reform. This change in its leadership, in turn, was the direct result of the mobilisation of the presidency itself in support of reform: Mr. Calderón took a much more active direct role in pushing for ISSSTE reform than did Mr. Fox.

The legislative phase of ISSSTE reform was complicated by the fact that no one party commanded a majority in either chamber of Congress after 2000. The president’s PAN could be counted on to support reform, and the left-wing Revolutionary Democratic Party (PRD) could be relied on to oppose it. While there were some smaller parties in Congress whose support the administration could court, the crucial position throughout was held by the PRI, which during the latter years of the Fox administration was by far the largest party in both chambers. The PRI was by no means wholly opposed to reform – it had, after all, reformed the much larger pension system for private-sector employees in 1997 – but it was split on ideological lines into more and less reform-oriented factions, and it was also divided over the political consequences of cooperation with the PAN. Even some relatively reform-oriented PRI leaders were reluctant to assist the administration and thus hand a victory to the president, while others argued that unremitting obstruction would ultimately hurt the PRI. They favoured allowing the PAN to take primary responsibility for a difficult but necessary reform, thereby sparing the PRI the need to deal with the problem once it returned to power.

Mexico’s public sector is heavily unionised, and consultations with organised labour were a major feature of the entire reform effort. The unions’ main concern throughout the process was to protect the acquired rights of older cohorts. While the reform team within the SHCP had over 100 unions with which to deal, two stood out above all: the teacher’s union SNTE – which, with 1.3-1.4 million members, was the largest union in Latin America – and the umbrella organisation for all public sector unions, the FSTSE. Both
were by tradition closely aligned to the PRI, and their leaders were in Congress as PRI representatives. However, there were growing tensions between the SNTE and the leadership of the PRI: SNTE leader Elba Esther Gordillo, one of the PRI’s congressional floor leaders in the early 2000s, broke with the PRI in 2005 and formed her own political party.

One set of actors might have been expected to influence the reform process but did not: private pension managers never emerged as a significant force in support of reform, although they represented a potentially significant pro-reform lobby. However, Mexico’s private pension fund managers, or AFOREs, were relatively new, having emerged only in the wake of the 1997 reform of the general pension system, and they had not yet developed any real corporate sense of themselves as a sector. The bigger AFOREs largely belonged to global financial groups, whose principal interests in Mexico lay elsewhere and for whom the ISSSTE reform was not a major concern, while the smaller, mainly local AFOREs had little influence. Moreover, the AFOREs’ performance was already subject to growing criticism in Mexico by the early 2000s, so it would have been risky to court further controversy by publicly supporting such a sensitive reform. In any case, the AFOREs stood to profit from the reform whether they backed it or not, and the state of ISSSTE’s finances made it unlikely that reform could be postponed indefinitely. The AFOREs thus has little incentive to mobilise for reform.

Issue history

Though established in 1959, the pension system for federal employees that existed in the early 2000s (Box 5.1) was largely a product of the 1970s oil boom. In an effort to maintain support among the growing army of federal workers, the ruling PRI began granting them very generous benefits, underpinned by a government commitment to cover any deficits in the system that might arise; contributions remained very modest indeed. The system, which operated on a pay-as-you-go basis, did not initially appear to be unsustainably generous: as late as 1980, ISSSTE supported one retiree for every 20 public sector workers. However, that ratio had fallen to 5:1 by 2003 and it was expected to shrink to 2:1 by 2020, so the federal commitment to cover the system’s deficits was rapidly becoming very costly. The deterioration of the system’s dependency ratio reflected not only the ageing of the federal workforce but also the increasing use of early retirement to downsize the public administration after 2000.

A small cash-flow deficit first appeared in the system in 1993, although its existence was not publicly acknowledged for some time. By 2002, the deficit had reached 15 billion pesos (0.24% of GDP), rising to over 37 billion (0.4% of GDP) in 2006. By then, the government was putting roughly six pesos into ISSSTE for every peso in contributions from federal workers. Almost 80% of the state’s contribution consisted of budgetary subsidies to cover the ISSSTE pension deficit; the remainder represented the state’s share of contributions on behalf of federal employees. ISSSTE constituted by far the single biggest source of future unfunded liabilities for the Mexican state: the actuarial debt of the ISSSTE pension system at end-2002 was estimated at around 45% of GDP and was reckoned to be rising by about 2 percentage points per annum. Altogether, the actuarial deficits of the other unreformed public sector pension systems were estimated at a further 45% of GDP.
Box 5.1. The Institute for Social Security and Social Services for State Employees (ISSSTE)

With 2.5 million enrolled employees and pensioners, as well as 10 million dependents, the federal ISSSTE is the second-largest social security institution in Mexico after the IMSS, which covers 14 million workers and around 45 million dependents. Smaller systems exist for Pemex and other parastatals (247 000 employees), the governments of Mexico’s federal states (900 000 employees covered by state-level ISSSTEs), the universities, the municipalities, the Development Bank and, finally, the 375 000 employees of the IMSS itself, who are covered by a system of their own. In addition to pensions, ISSSTE provides medical care, insurance, personal loans and other cultural and social benefits to federal employees and pensioners. It is overseen jointly by the federal government and the public sector unions, with both the unions and key federal ministries represented on ISSSTE’s board.

Pre-reform ISSSTE pensions were exceptionally generous. A worker with under 15 years’ service had no pension rights, but a federal employee with 15 years of ISSSTE contributions was entitled to 50% of salary from the age of 55. With 30 years of service, he could retire with a pension equal to final salary, regardless of age. Since promotions sometimes carried additional terms of pensionable service, which were credited but not worked, most federal employees retired after just 28 years on the job. Two other factors made this system steadily more expensive. First, life expectancy rose, exceeding 77 years by 2006, while the average retirement age in the ISSSTE system fell from roughly 62 in the mid-1970s to just about 56 by the early 2000s.* Secondly, a 2001 law stipulated that ISSSTE pensions should rise annually by an amount equal to consumer price inflation or the increase granted to the relevant public sector union, whichever was higher. ISSSTE pensions were not portable: an individual who left federal service before securing the right to a pension had no pension rights.

Despite the generosity of ISSSTE pensions, pension contributions were extremely low: the state and the employee each paid 3.5% of base salary; in addition, from 1992, the state paid a further 2% into an individual account under the system of retirement savings, making for a total of 9% of base salary saved, of which the state contributed just over 60%. In 2003, the SHCP estimated that the capitalised contributions for a federal employee with 30 years of service, including both employee and employer contributions, would amount on average to no more than 22% of the eventual cost of his pension. Closing this gap via contribution increases alone would have required a rise in contribution rates to 44.5% of base salary; holding contribution rates constant, the deficit could be eliminated by raising the average effective age of retirement by 20 years, to 75. Raising the retirement age to 65 would render the system sustainable at a contribution rate of 26.8%.

Prior to the reform, there was no separation of ISSSTE pension accounts from healthcare and other activities: workers and employers’ contributions, along with budgetary subsidies, went together to form a single budget, without any strict allocation of financing to particular functions. For many years, this was an attractive feature for those involved in the governance of ISSSTE, who had far greater discretion in managing the very large sums passing through the system than they would have enjoyed if the system had been based on individual accounts or clearer earmarking of funds. As the pension deficit mounted, however, the lack of separate financing for separate activities simply meant that pension payments dominated everything else. ISSSTE’s healthcare system suffered from chronic under-investment, even though it was better financed than most other non-pension activities, many of which were simply starved of cash to cover pensions and healthcare.

* According to the SHCP, the average public sector worker retiring in the early 2000s could expect to live a further 22 years on his or her pension.
The PRI administration of Ernesto Zedillo successfully reformed the Mexican Social Security Institute (IMSS) pension scheme for private-sector workers in 1997, transforming it from a defined-benefit plan to a defined-contribution system with individual accounts, private management and greatly reduced actuarial deficits. Under the 1997 reform, new workers were enrolled in the reformed system automatically but those already in the system before the reform were free to choose between the new and old regimes; moreover, they were free to make the choice only at the end of their careers. This implied a roughly 30-year transition, something Mexico could afford chiefly because it was still in the very early stages of demographic transition. The Zedillo administration initially hoped to reform the ISSSTE system as well and took some tentative steps in exploring that possibility, but it soon became apparent that the political price would be too high. The public-sector unions were strongly opposed, and the administration was reluctant to antagonise them ahead of what was expected to be a very competitive presidential election in 2000. In any case, ISSSTE’s financial problems, though growing rapidly, did not yet look so acute as to require immediate action. However, the generosity of the ISSSTE system looked even more anomalous in the wake of the IMSS reform, the implementation of which provided a template for subsequent reforms and led to the creation of a range of new institutions, including private pension funds and a new pension regulator, which could play a role in any public-sector pension reform.

Policy process: formulation

The search for solutions

In October 2001, the SHCP presented an actuarial study of ISSSTE to the Chamber of Deputies. This led to an informal debate and began the process of educating Congress and the public. Work on a reform bill began in earnest that same year, and the first rough draft of the reform was prepared by the end of the following year. Throughout 2003-05, these proposals were discussed with union leaders, key congressional figures and other stakeholders. These negotiations were facilitated somewhat by the unions’ direct involvement in the running of ISSSTE: while the trade unions remained firm in their defence of acquired rights, their leaders could see first-hand the disastrous state of ISSSTE finances and were therefore open to reform. Moreover, the incumbent head of ISSSTE was a former PRI politician with close ties to the SNTE. Whatever the views of the rank-and-file, union leaders themselves were increasingly aware that the status quo was unsustainable.

From the beginning, the SHCP officials preparing the reform aimed to model it on the IMSS reform. The broad intention was to migrate federal employees to essentially the same system as private-sector workers, even if certain institutional features of the public sector system remained distinctive. In part, this approach simply reflected the conviction underlying the 1997 reform – viz. that a defined-contribution system based on personal accounts offered the surest way to ensure both the adequacy of future pensions and the financial sustainability of the system overall. It would also introduce much-needed transparency and predictability into the system, since pension rights assigned to individuals as a matter of right could not be easily subjected to ex post political manipulation or adjustment. As generous as the old ISSSTE scheme was, it was based entirely on political discretion: individuals had no legally entrenched pension rights. There was nothing to prevent the government from cutting back on federal employees’ pensions in order to hold down budgetary subsidies to ISSSTE in the event of a future financial squeeze. The reformed system was also intended to reward longer careers in a
sector characterised by the prevalence of early retirement, while the harmonisation of the
ISSSTE and IMSS schemes was meant to allow portability of pensions between public
and private sectors, and thus to increase labour-market flexibility. Finally, the SHCP
sought to eliminate, or at least reduce, socially regressive cross-subsidies. The fact that
ordinary taxpayers on relatively modest incomes were financing the pensions of Mexico’s
comparatively well paid public sector workers was regarded as obviously inequitable.

In pursuing these goals, the SHCP officials designing the reform also determined to
adhere to two other basic principles: leaving current pensions completely outside the
reform and respecting the (implicit and explicit) acquired rights of workers already in
federal employment. However, the SHCP wished to limit the potentially open-ended
character of the latter commitment; whereas the IMSS reform allowed individuals to opt
for one scheme or the other at retirement, thereby inaugurating a 30-year transition
period, the SHCP wished to move federal workers onto the new system at once. With this
in mind, the reformers borrowed an element of the 1981 Chilean pension reform that had
not been used in 1997: the issuing of “recognition bonds” to workers entering the new
system. In essence, each federal employee was to be given a bond, equal in value to the
net present value of pension rights accrued under the old system, which would mature
when the worker retired. Indeed, the bonds would honour even those acquired rights
that were not yet secured in law: under the old system, a federal employee had no pension
rights for less than 15 years of service. If he left federal employment before that time, he
was not entitled to a pension. However, the SHCP proposed to offer recognition bonds for
all federal employees – otherwise, around 60% of incumbent federal workers would have
received no initial capital under the new scheme and would also have been deprived of
the benefits of pension portability, at least until they had done 15 years of service. Since
an individual became immediately eligible for a pension equal to 50% of final salary at
the 15-year mark, rising in linear fashion to 100% after 30 years, the calculations for the
recognition bond simply assumed a linear accumulation of rights from the first day of
federal employment.

While these basic principles remained at the heart of the SHCP’s work on the design
of the ISSSTE reform, it rapidly became clear in discussions with congressional and trade
union leaders that further concessions would be required. First, even with the recognition
bonds, there would be strong resistance to any move to compel all incumbent federal
officials to migrate to the new system. The reformers therefore opted to give workers a
choice between the old and new schemes. However, this choice would at least be subject
to a deadline – officials would not able to exercise their option on the point of retiring, as
the IMSS reform had allowed private-sector workers to do. Moreover, those officials
opting to remain with the old defined-benefit system would still face a gradual but
significant rise in contributions and a gradual increase in retirement ages (Box 5.2).

The question of who would manage ISSSTE pension accounts was a matter of great
concern to the unions, who opposed private management of federal employees’ pension
contributions. The unions had long participated in the management of ISSSTE and they
were clearly reluctant to sacrifice this role in the course of a reform. In 2003, the FSTSE
and its allies began pressing for the creation of a public AFORE to act as the sole
manager of federal employees’ contributions. The pension regulator CONSAR publicly
questioned the feasibility of such a solution, and the SHCP was decidedly unenthusiastic.
However, the unions continued to press for a designated state-owned AFORE to run the
system, and a compromise was reached on the creation of “PensionISSSTE”, a publicly
owned AFORE with a board made up of equal numbers of state and union
representatives. If the board was evenly divided, the casting vote would rest with the
state. However, PensionISSSTE would not be allowed to remain a monopoly fund manager for federal workers in perpetuity. Instead, federal workers opting to accept the recognition bond would have to remain clients of PensionISSSTE for three years; thereafter, they would have complete freedom of choice. Those entering federal employment after the reform took effect were to be allowed to opt for any AFORE they liked.

**Box 5.2. The ISSSTE reform plan**

Under the reform plan, individuals taking up federal employment after the reform took effect would contribute 6.125% of their basic salaries to an individual account managed by the AFORE of their choice; the state would pay an additional 6.625%, making for a total of 12.75%. The individual would then convert the accumulated account balance into a price-indexed annuity upon retirement. In addition, the plan offered federal employees a very generous voluntary savings pillar, in an effort to increase support for the reform: individuals would be permitted to set aside up to 2% of salary under this scheme, to which the state would make a contribution at a matching rate of 3.25 to 1. In other words, if an individual channelled the full 2% of salary into the voluntary pillar, the state contribution would amount to a further 6.5%.

Those already in federal employment at the time of the reform would be given the choice to opt into the new scheme, accepting a recognition bond reflecting the value of their acquired rights under the old system, or to remain with the existing defined-benefit arrangement. If they chose the latter, however, they would still face a rise in their 3.5% contribution rate of 0.5 percentage points per year until it reached 6.125%. From 2010, moreover, a minimum retirement age would enter into force even for those with 30 years of contributions; the old system had never had a minimum retirement age for those with full contribution histories. Beginning at 51 for men and 49 for women, the minimum retirement age would rise by one year every 24 months until it reached 60 and 58, respectively, in 2028. The reformers in the SHCP had initially aimed for a standard retirement age of 65, but the unions resisted this. The age at which individuals with fewer than 30 years of contributions could retire, which was 55 prior to the reform, would also rise by one year every two until it reached 60.

**Policy process: adoption**

*Frustration during the Fox years*

Securing passage of the reform through an opposition-dominated Congress was recognised from the beginning as a major challenge, particularly against a backdrop of economic weakness and steady advances by the PRI in state elections and the 2003 congressional mid-terms. Without the support of at least some elements of the PRI, reform would be impossible. By autumn 2003, however, a coalition in support of reform seemed to be coming together in the Congress, involving the PAN and some elements of the PRI, including the leaders of the SNTE and the FSTSE. Before the legislation could come to the chamber, however, an upheaval in the PRI led to the replacement of the SNTE’s Mrs. Gordillo as PRI floor leader and the reform package had to be shelved. In part, this was simply the outcome of a struggle for leadership of the PRI, but it also reflected differences within the PRI over strategy. PRI leaders were increasingly confident of recapturing the presidency in 2006 and thus had little incentive to cooperate with the administration, particularly if doing so would risk alienating public sector workers, a crucial constituency for the party.
The administration reacted to this setback by intensifying its efforts to win union acceptance of the reform plan, without which the congressional hurdle might never be overcome. However, the SHCP also continued to lobby Congress directly, in an effort to keep its political options open. This led to friction with some trade union leaders. Moreover, the administration pressed ahead with a much more modest reform of the pension system for IMSS employees, which was adopted in August 2004, over the objections of the union concerned.16 Taken together, these actions sent an important signal to the other public sector unions: while the SHCP badly wanted to find common ground with the unions concerning ISSSTE, its lobbying of Congress and its reform of the system for IMSS workers underscored its determination not to concede a de facto veto to organised labour. With the agreement of the FSTSE leadership, senior SHCP and ISSSTE officials began a series of consultations with over 20 major public sector unions, representing more than 85% of the affected workers. The consultations proceeded in stages throughout 2004-05. The unions’ reaction was heterogeneous but broadly positive – around 94% of union members participating in these consultations accepted the reform. Eventually, even the FSTSE leadership came out publicly in support of reform – a critical boost, given the FSTSE’s influence in the PRI.

The consultation process set the stage for a return to Congress in more promising circumstances. In December 2005, FSTSE and SNTE officials presented the reform in Congress. However, the SNTE pulled back from supporting the bill at the last minute.17 Accounts differ as to why the agreement fell apart at this late stage, but the main concern seems to have been a fear on the part of SNTE leaders that it was already too close to the 2006 elections to risk passing such a complex reform. Adopting it would require calling an extraordinary session of Congress in January, and it was not clear that the president would be prepared to risk convening a special session on such a sensitive subject just months before the poll. Given uncertainty about the president’s commitment to the reform and the knowledge that he could not stand for re-election, SNTE leaders and their allies had little incentive to take risks in order to help pass the reform; they could afford to wait.18

Adoption under Calderón

At the end of 2006, the incoming Calderón administration signalled its determination to proceed with ISSSTE reform, setting the goal of securing its adoption by the end of the first quarter of 2007. To the surprise of many observers, it met this deadline. Three factors seem to have made such a rapid outcome possible. First, Mr. Calderón made a priority of ISSSTE reform from the moment of his election. He appointed another former PRI official with close ties to the SNTE as Director General of ISSSTE, with an explicit mandate to pursue structural reform along the lines laid out in the SHCP plan. At the same time, Mr. Calderón set to work bringing the main public sector unions back on board. Negotiations with the unions at this point were centralised in the hands of the president and Secretary of Finance and Public Credit Agustín Carstens. Both men were much more directly engaged in the reform process than their predecessors had been. Mr. Calderón’s leverage vis-à-vis the unions and other actors was greatly enhanced by the fact that he was at the start of his six-year mandate. Any opportunistic behaviour or defection from prior undertakings at his expense would carry a potentially high political cost over the years to come. Only when agreement with the FSTSTE and SNTE leaders had been reached did the administration present the agreed compromise to the congressional parties.
Secondly, in gearing up to push the reform through Congress, the new administration adopted one significant change to the overall package: at the behest of the unions, substantial sums of money were made available to improve ISSSTE’s other services. These included some 8 billion pesos to improve healthcare services, 2 billion for personal loans and mortgages and smaller sums for various other activities. The unions’ view, which the government accepted, was that such an injection of funds would ensure that reform brought some immediate, palpable benefits. After the reform, moreover, the accounts for different ISSSTE activities were to be separated so that contributions were used to finance the activities they were meant to finance and the full range of ISSSTE services, including personal loans and other benefits, might once again be made available. If pensions alone were addressed, then the costs of reform would begin to be felt relatively quickly, whereas the benefits would be realised only over the very long term. Provided the additional funds were reasonably well spent, this proposal posed no problem for the SHCP, since ISSSTE was universally acknowledged to be hugely under-financed across the board. Thus, while the trade unions could take credit for this “concession”, the additional expenditure was not seen by the SHCP as much of a concession at all – it addressed real problems and in no way compromised the essential aims of the reform.

Thirdly, the revised bill went to Congress as a union initiative, rather than a government bill, with a request from the FSTSE and the SNTE to adopt it quickly. This made it far more difficult for PRI deputies, in particular, to oppose the bill. The PAN, the PRI, the Greens and the New Alliance all backed it, leaving the PRD as the only major source of opposition. Thus, despite the organisation of some anti-reform demonstrations in Mexico City, the bill made smooth passage through the chamber. In the Senate, it was held up by PRI demands for reforms to improve the competitive environment for the AFOREs, which had been the subject of considerable criticism for several years. While the regulation of the AFOREs was a distinct issue from the ISSSTE reform per se, concern about the performance of the existing AFOREs had been one of the arguments used by opponents of ISSSTE reform. On this point, the SHCP found it relatively easy to give way to the PRI’s demands, as it had been preparing such a reform anyway. The most important change was to make the structure of the AFOREs’ commissions more transparent. Here, too, then, the “concession” made to secure adoption of the reform proved relatively easy, precisely because it was coherent with the overall reform rather than cutting against it.

The government’s communications strategy

In presenting the reform to the unions, ISSSTE and SHCP representatives emphasised five things. First, they stressed that the status quo was simply untenable. Second, they emphasised that ISSSTE members’ benefits in the existing defined-benefit system were not guaranteed by law and could be changed by political decisions at any time. Given rapidly rising ISSSTE deficits, they stressed that benefit cuts would at some point be inevitable in the absence of reform. Individual accounts, by contrast, would create property rights, protecting individuals against ex post manipulation of their benefits. This was particularly important for younger workers, since it was unlikely that the existing arrangements would prove financially or politically sustainable to the end of their careers. It was in this context that officials emphasised the recognition bond, which offered a clear statement of the value of acquired rights already accrued under the old system. Third, great stress was also placed on the fact that no one would be compelled to take the recognition bond; employees would retain a right of choice under the reform and thus need not be threatened by it in any way. Fourth, the potential benefits of portability were
highlighted, particularly for younger workers. Finally, SHCP analyses consistently stressed that longer delay would mean deeper cuts when reform finally came, as each new cohort of workers joining the unreformed system was driving up its future liabilities.

While official advocacy efforts focused on the public sector unions, the SHCP also made use of the mass media to explain the ISSSTE problem to the general public. Estimates of actuarial net debt in excess of 100% of GDP for the unreformed public sector pension schemes were the focus of much attention. The additional funding for ISSSTE provided a strong further argument during the latter stages of the reform debate: while it was hard for the public to digest talk of astronomical “actuarial deficits” extending into the distant future, everyone who dealt with ISSSTE knew that its services were in terrible shape, particularly in the field of healthcare: the hospital infrastructure was decaying; treatment delays were normal; and medicines were in short supply. ISSSTE officials therefore laid great stress on the immediate benefits of reform in terms of healthcare and other non-pension services.

Policy process: implementation

Since federal officials entering service after the reform took effect on 1 April 2007 simply chose AFOREs for themselves and entered much the same system as private-sector workers, the key implementation issue concerned the disposition of incumbent officials, who had the option of remaining with the status quo or accepting recognition bonds which would form the basis of their AFORE savings. To some extent, these individuals confronted a choice about the risks to which they wished to be exposed. Those who remained with the old system would face greater political risk, since they did not have legally defensible acquired rights; their pension entitlements could be adjusted at any time prior to retirement. Individuals opting for the recognition bonds, by contrast, would have legally clear property rights but would be more exposed to financial risk, since their benefits would depend in no small measure on the performance of their AFOREs. This made it likely that younger workers would be more inclined to take the bond than their older colleagues: their investments would grow over a longer period of time and they were more likely to value the new portability of their pensions.

Federal workers were given until mid-2008 to choose whether or not to accept the recognition bonds. However, this deadline proved unrealistic, because ISSSTE’s records were not in good enough order to allow an accurate assessment of acquired rights. Enormous effort therefore went into reconstructing ISSSTE’s often erroneous or incomplete data, and the deadline for choosing whether or not to accept the recognition bonds thus had to be postponed until 14 November. In this context, the infusion of additional funds into ISSSTE itself turned out to be fortuitous, since a significant portion of this money went into upgrading its antiquated information technology and strengthening its management systems. The sums allocated for healthcare, mortgages and other purposes included allowance for overheads, which ISSSTE was able to spend on personnel training and management restructuring. The institute also undertook a major expansion of investment in infrastructure, equipment and personnel for ISSSTE hospitals.

In the end, the take-up of the recognition bonds was relatively low: only about 14.7% of those eligible opted for the bonds.\(^{21}\) To a great extent, this appears to have reflected an a priori conviction among many workers that, since the government had clearly not introduced the reform with a view to making ISSSTE pensions more generous, the new scheme would necessarily be less generous than the old PAYG system, even though the latter had undergone substantial modifications. Moreover, the communications strategy at
the implementation stage – in contrast to the adoption phase – proved to be rather weak, with many federal employees receiving inadequate or even inconsistent information about the choices they faced. Confusing and even contradictory advice would have tended to reinforce the status quo bias of risk-averse individuals. Finally, it is important to note that the vast majority of federal employees seem to have made their decisions in the final weeks before the deadline. While delays arising from problems with personnel records played a role here, this also appears to have reflected the existence of an option value to waiting: given the irreversible nature of the decision, officials uncertain about the best course to take had little incentive to make a choice before the deadline. This meant that large numbers of officials were making their decisions against the backdrop of growing international financial turbulence, which added to the perceived risks of the recognition-bond option – especially since the press was then reporting on the losses being incurred by the AFOREs as markets fell. Yet although acceptance of the recognition bonds was lower than had been hoped, the low take-up rate had little effect on the larger picture: only a small portion of the savings generated by the reform depended on the choices made by individuals, because the revisions of the PAYG system were substantial, too.

The arrangements for PensionISSSTE remain an issue. In principle, it is to be subject to the same regulatory requirements as other AFOREs, but problems and anomalies remain. PensionISSSTE was created by statute rather than by the usual procedures for creating AFOREs, so it is not clear that CONSAR could close it down if it came to that; there is even disagreement as to whether CONSAR can impose fines on it. The law says that PensionISSSTE cannot receive any public subsidy, but it also stipulates that the fund is to charge commissions that are below the average for all AFOREs, a requirement that critics believe may in due course necessitate hidden subsidies. There is also concern about the potential for investment and other decisions to be distorted by the policy priorities and interests of the unions and ministries on PensionISSSTE’s board. The scale of the potential problem will depend in part on the fund’s size, which depends on the uptake of recognition bonds and the choices of new federal workers: government officials expect PensionISSSTE to emerge as a middle-sized AFORE but it could be quite large, if the take-up of recognition bonds is unexpectedly high or new federal workers show a strong preference for it. The bigger it is, the greater are the risks that its peculiar status will affect competition in the sector as a whole.

From the outset, the administration’s intent was that the ISSSTE reform should set the stage for reform of the remaining PAYG public sector pension schemes, which covered around 1.5 million people and which are still under-financed and over-generous. Indeed, some were more generous even than the pre-reform ISSSTE scheme. The SHCP is actively trying to encourage the reform of these schemes, earmarking some oil revenues for use in financing pension reforms as an incentive to move, while indicating that those who resist reform longer can expect less generous treatment, as there are limits to the additional resources the government is prepared to make available to finance pension reform. Many state governments in 2007-08 did indeed begin discussing reform models with the SHCP.

Conclusion

Mexico’s pension reform path resembles those of many other countries, in which pension reform has proceeded in stages, beginning with the largest but most easily reformed segments of the system and working gradually towards those sectors where
resistance is likely to be fiercest. The 1997 IMSS reform set the stage for the ISSSTE reform in two important ways. First, it provided the basic reform template which the SHCP followed in the early 2000s; and, secondly, it made further reform harder to resist, as acceptance of the necessity and, indeed, inevitability of reform gradually spread, while support for the privileges of those whose pensions remained unreformed declined. In any case, the finances of ISSSTE were deteriorating rapidly, so even those with a vested interest in the status quo recognised that it could not be sustained over the long term. Crucially, however, those financial problems had not yet had reached a crisis point: by acting early, Mexico was able to execute the reform at lower cost than would have been possible if it had waited until a crisis was approaching.

Like the IMSS reform, the ISSSTE reform provided extremely advantageous terms for incumbent employees, honouring their acquired rights in full and shifting the full impact of reform onto newly entering cohorts. Such generous treatment of current workers was possible partly because the affected group was relatively small but also because the reformers were tackling the problem relatively early. The reform may have taken longer as a result – there was no threat of impending crisis to force speedy action, and the negotiations took years – but that, in turn, meant that much more effort was made to explain and sell the reform to those affected, which should reduce the likelihood of any reversal of the policy. It also helped the government to implement a major retrenchment of federal pensions without provoking the kind of public-sector strikes and protests that much more modest reforms have provoked elsewhere. Furthermore, although older workers were treated more favourably than their younger colleagues – as is commonly the case in major pension reforms – the use of a steadily sliding scale to calculate the value of recognition bonds meant that there was no sharp break in the treatment of different cohorts. Even younger cohorts of incumbent workers gained something – pension portability and confirmation of acquired pension rights for under 15 years of service, which the old system had not recognised. The state was also able to afford generous side payments, in the form of additional resources for ISSSTE, the voluntary savings plan and the unusually large share of pension contributions financed by the state.

Crucially, most of the concessions made to the unions tended to go with, rather than against, the grain of the reform and they therefore did not undermine its coherence. In addition, the intensive consultations with the unions that led to these concessions helped secure union support for the reform and to get it through a divided Congress. In the final stages of the debate, the direct involvement of a newly elected president in the negotiations with the major unions played a critical role.

Notes

1. OECD (2007:45).
4. On a broader definition of public debt, augmented to include liabilities created by the financial-sector rescue of the late 1990s and long-term investment schemes
(P RIDIREGAS), the public debt-to-GDP ratio was about 15 percentage points higher. See OECD (2004:51) for details.

5. It might be argued that the recognition bonds merely made implicit liabilities explicit, with no real impact on the government’s debt position. However, because pre-reform pensions were not legally secure – there was nothing to prevent the government from adjusting benefits at any time – the pre-reform liabilities were of uncertain value, whereas the recognition bonds conferred legal claims.

6. The president’s PAN, together with its allies, held a plurality in the Chamber of Deputies after 2000 but they were well short of a majority, and the PRI retained the largest bloc of votes in the Senate. After the 2003 midterms, moreover, the PRI emerged as the largest party in the lower house by a wide margin, falling just short of an outright majority. The PAN’s fortunes recovered somewhat in 2006, as did those of the PRD, but it remains the case throughout that legislation requires coalitions.

7. The Sindicato Nacional de Trabajadores de la Educación or National Education Workers’ Syndicate.

8. The Federación de Sindicatos de Trabajadores al Servicio del Estado or Federation of Public Service Workers’ Unions.


10. The SHCP estimated the contingent liability for the petroleum concern PEMEX alone at as much as 18% of GDP on certain assumptions (including constant employment levels).

11. On the IMSS reform, see Madrid (2003), chapter 3.


14. The bonds would carry a 3.5% real rate of interest, identical to that which was used to calculate the value of rights acquired to date when determining the nominal value of the bond.

15. The Comisión Nacional del Sistema de Ahorro para el Retiro or National Commission for Retirement Savings.

16. Though far less important than the ISSSTE reform, the IMSS reform was hotly contested, as it required IMSS employees to contribute to their own pensions for the first time. Their pensions had hitherto been financed entirely from general revenues.

17. Vázquez himself withdrew his signature from the motion.


19. Previously, the AFOREs had been permitted to charge a percentage of contributions, assets under management or a combination of the two. The bewildering variety of fee structures possible under such an arrangement made it difficult to compare effective rates of return across AFOREs. The SHCP and CONSAR took the view that limiting commissions to assets under management would make it easier for consumers to compare AFOREs. See Calderón-Colín et al. (2008); and Aguila et al. (2008).

20. At many meetings, participants were provided with tables that enabled them, on the basis of their age and years of service, to calculate the value of their recognition bonds on the spot.

22. The passivity of the public-sector unions may have contributed to poor communication at this stage of the reform: having supported the adoption of the ISSSTE reform, the unions subsequently worked hard to block attempts to overturn it, but they were largely indifferent to the choices made by individuals to take recognition bonds.

23. Mr. González Anaya, the architect of the ISSSTE reform, was subsequently put in charge of fiscal relations with the states inside the SHCP, in large part because the ministry sees pension reform as a central issue in its policy discussions with the states.

24. PensionISSSTE may prove an exception to this rule.
Bibliography


Summary and main conclusions

Poland introduced a comprehensive reform of its general old-age pension system in 1999, with a view to increasing work incentives, putting the system on an actuarially sound footing and ensuring that pension liabilities were fully financed. The reform established a defined-contribution, multi-pillar system involving: a pay-as-you-go (PAYG) pillar based on notional defined contributions (NDC) and administered by the Social Insurance Institution (ZUS), a mandatory funded pillar in which private pension funds manage individuals’ contributions, and a voluntary third pillar consisting of company pension plans and other savings vehicles. The 1999 pension reform stands out for having been designed and adopted under two successive governments of very different political orientations – perhaps the most important reform since 1989 to so transcend the partisan divide. This outcome reflected a number of factors:

- **A favourable economic context.** By the mid-1990s, the economy was growing relatively strongly, and Poland was still in a position to use a large share of its privatisation revenues to help finance the transition to a new pension system.

- **Broad agreement on the need for reform.** In contrast to the citizens of many West European countries facing pension reform pressures, ordinary Poles were increasingly dissatisfied with the existing system, which was widely regarded as both opaque and unfair, and were thus open to change. Public awareness of the need for pension reform was also growing as a result of widespread public discussion of demographic and fiscal pressures on the pension system.

- **De-politicisation and effective communication of the issues.** The creation of the Office of the Government Plenipotentiary for Social Security Reform (BPR) in 1996 enabled the authorities to de-politicise pension reform to some extent. The BPR was able to combine consultative and technocratic approaches to policy making: it was constantly engaged in public communications and dialogue with various stakeholders, but reform design was actually concentrated in a relatively non-political institution with a high level of expertise. The BPR thus played a critical role in educating both the elite and the public, creating a cross-party consensus in favour of reform and designing a reform plan. It also became the focal point for the provision of external advice and assistance, both technical and financial.

- **Concessions and “side payments” to potential opponents of reform.** The reform plan assigned a very important role in the new system to the existing Social Insurance Institution (ZUS), which helped to neutralise potential opposition from that quarter. Similarly, concessions were made to the trade unions over issues like
the retirement age, and they, too, were offered significant opportunities to be active participants in the reformed pension system – both major trade union confederations were among the founders of the pension funds created when the reform was implemented.

- **Deferral of the most difficult issues.** At certain points, both governments involved in passage of the pension reform legislation managed to maintain the momentum of reform by dealing with the less controversial issues first and deferring consideration of more contentious matters. This helped keep the reform process moving; indeed, the adoption of the first elements of the reform actually added to the pressure for action on the rest of the reform agenda. However, two of the most difficult issues – the question of “bridge pensions” for groups that will enjoy early retirement rights and the legislation governing payouts from the mandatory funded pillar – were deferred in 1997-98 and have not yet been resolved. The failure to resolve the early retirement question could prove costly over the long term.

### The context for reform

**Framework conditions for reform**

Cabinet instability has formed the backdrop to policy making for most of Poland’s post-communist history. The country had nine prime ministers during the 1990s, and the left-wing coalition in office during 1993-97, when the pension reform was designed and set in train, went through three prime ministers despite commanding large majorities in parliament. The fragility of governing coalitions, in turn, enhanced the deliberative role of parliament, particularly its lower chamber, the Sejm: party discipline tended to be weak, so even governments with solid majorities could find their proposals rejected or drastically modified by parliament. However, despite frequent changes of cabinet and fierce conflicts over specific measures, the need for reform was recognised by all parties and the overall direction of economic policy was fairly consistent, suggesting a degree of reform momentum.¹

Government instability apart, the context for pension reform in the mid-1990s was relatively benign. First, having weathered a sharp but short transition recession in 1990-91, Poland was enjoying steady real GDP growth of 5-7% per annum. Secondly, while the need for pension reform was increasingly clear, there was still time to act before Poland reached crisis point. The demographic situation, though set to deteriorate in the long term, was expected to remain relatively favourable for some years: the old-age dependency ratio (the ratio of retired to working-age population) was projected to remain roughly stable at around 23% through 2005 before rising to around 34% in 2020.² This comparatively low ratio stood in stark contrast to the high – and rising – system dependency ratio (the ratio of pensioners to employed persons), which by 1995 was above 60%.³ The high system dependency ratio made reform more urgent and pointed to the need for any reform to address incentives for early retirement, as well as the demographic challenges ahead.

Relatively favourable economic and demographic conditions could not easily be used to postpone pension reform for two reasons. First, by the mid-1990s, the system was widely regarded by ordinary Poles as unfair and untransparent.⁴ Secondly, Poland was still heavily indebted and running large budget deficits. Although the public debt-to-GDP ratio was falling rapidly from its 1993 peak of 88.7%, the country was being closely
monitored by creditors for any sign that the left-wing coalition elected in 1993 was relaxing its commitment to fiscal consolidation or market reform. Since budgetary subsidies to the pension system were widely seen as the principal cause of budget deficits (Table 6.1), there was a clear link between pension reform and fiscal consolidation. Finally, there was still an opportunity to use substantial future privatisation revenues to help finance the implementation of pension reform without imposing unacceptably high costs on specific cohorts.

Table 6.1. Expenditure on disability and retirement pensions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Retirement and disability pensions</td>
<td>8.6</td>
<td>12.6</td>
<td>14.6</td>
<td>14.6</td>
<td>15.4</td>
<td>14.6</td>
<td>15.2</td>
</tr>
<tr>
<td>Budgetary subsidies to the pension system</td>
<td>4.4</td>
<td>6.3</td>
<td>6.2</td>
<td>6.1</td>
<td>4.2</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Of which:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Insurance Fund (ZUS)</td>
<td>2.7</td>
<td>4.3</td>
<td>4.2</td>
<td>3.9</td>
<td>2.1</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Farmers’ Social Insurance Fund (KRUS)</td>
<td>1.7</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td></td>
</tr>
</tbody>
</table>


Actors and interests

Prior to 1996, the principal actors were the Ministry of Labour and Social Policy and the Ministry of Finance. The former was responsible for pensions policy and saw itself as the defender of pensioners’ rights, but fiscal pressure ensured that the latter also became deeply involved. Initially, there was relatively little outside input into policy, apart from popular protests against austerity measures and the proposals of a small number of individual experts. Some of the institutions that would play important roles later on emerged only in the course of the reform process itself. Thus, during the critical phase of the reform, the most important state actor was the Office of the Government Plenipotentiary for Social Security Reform (BPR), which had been created in an effort to alleviate the political conflict over pension reform. The ZUS itself and, later on, the newly created pension fund regulator UNFE also had considerable influence, particularly over implementation.

Poland’s major political parties were internally divided over the direction of pension reform. The post-communist Democratic Left Alliance (SLD) and the Polish Peasants Party (PSL), which governed in coalition together during 1993-97, both had strong factions opposed to structural pension reform. This was also true, albeit to a lesser extent, of the Electoral Action Solidarity (AWS)-Freedom Union (UW) coalition elected in 1997. Thus, the most difficult issues to resolve arose not between political parties but between the government and the trade unions, whose primary concerns were to prevent a sharp rise in the retirement age, to protect the early retirement rights of many of their members and to ensure the highest possible replacement rates. The two major union confederations were Solidarity and the All-Poland Alliance of Trade Unions (OPZZ). While oriented towards opposite ends of the political spectrum, the two unions made common cause over a number of pension reform issues, and when acting together, they were very influential. Moreover, while Solidarity could not be called a veto player, its links to the AWS meant that it was close to being one after 1997.

Numerous other internal and external actors were also involved. Research institutes and independent experts played a key role, particularly when it came to identifying solutions that could break the impasse that the reform debate seemed to have reached by 1995-96. Sectoral lobbies and, latterly, financial institutions – especially the newly
formed pension funds – also become important players at the implementation stage, and
the World Bank and other external donors played a key role in supporting the work of the
BPR both technically and financially. On the whole, non-governmental actors tended to
interact with the government rather than with each other. There was little dialogue among
them and certainly nothing like a big “concertative” negotiation. This may have made
reform easier, since it meant that no broad anti-reform coalition ever took shape.

**Issue history**

**Poland’s pension problems**

Post-communist Poland inherited from the old regime a PAYG, defined-benefit
system that was already out of balance. During the 1980s, the communist authorities had
extended early retirement rights to many professional groups in an effort to relax
labour-market pressures and calm the political situation. This trend continued in the
early years of the transition, as early retirement was used as form of passive
labour-market adjustment. By 1997, the effective retirement age had fallen to 58.3 for
men and 54.1 for women. Numerous other adjustments were made on an *ad hoc* basis
throughout the early 1990s, as governments and parliamentarians found themselves
captured between the need to hold down ballooning pension expenditure and strong popular
pressure for yet more generous benefits. While some changes were aimed at reducing
pension expenditure and strengthening the link between contributions and benefits, these
were more than outweighed by measures that increased the element of redistribution in
the system and its overall generosity. Replacement rates rose sharply, and pensions
spending nearly doubled relative to GDP (Table 6.1). The implicit pension debt (IPD) rose
to an estimated 220% of GDP by 1995. Contribution rates rose steadily but did not
come close to keeping up with the growth of pensions spending: the transfer from the
state budget to cover the social security system’s deficit exceeded 6% of GDP in 1993-94.

Increasingly, the *ad hoc* nature of pensions policy came to be seen as a major problem
in its own right. Constant tinkering with system parameters undermined public faith in the
system and added to the incentives for early retirement. Moreover, in the years prior to
1996, the Constitutional Tribunal repeatedly struck down specific changes to the system,
forcing the government to repay restored benefits in arrears. This further underscored the
need for a more systematic approach. The Tribunal consistently emphasised that its
rulings did *not* preclude a permanent, systemic change in the pension regulations,
provided that it was based on appropriate legislation. Its decisions were a major factor
in putting the question of more fundamental pension reform on the government’s agenda.

**Trends in public opinion**

Survey data in the mid-1990s showed steadily declining confidence in the pension
system and growing acceptance of the need for change. In early 1997, around 70% of
survey respondents found the system opaque, and nearly two-thirds said that they did not
understand the basis on which their pensions would be calculated. This lack of
understanding, moreover, contributed to a widespread perception that the system was
unfair: citizens understood that the weak link between contributions and benefits
closed a great deal of redistribution, much of which would have been hard to justify
on equity grounds, and most contributors seem to have believed that they were net *losers*
in this redistributive game – a conviction that gave defined-contribution proposals great
appeal. Finally, the increasingly gloomy picture of Poland’s pensions future presented by
the mass media convinced many Poles that the *status quo* was unsustainable as well as
unfair. This contrasted with the situation in many other OECD countries, where electorates remained deeply attached to unreformed systems, even if these were also seen to be unsustainable.

In principle, at least, ordinary Poles were open to fundamental change: heavy majorities favoured individual accounts, a strong link between benefits and contributions, and funded rather than PAYG financing. Moreover, opinion polls suggested that the public was receptive to proposals for private management of the second pillar. This openness to private-sector solutions reflected a lack of confidence in state institutions in general and in the ZUS in particular, as well as the relatively good reputation of Poland’s financial sector. For all its problems, Poland’s financial sector, though still immature, had been comparatively well regulated and had avoided major crises.

Policy process: formulation

Debates over more fundamental reform had been going on among experts and officials since 1991 but with little impact. The SLD-PSL coalition that took office in 1993 included a vague commitment to pension reform in its programme, but since pensioner support was critical to its election victory, the government’s initial inclination was to raise pensions. However, the Ministry of Finance began to advance proposals for a radical reform of the pension system, motivated in large measure by the need to commit the government to a credible path of fiscal consolidation and to show multilateral creditors and the markets that the coalition parties were committed to further economic reforms. The finance ministry’s ambitions were resisted by the Ministry of Labour and Social Policy (MPPS), which opposed both immediate measures, such as the indexation of pensions to prices rather than wages, and the finance ministry’s proposals for a structural pension reform involving a small basic pension financed on a PAYG basis, with most provision based on a mandatory funded pillar. The MPPS favoured a modest parametric reform that the finance ministry regarded as financially unsustainable. Government policy oscillated between the two throughout 1994-95, as neither ministry was able to impose its priorities on pensions policy.

A number of non-governmental reform proposals were also put into circulation around this time, of which the most significant were those put forward by the Solidarity trade union and the Institute of Labour and Social Affairs (IPSS); elements of both eventually found their way into the reform that was adopted. Solidarity proposed allocating privatisation bonds to all employees to help finance the creation of the funded pillar, and both Solidarity and the IPSS favoured a significant NDC element in the PAYG pillar.

De-politicising pension reform

The impasse was broken following a change of government in February 1996. The new prime minister, Włodzimierz Cimoszewicz, was a proponent of structural rather than a parametric reform, and he appointed a pro-reform minister, Andrzej Bączkowski, to the MPPS. As a former Solidarity activist, Bączkowski also provided the government with a link to the opposition. It was at this point that the government created the Office of the Government Plenipotentiary for Social Security Reform (BPR) under the auspices of the MPPS, with the minister serving as plenipotentiary. The BPR was given a mandate to prepare a pension reform plan for the government, thereby offering an exit from the inter-ministerial feuding that had characterised the preceding years. The BPR’s creation
reflected an awareness of the need for a broad cross-party consensus over pension reform – something not easily achieved in Poland’s often polarised political environment. Although formally subordinate to the MPPS, the BPR enjoyed a high degree of de facto independence. It also remained largely outside party politics, bringing together economists, lawyers, former officials and other experts from across the political spectrum. In November 1996, oversight of the BPR was transferred to the prime minister’s office, in an effort to enhance its autonomy further and shield it from those in the government who continued to oppose structural pension reform.28

The BPR soon became the principal conduit for external assistance in the field of pension reform. A large number of international donors and other agencies played an active part in the preparation of the reform, including the World Bank, USAID, the UNDP, the British Know-How Fund, IFC, PHARE and SIDA. The World Bank helped create and fund the BPR, which was run on a day-to-day basis by a Polish World Bank official on secondment, and BPR representatives worked closely with donor organisations to identify needs and facilitate coordination of their assistance. Sweden, a pioneer of NDC systems, helped finance the development of the NDC scheme and assisted in the creation of the model used to run simulations of it.

In addition to the BPR itself, the finance ministry, the State Treasury and the ZUS took part in the preparation of the reform. Though initially unenthusiastic about the BPR’s plans, the ZUS did not oppose structural pension reform. Its cooperation reflected in part the fact that the reform plans envisaged a substantial role for it in the new system, as the clearinghouse for all pension payments, the administrator of the first pillar and the agency responsible for managing the new system of individual accounts. Moreover, the ZUS president was a political appointee and could thus be disciplined for failing to back the government line.29

“Security through Diversity”

The BPR’s draft reform plan, called Security through Diversity, was presented in February 1997 (Box 6.1). Its primary aim was to ensure that the system could provide adequate old-age pensions to the population in a financially sustainable way. Putting an end to the array of sectoral privileges and other redistributive elements of the system that were reckoned to be inequitable was a second major goal. In addition, there was a clear awareness of the need to protect the acquired rights of older workers, in particular, in order to avoid further interventions by the Constitutional Tribunal; to encourage longer careers and individual saving; and to diversify the risks facing future pensioners. Creating a clear, stable link between contributions and benefits was seen as critical to ensuring that incentives to work longer and contribute more were effective.30 It was also hoped that such a link would limit the scope for political manipulation and the granting of special benefits to particular lobbies. Both the BPR and the World Bank also saw pension funds emerging as long-term institutional investors, capable of stabilising the financial system and stimulating growth.
Box 6.1. The pension reform plan

The reform plan devised by the BPR drew on elements of the various reform proposals already circulating by 1996. The old-age retirement system was to be separated from the rest of the social insurance system. The basic payroll tax rate would remain unchanged at 36.59%, but would be split roughly 50/50 between employer and employee, rather than being paid entirely by the employer. Just over half of this contribution, equivalent to 19.52% of gross wages, would be directed towards old-age pension provision, with the balance going towards other mandatory social contributions. Old-age pension provision was to be organised in three pillars:

- A substantial PAYG pillar would remain, retaining 12.22% of gross wages; these would, however, now be credited to individual accounts on an NDC basis. A demographic reserve fund was to be created to relieve pressure on the PAYG pillar caused by demographic shifts.¹
- The balance of the mandatory pension contribution (7.3% of gross wages) would go to a new, fully funded pillar, under the management of private pension funds. These funds were to be subject to rules on portfolio diversification.² Individuals participating in the second pillar were to choose pension funds by 30 September 1999 or be assigned to one at random.
- The third pillar would be voluntary, consisting of company pension schemes and other savings instruments.

First-pillar benefits would depend on lifetime contributions,³ uniform indexation rules and age at retirement. A minimum pension was to be guaranteed to anyone with 25 years of contributions,⁴ but the PAYG system would not otherwise be redistributive. The reformers argued strongly that redistribution in favour of the elderly poor was better executed outside the old-age pension system. Second-pillar benefits would depend on contributions, age at retirement and the efficiency of the fund’s investment strategy. The rules for calculating benefits would not change for those retiring through 2006, and benefits awarded before the reform would not be affected. The pre-reform acquired rights of those retiring after 2006 were to be converted into “initial capital” in the new system (Chloń et al., 1999:15).

Participation in the second pillar was to be mandatory for those born in 1969 or later,⁵ while those born prior to 1949 were to retire under the existing PAYG system. The cohorts born in 1949-68 could choose whether or not to join the second pillar; those who did not opt for a second-pillar account by the end of 1999 pay the full 19.52% of gross wages into their first-pillar NDC accounts. BPR estimates suggested that, for those around 50 years of age, there was not much to choose between participation and non-participation in the second pillar, but for older workers it looked risky (Góra, 2000). Although the “official” retirement age was to converge at 62 for both men and women, this was simply to be the age at which one could acquire entitlements; there would be no obligation to exit the labour market at any particular age, and, indeed, replacement rates were expected to rise significantly for those opting for longer careers (Table 6.2).

¹ This fund now accumulates and invests any surplus in the first pillar plus one percentage point of the wage bill transferred to it in 2002-08.
² Funds operating in pillar 2 could hold no more than 40% of assets directly in the shares of companies quoted on the stock exchange 40%, 5% in foreign assets, 10% in secondary market securities, 10% in National Investment Funds and 15% in local government bonds.
³ The reform made provision for contributions from the state budget on an individual’s behalf for periods of national military service, parental leave and nursing disabled children.
⁴ It was to be paid from age 65 to people pension entitlement was less than 49% of the average pension. A benefit financed from general revenues (not pension contributions) would top this amount up to the minimum level. This was intended to separate redistribution (via the budget) from the lifecycle reallocation of income (via the pension system).
⁵ This contrasts with the situation in Hungary, where the courts ruled against any attempt to mandate participation in the second pillar for those already working.
The reform drew on a number of foreign models, of which the Chilean and Swedish were most closely studied. However, the BPR’s proposed three-pillar system included a much larger first pillar than the Chilean model and a larger role for pension funds than the Swedish system. There was still considerable resistance to a funded system, and even advocates of pension funds were aware of the risks involved in over-reliance on funded schemes in Poland’s still rather immature financial system. Moreover, Poland’s very large implicit pension debt would have made it prohibitively expensive to shift rapidly from PAYG to a primarily funded system. Ensuring that a significant share of payroll taxes continued to flow into the first pillar in return for PAYG benefits later on was a way of holding down the visible pension system deficit during the transition period. Finally, the BPR argued that a more balanced division between pillars would allow for better risk management: the rate of return in the first pillar would be linked to the growth of the contribution base and would thus depend on labour-market and productivity performance, while the funded pillars would depend on rates of return on investment portfolios.

Table 6.2. Projected average replacement rates under the new system
As a per cent of final salary

<table>
<thead>
<tr>
<th>Age at retirement</th>
<th>60</th>
<th>61</th>
<th>62</th>
<th>63</th>
<th>64</th>
<th>65</th>
<th>66</th>
<th>67</th>
<th>68</th>
<th>69</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>First pillar</td>
<td>21.8</td>
<td>23.2</td>
<td>24.7</td>
<td>26.3</td>
<td>28.0</td>
<td>29.9</td>
<td>31.8</td>
<td>33.9</td>
<td>36.2</td>
<td>38.7</td>
<td>41.3</td>
</tr>
<tr>
<td>Second pillar</td>
<td>22.9</td>
<td>24.5</td>
<td>26.2</td>
<td>28.0</td>
<td>30.0</td>
<td>32.2</td>
<td>34.5</td>
<td>37.0</td>
<td>39.7</td>
<td>42.6</td>
<td>45.9</td>
</tr>
<tr>
<td>Total</td>
<td>44.7</td>
<td>47.7</td>
<td>50.9</td>
<td>54.3</td>
<td>58.0</td>
<td>62.1</td>
<td>66.3</td>
<td>70.9</td>
<td>75.9</td>
<td>81.3</td>
<td>87.2</td>
</tr>
</tbody>
</table>

1. Calculated on the assumption of a notional rate of return of 3% for the first pillar and 6% for the second.


In drafting the reform plan, the BPR left two highly problematic areas to one side – disability pensions, which were to be separated from the old-age pension system, and pensions for rural citizens belonging to the Farmers Social Insurance Fund (KRUS). Both omissions were substantial. The share of disability pensioners in the total number of pensioners had doubled in the first half of the 1990s, as such pensions were used in an attempt to limit the increase in registered unemployment, and by 1996 they accounted for more than 35% of all pensioners in Poland. Reform of disability pensions was nevertheless postponed until after the reform of old-age pensions. The reform was initially intended to encompass the farmers, too, but the BPR was forced to drop this idea at an early stage, partly because it presented particular technical difficulties but chiefly because the power of agrarian lobby in general and the PSL in particular made it impossible to touch the KRUS.

Policy process: adoption

The government initially intended to bundle reform of the first (PAYG) pillar with the creation of the new second pillar and the re-regulation of supplementary private pensions. However, with elections approaching, the government reversed course and opted for a strategy of un-bundling the various strands of reform, leaving the more controversial issues to be resolved after the elections.
The passage of the pension reform bills

Security through Diversity was approved by the government in April 1997 and then by the Tripartite Commission on Social and Economic Affairs, the body created in 1994 to provide a forum for policy discussions involving unions, employers and the state. The commission gave its overall backing to the plan despite union misgivings about specific aspects of the reform. These largely concerned the PAYG pillar and the reduction or elimination of sectoral early retirement rights. Anxious to get as much of the reform adopted as possible before the elections, the government left these issues to one side and focused on the less controversial elements of the reform – the three statutes that were to govern the new second and third pillars, as well as to regulate the use of privatisation revenues to make up the losses to the PAYG first pillar arising from the diversion of some social security contributions to the second.35 This, it was hoped, would oblige whatever government emerged from the elections to press ahead with the reform. Thus, the first three pension reform bills were adopted in the summer of 1997 with roughly 90% support in parliament.36 However, the second pillar legislation covered only the “accumulation phase”: the bill on how second-pillar benefits were to be paid proved contentious37 and was left to the next phase of the reform.

The unions’ acceptance of the defined-contribution principle probably reflected the fact that many unionised workers expected to gain from a less redistributive system, while those workers with early retirement rights were more focused on preserving them than with maintaining defined benefits. The principal losers from the move to an NDC system were likely to be low earners and those with shorter contribution histories, two groups that had been treated relatively generously under the old scheme. Since women tended both to earn less and to have shorter and more frequently interrupted careers, women in low-wage sectors were probably the group with the most to lose in relative terms, but their interests were weakly represented in the pension reform debate.38 The creation of a privately managed mandatory funded pillar won the support not only of Solidarity, which had called for a funded pillar in its own proposals, but also of the OPZZ. This reflected, at least in part, Solidarity’s generally pro-market orientation, but it was also partly the product of the BPR’s deliberate strategy of involving the unions in discussions of reform from an early stage and of the opportunities that the reform offered to them. When the reform was implemented, both Solidarity and the OPZZ emerged among the founders (in international joint ventures) of new pension funds.39

The BPR’s efforts to build cross-party support for reform turned out to be critical, because the left lost the September elections, and the reform was left to be completed by the AWS-UW coalition. Despite its enormous importance, pension reform did not loom large in the election campaign, largely because all the major parties had all backed the first-wave legislation just a few months earlier and the leading pension policy experts in both coalitions supported the BPR plan. Two pensioners parties did contest the election, one on the left and the other on the right, but neither won seats in parliament.40 Nevertheless, continuation of the reform was not guaranteed, as many AWS parliamentarians wanted to revise the plan.41 However, the UW supported the BPR’s approach, as did most AWS officials directly involved in pension policy. In the end, therefore, the coalition confirmed its commitment to the overall approach of Security through Diversity. The new cabinet was in any case under pressure to proceed, because the adoption of the legislation on the second and third pillars had – as the BPR hoped – strengthened the pressure for change. A pension-fund regulator (UNFE) was created under the law on the funded pillar in early 1998 and licensing of pension funds began in the summer. This led to the rapid emergence of an important new pro-reform
constituency, as domestic and foreign financial institutions mobilised to enter the new pensions business. Abandoning the reform would have damaged Poland’s reputation in the eyes of major financial institutions.

The government wished to move fast once it had made its decision, since the AWS aimed to launch four major reforms – of pensions, healthcare, education and local government – during its first year in office, in order to allow them to begin bearing fruit before the next elections, due in 2001. However, the first-pillar legislation proved far more difficult to pass than the earlier bills:

- The unions demanded, and won, some significant changes, the most important of which was the decision to leave the retirement age unchanged: women could claim benefits from age 60 and men from 65. However, no agreement was reached on the critical question of early retirement rights for specific sectors. The government proposed to retain these only for older cohorts of workers, but the unions wished to keep them indefinitely. Both Solidarity and the OPZZ therefore gave qualified negative assessments of the draft legislation to parliament. However, the government’s concessions on the retirement age and other issues were sufficient to ensure that they did not mobilise against the reform.

- Some AWS parliamentarians remained sceptical of reform, and the opposition SLD, while expressing support for the reform concept as a whole, contested numerous specific provisions and backed the unions’ defence of early retirement rights. While the government commanded a majority sufficient to override any SLD amendments, maintaining party discipline proved difficult on issues where AWS deputies were inclined to back the unions.

Faced with line-by-line challenges to the bills in the Sejm, the government reverted to the approach of the previous cabinet, tackling the least controversial questions first. Discussion of the first-pillar law was suspended and attention shifted to a less contentious framework law on the Social Insurance System, which was passed in October with few significant changes.

When parliament returned to the first-pillar law and the vexed question of occupational early-retirement schemes, the trade unions mobilised, and representatives of the Solidarity miners’ union even occupied the labour ministry for two days in early December. Negotiations involving employers, the unions and the government produced partial agreement on relatively generous transitional arrangements for older workers. The fate of younger cohorts was left open, as it was agreed that after the reform was adopted, a special commission would devise an equitable way to deal with sectoral early-retirement rights. In order to give credibility to this promise, the first-pillar law was amended to require that so-called “bridge pensions” be addressed in a subsequent law. Eligible workers would be able to retire early, but the ZUS would not be responsible for their pensions until they reached the statutory retirement age. Instead, “bridging” schemes covering the period between early retirement and the statutory retirement age would be financed separately from the ZUS system, by additional contributions from employers and the state budget. Absent such a legislative commitment, the unions would probably not have agreed to negotiate after adoption of the legislation, since their bargaining position would, in normal circumstances, have been greater prior to adoption. The compromise also owed much to the Solidarity trade union’s links to the government. The deal over bridge pensions opened the way to adoption of the final piece of pension legislation. Yet again, however, legislation on second-pillar pay-outs was postponed: a
bill was prepared but then withdrawn, because the government was dissatisfied with it. Since there would be no benefits payable under the second pillar until 2009, it was thought better to take more time to resolve this issue.

**The government’s communications strategy**

From the beginning, the BPR saw public communications as a key part of its role, and its efforts to shape the public debate over pension reform began even before *Security through Diversity* was unveiled. Successive plenipotentiaries and other BPR-affiliated experts published widely in the press. When the first legislation was prepared for submission to the Sejm, the BPR launched a major public information campaign in support of the reform. BPR and ZUS staff were trained in public communications, and seminars were organised to target various audiences, including politicians, journalists, employers and union officials. Brochures addressing different constituencies were published. These public-relations efforts appeared to enjoy some success: press coverage of the reform plan was broadly favourable, and polls showed majority public support for reform, although survey responses suggested that this was based more on doubts about the old system than understanding of the new. The focus of these communications efforts shifted as the debate evolved. Thus, the initial phase of the campaign, in 1997, focused on the need for, and inevitability of, reform. In 1998, the BPR devoted its communications resources to making case for the reform of early retirement privileges and explaining the new legislation on pension funds to the public.

Throughout the discussion, the BPR laid great stress on the essential clarity and fairness of an NDC system, rather than its distributive consequences. They also emphasised the need to ensure its financial sustainability. These were arguments that polling data indicated would resonate with Polish citizens. The BPR and the government also avoided drawing attention to the fact that reform would leave many Poles with a less generous relationship between contributions and benefits. They emphasised instead that fairly high replacement rates could be sustained, provided individuals worked full careers and that they invested their second-pillar contributions in successful funds. The complexity of the changes adopted – and, in particular, uncertainty about the rates of return that second-pillar contributions would earn – in any case made it very difficult to estimate the savings generated by the reform. The government estimated that it would reduce the value of ZUS’s pension liabilities to 2050 by 268% of GDP, but most of this reduction simply reflected the shifting of resources to the second pillar. How much pension wealth was actually eliminated would depend chiefly on the performance of that pillar. The government’s projections suggested that the total value of first- and second-pillar pension liabilities to 2050 would fall by around 59% of GDP, a reduction of around 13% in pre-reform pension wealth. However, since the reformed system looked increasingly unsustainable, reform advocates could argue that the system was bound to become less generous over time – the question was whether it would do so as a result of a reform or a crisis.

**Policy process: implementation**

**Management**

Although the final legislation was not adopted until December 1998, the new arrangements took effect in April 1999, having been postponed from 1 January of that year at the request of the ZUS. Electoral considerations seem to have played a role in
the decision to introduce the new system so quickly – the government wished to ensure that any initial disruptions associated with controversial social reforms were forgotten by the time of the next parliamentary elections, due by late 2001. However, the more urgent concern seems to have been the government’s desire to synchronise pension and healthcare reforms. Because both reforms were to require quite radical changes in the operation of the ZUS, which also handled healthcare contributions, it was important to coordinate them. This was no easy task, since the technical requirements the two reforms imposed on the ZUS were somewhat different. The complexity of this challenge might have pointed to the need for a longer transition period, but the finance ministry was determined that healthcare reform should proceed on schedule, for budgetary reasons, and it feared that any postponement of the implementation of the pension reform would entail costly delays in healthcare reform as well.

The tight implementation timetable meant that the new system got off to a difficult start. The first months of 1999 revealed the lack of readiness on the part of the ZUS to administer the new system: its offices were overwhelmed by customers anxious to obtain information, forms and software. Many employers were also ill prepared to handle the administrative tasks the reform imposed on them. The organisation of individual accounts lagged, so that by end-1999, the ZUS had transferred barely 50% of accumulated second-pillar contributions to the pension funds, and around 40% of accounts were still dormant. Legal anomalies remained in certain texts prepared for the reform, and some 17 legislative amendments were needed in 1999-2001 to address technical problems. A strategic plan for restructuring the ZUS was adopted in late 1999. James and Brooks (2001) argue that these disruptions were the price the government paid for conceding a key role in the new system to the ZUS, which might otherwise have opposed the reform. However, the very tight implementation timetable was dictated by the determination to proceed with healthcare reform on schedule, and it is not clear that any other agency could have been set up quickly enough to get the new system up and running faster than the ZUS did. The UNFE, after all, was a brand new agency, and it, too, attracted much criticism, not least from some of the reform’s own architects, who insisted that it was both overly politicised and under-prepared for its task.

Communications

Implementation, too, was accompanied by a large-scale government information campaign aimed at providing citizens with the information needed to make decisions about their pensions and allaying fears about what the reform might imply for different groups. This was particularly important for intermediate cohorts, who had to decide whether or not to join the second pillar. Polls in late 1998 found that only 30% of respondents thought they understood the reform. In 1999, the focus of the BPR’s communications efforts therefore shifted from opinion leaders to the general public and from advocacy to explanation. Specific messages were aimed at specific cohorts, on account of the differential impact of reform on different age groups. In addition, the ZUS conducted its own information campaign, mainly to explain the new reporting obligations and other arrangements.

The new pension funds’ advertising campaigns, which were launched in the second half of 1998, complicated the authorities’ efforts. While the private funds’ marketing efforts served to raise awareness of, and perhaps even support for, the reform, they often gave rise to unrealistic expectations or simply confused the issues facing future pensioners. The UNFE was criticised for failing to police these media campaigns more effectively, but its powers in this field were limited. With the funds spending about
20 times as much as the government on public information, there was a risk that the government’s explanation of the reform would simply get drowned out. However, the BPR made repeated adjustments to its media strategy as the year went on, and by late 1999, focus groups were beginning to show a growing awareness of difference between funds’ advertisements and government information. Overall, the proportion of Poles saying that their knowledge of the reformed system was adequate rose from under half in March 1999 to 80% a year later.59

Issues left unresolved

As noted above, some significant design decisions were left to the implementation phase, the most important of which concerned bridge pensions. In principle, early retirement was to be limited to those working in conditions that might be particularly deleterious for the health of the employee (e.g. miners) or for those whose work required exceptionally good physical or psychological fitness (e.g. pilots). While there were those in the BPR who rejected bridge pensions and early retirement as the solution for such sectors, these criteria commanded broad acceptance. In practice, though, successive attempts to apply them foundered on resistance from those threatened with losing their early retirement privileges. The issue of early retirement was only resolved in more or less definitive fashion in respect of the miners, who in 2005 were granted a special regime of their own following violent protests in Warsaw. The government itself tried to block this change, which was adopted by parliament and signed by the president – against the cabinet’s wishes – during the run-up to the 2005 parliamentary elections.60 In 2008, the new centre-right government returned to the bridge pension issue, putting forward draft legislation that would restrict such pensions to older workers in a very limited range of occupations – effectively preserving early retirement rights for around 130 000 workers, out of an estimated 1 million entitled to early pensions under the old system. However, the proposals encountered strong trade union opposition, and the legislation finally adopted preserved these rights for 270 000 workers.61

Other issues that were postponed in order to facilitate adoption of the reform likewise remained largely unresolved a decade later, although the government had begun to put forward its legislative proposals. With the first benefits from funded schemes due to commence from January 2009,62 the government at the end of 2008 adopted an interim solution, requiring the pension funds themselves to make the relevant payments from 2009 to 2013, but legislation on the creation of special “pension houses” to handle the second-pillar pay-outs beyond 2013 was vetoed by the president, leaving the issue unresolved. An attempt to reform the farmers’ social security scheme in 2003-05 achieved little, leaving the KRUS system almost entirely untouched, although the government returned to the issue in 2008. There was some progress on the reform of disability pensions beginning in 1999, and both inflows into and expenditure on disability pensions fell; however, a large part of this reduction was offset by the growth of “pre-retirement benefits”, which are paid by the ZUS to unemployed older workers who have made pension contributions for a certain number of years but are too young to draw old-age pensions.63 These benefits seem to have replaced disability as an early exit channel for many workers.64

Conclusion

In many respects, political and economic conditions in Poland in the second half of the 1990s were exceptionally favourable for pension reform. Public dissatisfaction with
the inefficiencies and inequities of the existing pension system was substantial, and both
fiscal and demographic pressures were contributing to a growing consensus on the need
for change. Ad hoc changes had made the system more complex and less fair. The public
was thus open to reform on equity grounds, and the principles of a defined contribution
system were relatively easy to explain to citizens. That said, a degree of creativity and
political leadership was required to break the reform impasse in 1996. Here, the work of
the BPR was critical: its independence, its emphasis on consultation and public
communications, and its ability to bring together experts and officials from across the
political spectrum meant that the reform process was characterised by a high degree of
continuity despite the change of government in 1997. Staffing and operations changed
little under successive plenipotentiaries, all of whom shared a basic commitment to
structural pension reform. While it is not possible to de-politicise an issue like pension
reform entirely, the BPR succeeded in greatly reducing the degree of partisan polarisation
around the issue. In assigning the BPR the role they did, the Polish authorities also
managed to combine participatory and technocratic approaches to policy making: while
technical design was concentrated in a relatively de-politicised institution with a high
level of expertise, numerous actors and groups were involved in discussion of the
reform.⁶⁵

Passage of the reform still required a degree of tactical finesse on the part of the
governments involved. In some cases, potential opponents, like the ZUS and the trade
unions, were drawn into the process and allowed to play important roles in the new
system. Moreover, both cabinets sought to maintain the momentum of reform by adopting
the least controversial elements first, postponing more difficult issues. They calculated
– rightly, up to a point – that progress on some issues would add to the pressure for
movement on others. However, this was not always the case. The most difficult issues
were deferred until the implementation phase, and some of these – most notably the
legislation on second-pillar pay-outs – were still awaiting resolution in early 2009.

Notes

1. See OECD (1997) for an overview.
2. GUS estimates cited in Góra and Rutkowski (1998); somewhat less pessimistic
forecasts may be found in Styczeń (1999), but even these show the demographic
dependency ratio rising over the period, to exceed one-third by 2050. See also Cain
and Surdej (1996).
4. See below for details on public opinion.
6. Góra (2000). Since privatisation revenues were of a one-off character, they could not
resolve the larger fiscal or pensions problems, but they could cover the one-off costs
associated with the transition to a new pension system.
7. The OPZZ had been set up by the communist authorities in 1982 as an alternative to the then banned Solidarity, and it remained close to the post-communist SLD in the first years of transition.

8. Of particular prominence were the Centre for Social and Economic Research (CASE), the Institute of Labour and Social Affairs (IPSS) and the Gdansk Institute for Market Economics (IBnGR), as well as leading experts from Polish universities.

9. Bialas et al. (2001) base this claim on interviews with participants in pensions policy making.

10. Bialas et al. (2001) compare the structure of action to a hub and spokes arrangement, with the state at the centre, rather than a more complex network.

11. By 1989, almost 25% of working-age population had the right to retire ahead of the statutory retirement age (60 for women and 65 for men) without any reduction in benefits. Since the benefit formula gave little weight to length of service, there were few, if any, incentives to work longer.

12. Both age- and disability-based retirement: by 1995, there were some 3.4 million disability pensioners, equivalent to around 37% of the pensioner population – roughly double the proportion of 1989; 70% of these were in category 3 (partial invalidity), the category most open to abuse. See OECD (1997:107).

13. The few measures aimed at strengthening the system’s finances were more than outweighed by those that increased the system’s generosity. See Chłoń (2000) for details.

14. For a summary of these early changes, see Chłoń (2000).

15. The average replacement rate rose from 52% in 1989 to 64% in 1994; this compares with levels of 57% in Hungary, 49% in the Czech Republic and 42% in Slovakia at that time (OECD, 1998:65).

16. Including old-age and disability pensions and both the employees’ and farmers’ social security systems (i.e. ZUS and KRUS, respectively).

17. The present value of accrued rights of pensioners and workers under existing pension rules.


19. The prospect of benefit cuts or other adverse changes in the rules encouraged older workers to take their pensions before such changes could be applied to them.


21. Unless indicated otherwise, the polling data cited in this paragraph are drawn from Chłoń (2000). See also Bialas et al. (2001).

22. Bialas et al. (2001:26) cite data from a 1996 survey giving a similar result.

23. On the tenor of press coverage, see Bialas et al. (2001).


25. Numerous proposals were floated, including some that were quite radical: as early as 1991, ZUS President Wojciech Topinski and actuary Marian Wisniewski proposed a multi-pillar system modelled loosely on that of Chile but the proposal did not get serious political attention.
26. In 1991, Poland began indexing pensions to wage growth; this made for significant savings at first, as real wages were falling at the time, but it became very costly after real wage growth resumed in 1993.

27. See Orenstein (2004:44) for a comparison of the finance ministry, MPPS, Solidarity and IPSS proposals.

28. Bączkowski’s sudden death led to the appointment of a new plenipotentiary, Jerzy Hausner, who did not, however, succeed Bączkowski as minister, so the transfer effectively enhanced the BPR’s status and ensured that the new plenipotentiary would not be subordinated to the new labour minister, who was openly critical of the BPR’s approach to reform.

29. Orenstein (2000:43) contrasts Poland’s experience with the situation in Hungary, where the reform plans threatened the role of the Pension Insurance Fund, which consequently opposed reform, and where the fund itself was governed by an independently elected body with significant trade union representation.

30. If the contributions-benefits link is opaque or unstable, the combination of a defined contribution system with actuarially fair calculation of benefits may not affect individual incentives much.

31. In this respect, it also differed from the three-pillar model proposed by the World Bank (1994).

32. See, in particular, James (1998); and Orenstein (2000).

33. The larger the implicit pension debt, the greater appeal of such an approach. James (1998) and James and Brooks (2001) provide empirical support for the claim that countries with large IPDs in the old system (owing to generous, earnings-related benefits and universal coverage) tended to retain large public PAYG pillars after reform and to rely less on funded schemes.

34. While both pillars would be subject to common aggregate shocks, the BPR sought to provide some hedge against those risks more likely to affect one or the other. For example, they argued that the PAYG pillar would be more susceptible to political bargaining, population ageing and poor labour-market outcomes, while the funded pillar would be more exposed to the risks of inflation and financial market volatility. See Chłoń et al. (1999) and Góra (2000).

35. There was general agreement on the attractiveness of using privatisation revenues to finance the pension reform; the question was how to use them. In contrast to Solidarity, the reformers opted to use these revenues to finance the first pillar.

36. Only a few right-wing deputies from the “Christian-national” wing of the Solidarity movement continued to oppose fundamental reform. See Orenstein (2000:45).

37. There were disagreements concerning the institutions to be involved in paying second-pillar benefits, as well as to the rules governing, e.g., the actuarial tables to be used and the fees that might be levied.

38. Zajicek et al. (2007). The gender bias at work here is substantially offset by the use of unisex life expectancy tables when calculating benefits at retirement. Since women live longer, the calculation of identical benefits for men and women with identical contributions histories benefits the latter.
39. The fact that both unions represented private- as well as public-sector workers probably played a role as well: efficiency and low payroll taxes are likely to matter more to private-sector workers in competitive environments than to public-sector employees. See James and Brooks (2001).


41. In particular, a number of them wished to revisit the Solidarity trade union’s proposals for allocating privatisation bonds to employees’ funded pension accounts.

42. Having initiated the reform, the SLD could hardly come out against it now. However, the balance of power within the SLD’s Sejm fraction was shifting towards opponents of the reform, so fighting the government on the detail of the legislation was an attractive strategy.

43. The law on the social insurance system covers such basic issues as: eligibility for mandatory social insurance; the level of contributions and their allocation among various purposes; and allocation of responsibility for payment of contributions. The most noteworthy changes in this law were the inclusion of the uniformed services in the system and the exclusion of prosecutors; judges had already pressed for, and won, exemption from the system on constitutional grounds.

44. The miners’ concerns were intensified by the on-going restructuring of the coal industry, which was resulting in mine closures.

45. It was this amendment that induced the miners to leave the ministry building.

46. Other miners groups – including the OPZZ-affiliated union and some small independent branch unions – called for further protest, but they, unlike Solidarity, were unable to affect the legislative process. See Orenstein (2000:49).

47. Bialas et al. (2001) provide the best overview of communications efforts of successive governments in this field.


49. The World Bank’s IPD estimate of 220% of GDP covered only the value of liabilities accumulated by pensioners and workers at the time of calculation; the government’s estimate of 462% to 2050 included the liabilities that would be accumulated in future as well. A further small discrepancy arises from the fact that the government’s estimates are somewhat later than the Bank’s, and the IPD was rising before the reform.

50. In mid-1998, the ZUS indicated that it would not be ready by January. However, the delays in securing adoption of the final legislation probably made a postponement inevitable anyway.


52. The ZUS had to set up individual accounts for all participants in both systems. However, the two systems encompassed different groups of people (not all participants in one were enrolled in the other), as well as different contribution bases in terms of participants’ assessed income, and the requirements for managing contributions collected were very different.

53. See Chłoń-Domińczak (2004:271-75) and OECD (2000) for details on these initial “teething pains”.

---

54. In the event, a large majority of those in the middle cohort joined the second pillar.

55. The campaign relied on a variety of media, including press and television advertisements, brochures and the organisation of call centres to answer citizens' questions. The BPR published simulations showing circumstances under which individuals could expect a better pension under each option. The BPR also put software on the web allowing people to calculate pensions under both options for themselves. See Chłoń (2000) for details.

56. Advertisements depicting retirees enjoying exotic island holidays, for example, attracted much criticism: they did not actually promise spectacular pension benefits, but the implication was clear.

57. For example, by excluding the effect of administrative charges on annuity values from their calculations.


59. See the polling data cited in Chłoń (2000).

60. The MPPS estimated that the miners' package would cost PLN 10 billion in 2020, equivalent to PLN 500 per head of population. In fact, the real winners were not the miners but the mining companies. There was never any real doubt that any resolution of the early retirement issue would leave the miners with early retirement privileges. However, a bridge pension solution would have involved employers' contributions as well as budgetary revenues. Under the August 2005 legislation, the budget picked up the costs of the miners' privileges.

61. The bridge pension legislation adopted by the centre-right government was vetoed by the president, but parliament overrode this veto in December 2008, with leftist parties backing the government against the president.

62. The payments in question would initially be very small, as they would affect only the minority of individuals born in 1949 who had opted to participate in the second pillar; such retirees would have accumulated only limited claims under the new system. Nevertheless, it was critical that arrangements for paying these benefits be put in place.

63. Eligibility requirements are complex, depending on age, gender, contribution record, employment history and reasons for redundancy. For details, see ZUS (2007).


Bibliography


Chapter 7

Poland: Reform of the farmers’ social security system, 2003-05

Summary and main conclusions

Poland operates a separate social security system for farmers, administered by the Farmers’ Social Insurance Fund (KRUS). It covers pensions, healthcare, disability and other benefits for the large part of the population that meets the rather broad criteria for inclusion in the farmers’ system. The KRUS-administered system has long been criticised for its perceived inequities, the perverse incentives it creates, which distort both labour and land markets, and the fact that the system is nowhere close to being self-financing: budget subsidies to the KRUS in the early 2000s were of the order of 2% of GDP each year, far in excess of non-KRUS spending on agriculture. Yet the power of the agrarian lobby ensured that the system was largely untouchable until late 2003, when fiscal pressures compelled the government to propose changes to the system as part of a larger fiscal adjustment plan, the so-called “Hausner plan”. These reform proposals were debated – and gradually scaled back – over the course of the next two years. In the end, only minor changes to the administration of the farmers’ social security system were adopted. Indeed, the 2006 budget actually increased its generosity.

The frustration of the attempt to introduce more fundamental changes to the KRUS system reflected a number of factors, the most important of which included:

- **An unfavourable political context.** A surge in support for populist parties of both left and right made reforms even more difficult to pursue than they had been in the 1990s. There was considerable resistance to the austerity package, of which the KRUS reforms were a critical part, within the governing coalition and among the electorate. This eventually brought down the government. The minority cabinet that followed did try to press ahead with KRUS reform but lacked both an electoral mandate and the necessary support in parliament.

- **Declining pressure for immediate action.** Accelerating growth and an improving fiscal position in 2004-05 reduced the sense of urgency about reform among parliamentarians, who were increasingly focused on the approaching elections and who proved unwilling to accept austerity measures that they had previously seemed ready to adopt. Yet the improvement in public finances owed much to cyclical and one-off factors and did not mitigate the structural distortions created by the KRUS system.

- **The strength of vested interests.** Poland’s agricultural sector is large, politically powerful and committed to protecting the KRUS system. No government since 1989 has commanded a parliamentary majority without the support of at least one agrarian party, and rural interests are well represented inside most of the larger, more urban parties as well. Moreover, farmers’ trade unions and other agrarian...
organisations enjoy considerable influence in the Ministry of Agriculture, which is usually headed by a farmer, and they are directly involved in running the KRUS itself.

- **Mixed motives for reform.** Because the immediate impetus for reform was the need for fiscal consolidation and the measures proposed aimed to curtail KRUS subsidies quickly rather than to address the system’s structural defects, the case for structural reform was not clearly articulated in its own right. The focus on short-term savings also left the government little scope for softening the impact of reform via “side-payments” such as increased spending on other rural programmes.

**The context for reform**

*Framework conditions for reform*

*Fragile coalitions*

Although the need for both fiscal consolidation and a reform of the farmers’ social security system was widely accepted by early 2003, the political context for reform was otherwise unfavourable. Opinion polls showing a surge in public support for populist parties of both left and right, along with the fragmentation of the centre-right as a whole, made reforms even more difficult to pursue than they had been in the 1990s, and the two governments that held office while the Hausner package was under consideration had very limited parliamentary and popular support:

- In 2003, when the Hausner plan was first conceived, the governing Party of the Democratic Left (SLD) no longer commanded a parliamentary majority, having evicted the Polish Peasants Party (PSL) from the coalition a few months earlier. The popularity of the government was extremely low by this time, owing to disappointing economic performance, corruption scandals, and widespread criticism of its handling of issues like healthcare reform. Moreover, the SLD itself was increasingly divided over economic policy, particularly the Hausner plan. In March 2004, the party finally split and the government fell.

- A minority cabinet under former finance minister Marek Belka held office from May 2004 until the September 2005 elections. The Belka cabinet had no electoral mandate and even less parliamentary support than its predecessor. While the government did last until the end of the parliamentary term, it was not always clear that it would do so, and the very real prospect of early elections made fiscal consolidation still more difficult.

*Fiscal pressures*

Although the need to reform the farmers’ social insurance system had long been recognised, it was fiscal pressures that finally prompted the government to try to act. The sharp slowdown of the early 2000s contributed to a rise in government deficits, which had been trending slowly downwards until 1999 (Table 7.1). The public debt-to-GDP ratio began to rise, and in 2003, it exceeded 50% on the methodology set out in the Act on Public Finance (APF). The government projected that it would breach the 55% level in 2004-05 and hit 60% in about 2006. This was a critical concern, because the APF imposed a series of automatic adjustments at these thresholds. At 50 and 55%, the rules
were not too constraining, but once the debt rose above 60% the authorities would be obliged to impose severe restrictions on public expenditure; this, it was feared, might trigger a recession. There was also considerable concern that EU entry would lead to a rise in the budget deficit. Finally, the Monetary Policy Committee (RPP) of the National Bank of Poland (NBP) cited fiscal uncertainty in explaining its reluctance to loosen monetary policy more rapidly. Poland had begun to bounce back from the 2001-02 slowdown, but growth was still sluggish, and the government believed that the RPP was not moving aggressively enough to support the recovery. By early 2003, therefore, the state of public finances had become a source of major concern to the government.

Table 7.1. Growth, deficits and public debt

<table>
<thead>
<tr>
<th>Year</th>
<th>Real GDP growth</th>
<th>Deficit</th>
<th>Public debt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>4.8</td>
<td>-2.4</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>4.1</td>
<td>-1.7</td>
<td>41.9</td>
</tr>
<tr>
<td>2000</td>
<td>4.1</td>
<td>-2.9</td>
<td>38.7</td>
</tr>
<tr>
<td>2001</td>
<td>1.1</td>
<td>-2.9</td>
<td>39.7</td>
</tr>
<tr>
<td>2002</td>
<td>1.4</td>
<td>-3.9</td>
<td>45.2</td>
</tr>
<tr>
<td>2003</td>
<td>3.8</td>
<td>-4.1</td>
<td>51.5</td>
</tr>
</tbody>
</table>

1. Data as of April 2004. Substantial revisions to GDP and fiscal data have since yielded much higher deficit figures for these years but lower estimates of the debt. However, these revisions came too late to affect the politics of the Hausner plan. For the revised figures, see OECD (2008:62).


This sense of urgency did not last. Both growth and fiscal performance improved during 2003-05. Real GDP expanded by 3.8% in 2003 and 5.3% in 2005. The reported deficit fell to 2.5% of GDP in 2005 and the debt-to-GDP ratio remained roughly stable, at about the level of end-2003. This apparent improvement in the deficit and debt positions was largely the product of cyclical and one-off factors, but its effect was to reduce the pressure for immediate fiscal adjustment and thus to undermine support for the austerity package, particularly in the run-up to the 2005 general election.

The Polish farm sector

Debates over the reform of farmers’ social security have long been shaped by the relative backwardness of Poland’s rural sector, which depends heavily on income transfers provided by the Social Insurance Institution (ZUS), which administers Poland’s general social insurance scheme, and the Fund for Farmers’ Social Insurance (KRUS), which runs the special social insurance regime for farmers. Pensions and other public transfers accounted for around 23% of farm households’ incomes in 2000, and for more than 40% of income received by the more than 70% of Polish farms of less than 7 hectares (Table 7.2). For most farms, KRUS benefits accounted for a larger share of income than farming. Any reduction in the generosity of the KRUS system was therefore bound to have a direct and palpable impact on the incomes of rural households, so reform could be expected to provoke significant resistance.
Table 7.2. Sources of farm income by farm size

<table>
<thead>
<tr>
<th>Source of Income</th>
<th>0-1 ha</th>
<th>1-7 ha</th>
<th>7-15 ha</th>
<th>15+ ha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household annual income</td>
<td>88</td>
<td>80</td>
<td>81</td>
<td>233</td>
</tr>
<tr>
<td>Per capita</td>
<td>93</td>
<td>60</td>
<td>65</td>
<td>160</td>
</tr>
<tr>
<td>Per cent of national average earnings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net farm income</td>
<td>11</td>
<td>7</td>
<td>44</td>
<td>73</td>
</tr>
<tr>
<td>Net income from non-farm business</td>
<td>2</td>
<td>10</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Non-farm wage income</td>
<td>43</td>
<td>43</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>Income from transfers</td>
<td>46</td>
<td>38</td>
<td>30</td>
<td>11</td>
</tr>
<tr>
<td>KRUS benefits</td>
<td>14</td>
<td>16</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>Other social insurance benefits</td>
<td>25</td>
<td>19</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Share in total number of farms</td>
<td>–</td>
<td>70</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

1. World Bank rural household survey data divided by GUS whole economy data.
2. World Bank rural household survey data divided by World Bank rural household survey data.
3. Data for farms of at least 1 hectare only.


The farmers’ social security system

Though described as a form of social insurance, the KRUS system is in fact a non-means-tested social welfare programme for the rural sector. KRUS pensions are on average somewhat smaller than those of the general social insurance system, but KRUS contributions are minimal and do not vary according to income. Moreover, non-pension KRUS benefits are broadly comparable to those of the general social insurance system. Consequently, contributions cover only a small fraction of KRUS expenditures, over 90% of which are financed by the state budget. The system thus places a considerable strain on public finance, while doing little to foster needed restructuring in the farm sector.

Eligibility. To be eligible to participate in the KRUS system, an individual must ordinarily be: a farmer owning at least one hectare of arable land or engaged in agricultural production that does not require ownership of such a quantity of land; the spouse of a qualifying farmer; or an adult member of a qualifying farmer’s household who either derives his income from farming or from other work related to the agricultural activity of the farm, or is otherwise inactive and a dependent of the farmer. Eligible individuals whose farm is less than two hectares in size are exempt from unemployment insurance premia but eligible for unemployment benefits.

Contributions and benefits. A KRUS contributor pays a quarterly contribution equal to about one-third of the contribution paid by a minimum wage earner in the general system. KRUS offers farmers and their extended families a range of social insurance benefits (old-age and disability pensions, health, sickness and accident insurance) roughly comparable to those offered by the general regime. All members of a KRUS household qualify for health insurance, even if they have no income, which is not the case with the ZUS. Men are eligible for retirement at 65 with 100 quarters of contributions or at 60 with 120 quarters. Women may retire at 60 with 100 quarters. On retirement, farmers
must give up all of their land in excess of one hectare, but they may transfer it to another family member and they may continue farming. KRUS pensions tend to average about 134% of the level of the minimum pension in the general scheme, and they vary little in size since the number of contribution years is the only source of difference and its weight in the benefit formula is limited. KRUS disability pensions are relatively easy to obtain, because the applicant must show only that he/she is unable to farm (fitness for other work is not assessed). Moreover, there is only one class of disability pension under the KRUS: no partial disability benefits are available. Thus, the KRUS covers around 19% of the workforce but accounts for 25% of disability pensions. In 2005, 43% of KRUS pensioners were on disability pensions, and the number of newly granted disability pensions was substantially higher than the number of newly granted old-age pensions.

**Organisation and finance.** The KRUS’s peculiar legal status has generated criticism. It is not a special-purpose fund (like the Social Insurance Fund), a budgetary institution or a state-owned company. Indeed, it is not a legal person at all, a fact which critics argue makes effective monitoring of its finances and activities more difficult. The KRUS’s president, however, is part of the state administration: he is appointed by the Prime Minister after consultation with the Council of Farmers, and for most of KRUS’s history, he has reported to the Ministry of Agriculture. The KRUS administers a number of distinct funds, of which the two most important are the pension fund and the so-called “contribution fund”. The latter is in some ways the most controversial. Unlike the pension fund, which relies on budgetary subsidies to cover over 95% of its expenditure, the contribution fund is, as its name implies, mostly contribution-financed. It covers short-term benefits (accident, sickness and maternity), as well as such activities as accident prevention and rehabilitation, and the provision of summer holidays for farmers’ children (up to 140 000 per year). Critics have long argued that KRUS management has too much discretion managing this fund and that it is often used to promote the local interests of the various parties, unions and farming organisations represented on the Council of Farmers, which oversees the KRUS.

**Actors and interests**

Poland’s rural sector has traditionally been quite powerful, owing to its size – although agriculture generated only about 3% of GDP in the early 2000s, it continued to account for 20-25% of total employment – and to the existence of a wide range of farmers’ trade unions and other rural organisations. There are agricultural trade unions descended from both Solidarity and the old communist trade union, OPZZ, as well as numerous regional bodies and a nation-wide network of “agricultural circles” represented at national level by the National Union of Farmers, Circles and Agricultural Organisations, one of the most powerful lobbies in Poland. In addition, Poland has long had parties of both left and right dedicated to defending rural interests. No government since 1989 has commanded a parliamentary majority without the support of at least one such party, and rural interests are well represented within the major “urban” parties as well. In the 2001-05 parliament, predominantly agrarian parties held almost 30% of the seats in the lower chamber, the Sejm. In addition, the governing SLD had a substantial agrarian wing of its own. Not surprisingly, therefore, KRUS reform had never been a prominent part of the major parties’ electoral platforms.

The farming lobby can generally count on strong support from the Ministry of Agriculture – most ministers have been farmers themselves, and the traditional orientation of the ministry has tended to be to look after the interests of its rural clients, not least in resisting pressure for cuts in the KRUS system from the Ministry of
Economics, Labour and Social Policy (MPPS) and the Ministry of Finance, the two ministries that have tried to advance the reform of the KRUS system. Except for a brief period in 2004-05, it has been the agriculture ministry that oversees the KRUS. The KRUS itself was also an active participant in reform debates of 2003-05. While not a policy-making body, it was represented in the working group on reform formed within the MPPS, and its assessments of various approaches to reform proved influential. This marks an important contrast with Poland’s experience of reforming the general pension system: unlike the ZUS, the KRUS was never won over to the cause of reform, and there was no analogue to the Office of the Government Plenipotentiary for Social Security Reform to overcome, or at least circumvent, opposition from within the ministries concerned.

**Issue history**

**Social spending and the pressure on public finances**

During the 1990s, government spending as a share of GDP declined gradually but steadily, from around 46% in the mid-1990s to stabilise at around 42-43% in the early 2000s, rather high for a country at Poland’s level of per capita income. High levels of public expenditure largely reflected the maintenance of an extensive but inefficient system of social transfers. Altogether, social spending accounted for around 45% of general government expenditure. The finance and economics ministries thus took the view that sustainable fiscal adjustment would require structural changes that reduced the burden of social expenditure. This implied a need to address the budgetary costs of the KRUS system. By the early 2000s, government transfers to the KRUS were covering over 90% of its outlays (Table 7.3). The subsidy to the KRUS in 2002 accounted for 8% of the state budget, or around 2% of GDP. This figure dwarfed the rest of public funding for agriculture – production support in 2003 amounted to only about 15% of budgetary expenditure in support of the KRUS.

<table>
<thead>
<tr>
<th>Table 7.3. Expenditures and revenues of the KRUS</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Table 7.3 Expenditures and revenues of the KRUS" /></td>
</tr>
</tbody>
</table>

1. Including both disability and old-age pensions.


**The KRUS system as an obstacle to restructuring in agriculture**

Yet while fiscal pressures provided the immediate impetus for the attempt to reform the KRUS system in 2003, reform was needed in any case in order to foster the long overdue restructuring of Polish agriculture and to correct the manifest inequities of the system. Productivity in agriculture was seen as low and declining, and restructuring was proceeding very slowly. The KRUS system was by no means the only reason for this, but...
it played an important role in preserving existing inefficiencies in the farm sector. The system impeded the reallocation of both land and labour:

- Poland has too many very small farms. In 2002, around 90% of farms were smaller than 15 hectares, reckoned to be the minimum size needed to support a household. The potential gains from consolidation were enormous. Yet the land ownership requirement for KRUS participation created incentives for farmers to hold onto small plots of land and even encouraged some urban dwellers to buy small tracts in order to join KRUS and avoid contributing to the general social insurance scheme. Thus, the number of small uneconomic farms was increasing in the early 2000s. However, attempts to curtail abuses by making it harder for those with non-farm incomes to join the KRUS risked discouraging farmers from diversifying their activities – something the government wished to promote.

- Savings on contributions easily exceeded the difference in benefits, so moving from the KRUS to the ZUS imposed a heavy tax on a worker’s exit from agriculture. Payroll taxes for non-farm employees in the early 2000s amounted to 32.4% of gross wages, while the contributions paid to the KRUS by qualifying farmers typically amounted to around 3-4% of the farm’s household income. Moreover, the move into non-farm employment could impose costs on other members of the household (inactive or in shadow employment) who depended on KRUS-provided healthcare and other benefits.

The KRUS was also open to criticism on equity grounds. The absence of means-testing meant that urban taxpayers were financing subsidies for wealthy farmers as well as poor ones. Moreover, the system facilitated tax evasion, since KRUS benefits provided basic social protection to individuals in shadow employment, provided they were employed at least part of the time in agriculture or were qualifying members of KRUS households. While direct evidence of tax evasion or benefit fraud was scarce, household surveys in the early 2000s found higher incomes on very small farms than on middle-sized ones (Table 7.2) and recorded exceptionally high concentrations of non-farm rural employment in the sectors known to rely most heavily on informal employment relationships.

**KRUS reform debates in the 1990s**

By the mid-1990s, there was a growing awareness that the KRUS system was inefficient and abuse-prone. It was hoped that the system might be restructured in the context of the broader reform of pensions undertaken in 1996-99. The Office of the Government Plenipotentiary for Social Security Reform harboured such ambitions, and its pension reform proposals did initially include important changes to the operation of the KRUS. It soon became apparent, however, that the general pension reform would not go forward unless the KRUS was exempted; the PSL was the junior partner in the ruling coalition at the time and it refused to countenance such a reform of the KRUS system. Towards the end of the decade, the government proposed tightening the eligibility rules for KRUS disability pensions, as well as a re-qualification benefit to finance transitions out of the sector for those who had to leave farming on account of health problems but who could still be employed in another sector. These proposals were modelled on changes to disability pensions in the general scheme that had been reckoned to be a success, but they were never adopted.
The apparent untouchability of the farmers’ social security system reflected more than just the influence of the farming lobby, important though that was. First, the authorities did not welcome the prospect of further labour shedding in, and out-migration from, the rural sector at a time of high unemployment. Secondly, despite the KRUS system’s inefficiency and inequity, there was a concern that incorporating the farm sector into the general social security system would simply be too technically complex and politically contentious to be worth the effort, especially after the general pension scheme was reformed in 1999. It was not clear that farmers would be able to cope with the onerous record-keeping and reporting requirements of the new general pension system. Finally, the limited understanding of the KRUS system among those not directly involved in it created an information asymmetry in favour of system insiders that made it relatively easy for them to discredit reform proposals in the eyes of public and elite opinion. Lack of public understanding of the system also meant that there was relatively little debate on KRUS reform among the wider public: while the Hausner package was widely discussed in the media, little attention was focused on the structural issues involved in overhauling the KRUS system.

Policy process: formulation

The debate over KRUS reform options

In early 2003, the government was under increasing pressure to check the growth of public debt, while also finding resources to finance the new spending necessitated by EU entry, by then just over a year away. The Ministry of Finance’s initial set of proposals for adjustment included changes to the indexation of pensions and other benefits, higher contributions for KRUS members, the transfer of some wealthier farmers to the general social insurance system, and the liquidation of some discretionary off-budget funds administered by the KRUS. Just days after the finance ministry proposals were unveiled, the PSL was ejected from the governing coalition in a dispute over road-tax legislation. Once in opposition, the PSL joined other agrarian parties in rejecting any reduction in support for the farm sector. The government’s loss of its Sejm majority31 significantly reduced the chances that the fiscal package would be enacted swiftly.

The Ministry of Economics, Labour and Social Policy (MGPPS) also criticised the finance ministry’s package publicly, triggering a conflict between the two ministries that ran for the rest of the spring. The MGPPS argued that boosting growth must be the government’s first priority and that fiscal adjustment would be easier once Poland had returned to growth rates of around 5%. On this basis, the ministry began preparing an alternative plan, which envisaged a temporary rise in the deficit in 2004-05, ahead of a fiscal consolidation to take place in 2006-08.32 This proved appealing to the government, particularly in view of the SLD’s unpopularity, the approach of the EU referendum and the growing probability of early elections. By mid-June, the government was clearly shifting towards the MGPPS line, and its Minister, Jerzy Hausner, emerged as the cabinet’s chief economic policy maker.

From the beginning, it was clear that changes to the KRUS system would be part of any package that emerged from the MGPPS. The declared goal was to separate social assistance from social insurance, while retaining adequate welfare provision for the rural poor. The MGPPS argued that EU accession created a unique opportunity to address these challenges. When Poland entered the Union, anyone owning at least one hectare of agricultural land would be entitled to direct payments from the Common Agricultural Policy (CAP). Initially, these would be at a reduced rate, but they would rise over time.
This would provide a substantial boost to incomes in the sector and, because the CAP funds would flow to landowners regardless of effort, it would make it all the harder to justify the retention of the existing KRUS-related subsidies.

The working group on KRUS reform and the preparation of the Hausner plan

A working group on KRUS reform had already been formed under the aegis of the MGPPS in early 2003. It brought together experts on agricultural policy, social policy, farm incomes and other relevant issues. The MGPPS working group identified a number of priorities for reform, including: more restricted access to the KRUS system; closer government oversight of a number of KRUS funds that were largely at the disposal of the Council of Farmers; the introduction of some element of progression in KRUS contributions, linked either to income or farm size; an increase in total KRUS contributions; and a shift in budgetary subsidies to the KRUS from benefits to contributions. This last measure was seen as a necessary concomitant to any attempt to link KRUS contributions to income. Contribution rates would be raised overall, but budgetary subsidies would be used to pay the part of the new, higher contributions of poorer rural households. Better-off KRUS contributors would have to pay their way or switch to the ZUS. The MGPP believed that CAP payments to producers would provide them with considerable additional income from which to pay KRUS contributions, making EU accession a particularly propitious moment for reform.

The working group considered the question of aiming for the liquidation of the KRUS, but came to the conclusion that it would be easier, both technically and politically, to reform it, at least in the first instance. Nevertheless, the liquidation issue remained important in the political debate, since many opponents of reform later argued that the working group’s ultimate goal was indeed the elimination of the KRUS system. The group soon conclude that a real reform of the system would be much easier if oversight of the KRUS itself were transferred from the Ministry of Agriculture to the MGPPS. This reflected both the conviction that the agrarian lobby was too influential in the Ministry of Agriculture – there were doubts about whether effective oversight of KRUS’s contribution fund could be exercised via that ministry – and the view that KRUS reform should be seen primarily as a question of social policy, which was the responsibility of the MGPPS, rather than agricultural policy.

In September 2003, Minister Hausner presented the broad outlines of the plan (Box 7.1). The proposed 2004 budget was actually expansionary, providing for a fiscal stimulus estimated at 1.7% of GDP in an effort to meet the government’s 5% real GDP growth target. Over the medium term, however, the Hausner plan aimed to rein in the deficit and to do so chiefly via structural measures on the spending side instead of tax increases, ad hoc cuts and the postponement of expenditure. It was meant to achieve an eventual improvement in the structural balance of 1.3% of GDP.

Policy process: adoption

Government weakness and the erosion of the Hausner package

Knowing that the unions, in particular, did not like the proposed cuts in social spending, the government sought to broaden the civic dialogue over the fiscal adjustment plan, moving it away from the Tripartite Committee for Social and Economic Issues, the usual venue for such consultations. A number of special discussion fora addressing different aspects of the plan were organised under the auspices of the MGPPS: there were
groups concerned with disability policy, special retirement regimes, the retirement age and KRUS reform. Each brought together experts and representatives of NGOs, as well as representatives of the trade unions and other social partners represented on the Tripartite Committee. This enabled the government to organise consultations with a different group of interlocutors for each major set of issues. The groups met weekly for about two months in late 2003, but, far from fostering a broader consensus, the exercise antagonised the major unions, who saw it as an attempt to dilute their influence. Some unions simply refused to participate.

Box 7.1. The Hausner plan

The package that came to be known as the “Hausner plan” envisaged three groups of measures: savings from public administration, some increases in tax revenues, and savings in the personal transfer system, including the KRUS. The plan included a number of proposed reforms of the KRUS, including:

- tightening eligibility for KRUS membership by requiring the wealthiest farmers to join the ZUS system and clarifying the definition of “farmer” so as to close off access to the system for land-owners employed in the urban sector or in “quasi-agricultural” activities;
- raising contributions from those wealthier farmers who remained with the KRUS, by linking them to income;
- linking pension benefits to contribution history, in an effort to encourage longer careers;
- introducing partial disability pensions for farmers who could work outside the farm sector, while reserving full disability for those who could not work at all;
- providing less generous indexation of non-pension benefits, which were to be adjusted only when cumulative inflation passed 5% rather than rising automatically with the CPI; and
- simplifying the KRUS’s structure and aligning it with that of the main Social Insurance Institution (ZUS), chiefly in an effort to strengthen government control over the premium fund and the prevention and rehabilitation fund.

Altogether, the Hausner package aimed to cut social expenditures by roughly 0.8% of GDP per annum during 2005-07. Measures affecting the KRUS were expected to account for around one-quarter of this total. Farmers’ organisations argued that this placed a disproportionate share of the cost of adjustment on the agricultural sector and was thus inequitable, but the MGPPS took the view that the allocation of the cuts reflected the biases of existing social expenditure. As a sector that received an exceptionally high share of non-means-tested social transfers, the farm sector could not but feel the effects of curbs on social spending more than most. These cuts would in any case be offset by the additional funds that would flow into the sector after EU entry.

1. The measures under this rubric included job cuts, liquidation or consolidation of state institutions, and better management of extra-budgetary funds and state agencies.
2. The plan included proposals to increase the contribution base for self-employed high-income earners; and the social security contributions of disability pensioners and the self-employed.
3. Such as breeders of aquarium fish and other pets.
4. This, in turn, would have necessitated some clarification of the rules for calculating farm income.
5. The original Hausner plan envisaged cuts in social expenditure equivalent to 0.6, 1.0 and 0.8% of GDP in 2005, 2006 and 2007, respectively.
Opposition from the unions was all the more significant in view of the government’s lack of a majority in the Sejm. When the package was first presented, prospects were not too bad. The government could still, if it maintained internal party discipline, muster almost enough votes for passage, and most of the opposition parties were willing to support some elements of the plan, even if none was ready to back it in its entirety. In early March, therefore, the government won initial approval for proposals to subordinate the KRUS to the MGPPS and transfer wealthier farmers to the ZUS system. This prompted loud protests from farmers’ organisations, which claimed they had not been consulted and charged that the transfer of the KRUS to the MGPPS was the first step towards its liquidation. The government promised consultations with farmers before proceeding further and indicated that it would pass the bulk of the legislation in April-May. However, tensions over the package finally provoked a split within the SLD. A breakaway group of SLD deputies left the party in late March to form Polish Social Democracy (SdPl), largely in protest against the austerity measures. This, in turn, triggered a change of government at the beginning of May.

The new government was even weaker in the Sejm. Distrusted by both left and right, the Belka cabinet was sustained in office largely by the major parties’ desire to avoid an early election rather than by any strong support for its programme. It was not in fact clear that a snap election could be avoided, though, and the prospect of one reduced deputies’ incentives to support fiscal adjustment. It soon became apparent that a marked softening of the Hausner package was to be the price the new government paid for holding office. The cabinet quickly abandoned around PLN 2.9 billion in planned cuts in social expenditure, the first of many revisions to the plan, virtually all of which involved concessions on social spending. The SdPl and its allies insisted that social policy be removed from the MGPPS, while the PSL indicated that it would support the government only if the plan’s proposals for the rural sector were dropped. With the agrarian-populist Samoobrona surging in the polls, the PSL could hardly compromise on KRUS reform without risking its position among rural voters. Despite these obstacles, the government pressed ahead with those elements of the plan that it believed it could push through the Sejm. In July, it won approval of a bill to eliminate the annual adjustment of disability and other benefits, replacing automatic indexation to the consumer price index with a mechanism providing for indexation every three years or whenever cumulative inflation hit 5%.

The rejection of income-linked contributions

Supporters of KRUS reform continued to hope that EU funds would make it palatable, and the government proceeded cautiously with a revised set of proposals. A new KRUS bill was finally approved for submission to parliament at the end of September. The bill’s main emphasis was on raising KRUS contributions and linking them to farm size, the closest available proxy for income available, given that farmers did not pay income tax and that most did not keep books. Its preparation involved extensive discussions with farmers’ organisations, and the government believed that they had accepted the proposals, albeit very reluctantly. However, when the bill reached the Sejm in mid-October, the farmers’ trade unions and other agrarian organisations mobilised against it. By that time, it was in any case apparent that growth was accelerating and the fiscal position was improving. With the prospect of a crisis or a wrenching adjustment under the terms of the Act on Public Finances receding, a growing number of parliamentarians were openly questioning the need for austerity. In any case, the relatively modest proposal put forward by the government offered only limited fiscal
benefits and would do little to correct the structural distortions created by the KRUS system.

The farm-size criterion was fiercely attacked, with representatives of the agrarian parties arguing that size and profitability were not necessarily linked and that such a scheme would hit middle-sized farms harder than large ones. KRUS officials also pointed out that using the farm-size criterion would create new incentives to subdivide holdings, at least on paper, and that this was contrary to government policy. Farmers’ organisations argued that higher premiums would in any case lead to lower compliance and that up to one-third of farmers would simply stop paying their KRUS contributions, a view that both the KRUS itself and the agriculture ministry publicly shared. Both KRUS officials and farming organisations also criticised the government’s desire to raise contributions rapidly rather than phasing them in over a long period. However, since the overriding concern of the MGPPS was to cut the KRUS subsidy, a long transition period would have been unacceptable.

In the end, no parliamentary fraction was willing to endorse the bill, which was withdrawn without ever coming to a first reading. The bill might have had some chance of passage if the government had offered increased benefits to those who paid higher contributions. This was what many farmers were demanding, and in any normal social insurance scheme, it might have made sense. However, the KRUS did not really function as a social insurance scheme. Since all KRUS contributors were heavily subsidised, the MGPPS took the view that it made no sense to offer higher benefits to any of them until their contributions were at least self-financing.

The government worked further on amendments to the legislation and held additional consultations with farmers, but the agrarian lobby was not to be persuaded and the Sejm was not willing to upset rural constituencies in the run-up to the elections. The government’s initial proposals for the 2006 budget actually changed the criteria for the allocation of social transfers to farmers in such a way as to make the system more generous. The government did win some changes in the organisation of the KRUS and a tightening of the eligibility requirements for individuals with non-farming incomes, but that was all. The latter measure brought no immediate reduction in the budgetary subsidy to the KRUS, since the individuals affected were contributors. After the elections, the new government announced that supervision of the KRUS would be transferred back to the agriculture ministry. This change was part of the agreement that the governing Law and Justice party concluded with Samoobrona in order to win its support.

Conclusion

The fate of the KRUS proposals must be seen in the context of the gradual but steady erosion of the Hausner plan: throughout 2004, the package was repeatedly revised, as various parties demanded the abandonment of proposed cuts. While the government managed to achieve over 80% of the plan’s targeted revenue increases and around 74% of the projected cuts in spending on defence and public administration, only 47% of the projected savings on social expenditure were ever realised, with two-thirds of this amount arising as a result of the changes to benefit indexation rules alone. Indexation apart, the KRUS emerged from the entire episode more or less unchanged, apart from some minor revisions to its governance.

The frustration of the attempt to reform the KRUS system reflected three factors above all. First, the governments in power during 2003-05 were too weak in parliament...
and too unpopular with the public to push the plan through, while the improving economic and fiscal situation – though clearly good news in its own right – reduced the sense of urgency about reform. Given the public’s dislike of the Hausner plan, Sejm deputies looking to the coming elections had strong incentives to treat the improved circumstances as a reason to postpone fiscal adjustment. Secondly, the power of the agrarian lobby ensured that those elements of the Hausner plan that were adopted did not target the agricultural sector. Finally, the immediate stimulus for reform was the need for fiscal consolidation. As a result, the measures proposed focused on curtailing KRUS subsidies quickly rather than on addressing the system’s structural defects.

The fiscal and structural issues were not unrelated, but the focus on immediate fiscal concerns meant that the case for structural reform was overshadowed by the need to reduce subsidies to the farm sector. While the architects of the initial reform package were well aware of the rent-seeking opportunities created by the KRUS system, as well as the impediments to agricultural restructuring that it created, little was done to communicate the case for reform on structural grounds to a wide audience. The emphasis on fiscal objectives was bound to antagonise agrarian interests while leaving the government little scope for softening the impact of reform via phased increases in contributions or “side-payments”, such as increased spending on other rural programmes. The irony here is that the reformers to some extent saw CAP funds as a sort of side-payment, inasmuch as the inflow of EU cash was expected to smooth adoption of the reform. However, since the CAP payments were in no way contingent on the reform, farmers rejected such linkage, opting to take the CAP funds while still defending their existing KRUS rents.

Notes

1. The “Rywingate” scandal concerned allegations of influence-peddling in the SLD. While public hearings did not turn up sufficient evidence to charge anyone associated with the party of wrongdoing, they did significant damage to the government’s public standing.

2. Precise evaluation of the stance of fiscal policy in Poland has been complicated in recent years by the wide range of definitions and concepts employed. Deficit figures are complicated by the debate over whether or not to count as government revenue the compulsory contributions to individuals’ privately managed retirement accounts. Debt figures are also calculated according to two different methodologies: one based on the Maastricht criteria for Economic and Monetary Union and the other based on the broader concept of all general government financial liabilities, calculated under rules contained in the Polish Act on Public Finance (APF). Unless indicated otherwise, this case study uses the ESA95 deficit estimates then being reported by Poland to the EU (i.e. including the pension contributions and the associated transfers in the public sector) and the national debt estimates. As will be seen, the national debt estimates were highly relevant, because they formed the basis for the application of legal provisions concerning limits on public debt. See OECD (2004:75-76) and OECD (2006:65, 70) for more detail on these definitional issues.
3. The definitions used in the APF differ in several respects from the public debt calculations used under the Maastricht Treaty. In particular, the latter excludes risk-adjusted guaranties and treats the matured liabilities of the public sector differently; see OECD (2004).

4. Calculated according to the methodology used for assessing compliance with the Maastricht criteria for Economic and Monetary Union (EMU), the debt was far lower, at around 42-43%, but this mattered less at the time, as Poland was still some years from EMU entry, and it was not the debt criterion that posed the main obstacle to eventual entry into the euro area in any case.

5. No new state guarantees could be issued; sub-national governments would be required to pass balanced budgets; the government would have to submit a fiscal consolidation plan to the parliament; and it would be legally impossible to pass a resolution authorising a budget deficit. For more detail on the provisions of the APF then in force, see OECD (2002:79).

6. While EU transfers to Poland would easily exceed Poland’s contribution to the EU budget and institutions, it was feared that co-financing of EU programmes and other new spending necessitated by EU accession might more than offset this. See OECD (2004:80-83) for details.

7. See OECD (2006:66-68) on this issue. Subsequent data revisions also led to much higher deficit estimates than were available in 2004-05; had these been available at the time, the politics of the Hausner plan might well have been rather different.

8. See, e.g., Gazeta Wyborcza, 14 and 16 October 2004; Rzeczpospolita, 16 December 2004.

9. In some circumstances, individuals who do not meet the definition of “farmer” may apply to join the KRUS system.

10. World Bank (2001) estimates the KRUS contribution at one-sixth of the average contribution paid to the general scheme. It should be noted also that most farmers are not subject to VAT and that farmers pay income tax only on specific agricultural products (such as greenhouse production, poultry, mushrooms, silk and fur). The proportion of farmers producing such products was estimated at 2-5% in 2004. Income tax on such products may be assessed based on average production norms or cost accounts.

11. In the latter case, they must transfer their land to a successor (possibly a family member) under the age of 40. This policy aims to improve the age structure of the agricultural labour force.

12. The minimum pension in the ZUS system is defined as the basic pension for KRUS calculation purposes; for each year of contributions, 1% of the basic pension is added to the benefit. There is a non-contributory component of the pension benefit that ranges between 80 and 90% of the basic pension (it declines as contribution years increase), which sets an effective floor on the value of KRUS pensions.

13. Although it relies heavily on subsidies from the state budget, the KRUS is not supposed to be budget-financed and only some of the funds it administers are covered by the law on the state budget.

14. Estimates of farm-sector employment in 2002-03 varied widely, from around 19% on the Central Statistical Office (GUS) data to as much as 26-27%, according to the

15. “Agricultural circles” are a form of cooperative which also provides services, including tractor hire.

16. The two largest agrarian parties, the PSL and the Self-Defence of the Republic of Poland (Samoobrona), together commanded over 20% of the seats in the Sejm, and the right-wing League of Polish Families, which was also disproportionately reliant on rural votes (though not an avowedly “rural” party), held another 8%.

17. The SLD was in fact the largest party in rural Poland. See the election results and exit poll data presented in Szczepanik (2003).


19. Including pensions, housing, health, active labour-market policies and a range of different benefits (family, survivors, disability and unemployment).

20. Such high levels of social spending are not, in principle, entirely surprising in a country with chronic high unemployment, but, despite Poland’s very high rates of unemployment, very little of this expenditure was on unemployment benefits or active labour-market measures.

21. The KRUS’s financing gap had to be filled by the budget, because the fund itself could not borrow. When the KRUS was set up in 1991, the aim was to increase the share of farmers’ contributions from roughly one-quarter of outlays to around one-third, but instead this share fell sharply over the decade that followed. See Chłoń (2000:1). KRUS officials point out that a large part of the subsidy is used to pay benefits granted prior to 1989 to individuals who had never contributed to the system but who had transferred their land to the state on retirement.

22. The ratio of KRUS expenditures relative to GDP peaked about at 2.5% in the 1990s before falling back to stabilise at around 2% as a result of the falling value of KRUS pensions relative to wages, a decline in the number of beneficiaries, and a decline in the number of new pensions granted. See Chłoń (2000).


25. OECD (2004:166-7) estimates that that over 90% of Polish farms need to shed labour and that consolidation could yield increases of 12-25 times in output per worker on an FTE basis. Economy-wide returns from agricultural restructuring could increase potential output growth by up to 1% per annum for a period of several decades. See also World Bank (2001).

26. There was some consolidation of medium-sized farms into larger ones, but smaller farms showed a tendency to multiply. Thus the number of KRUS contributors rose by 15% during 1996-2002.


28. Indeed, the KRUS system encouraged low-productivity workers to enter the agricultural sector, and was seen as one reason why rural outmigration fell from sharply in the late 1990s and then even reversed itself for a time in the early 2000s.

30. The most significant of these was the stipulation that the subsidies needed to guarantee a minimum pension to a farmer would provided against be the security of the farmer’s fixed assets (land, buildings, machinery). At the pensioner’s death, a part of those assets, proportional to the subsidy, would revert to the state.

31. The SLD and its remaining partner, the Labour Union (Unia Pracy), retained a comfortable majority in the Senate.

32. In fact, there were to be substantial savings relative to the baseline in 2005, but the deficit would not begin to fall until 2006.

33. Hausner, it should be noted, was the Government Plenipotentiary for Social Security Reform in 1997, when resistance to from the PSL and its allies had necessitated leaving the rural sector out of the general pension reform. He was therefore familiar with the KRUS system and reportedly regarded it as part of the “unfinished business” of the 1999 pension reform.

34. A World Bank analysis of the KRUS system came to the same conclusion; see Fitzpatrick (2005).


36. The Tripartite Committee had sharply criticised the Kołodko proposals in the spring, with Minister Hausner himself in the chair.

37. It dropped proposals to: reassess disability benefits granted for an unlimited period; begin a phased increase in the retirement age for women; and reduce the generosity of family benefits. It also agreed to soften the rules on senior citizens and disability pensioners earning extra income Gazeta Wyborcza, 25 and 26-27 June 2004.

38. The SLD’s Izabela Jaruga-Nowacka was named deputy prime minister for social policy; she consistently resisted cuts in social spending.

39. Farm size was to be expressed in “standardised” hectares, which allowed for adjustment to take account of such factors as the quality of the land. This was intended to strengthen the link between (calculated) farm size and profitability.


41. The Council of Farmers was reduced from 50 to 25 members and the remuneration of its members cut, a move that generated some resentment among farming organisations. In addition, the requirement that the prime minister consult the Council before naming the president of KRUS was dropped.

42. People who combine farming and non-farm income may now remain in the KRUS, subject to some limits on non-farm income, if they were already qualifying farmers for three years.

43. As noted above, the target shrank by PLN 2.9 billion as a result of the change of government in May. In July, parliament eliminated another PLN 3 billion or so in projected savings and in November its rejection of a government bill on social security contributions for urban entrepreneurs and its amendment of other measures removed a further PLN 7.2 billion or so from the package. The KRUS bill was withdrawn in December.

44. See OECD (2006:77).
**Bibliography**


Chapter 8

United States: The 2005 Social Security reform debate

In February 2005, the Bush administration presented a Social Security (pension) reform proposal that would have allowed workers to divert a portion of their payroll taxes to personal savings accounts. The principal and interest in such accounts would in due course be disbursed to retirees and would supplement traditional Social Security benefits. A retiree’s traditional benefit would be reduced by the annuitised value of the diverted payroll taxes accrued at a real interest rate of 3%.1 The proposal met fierce opposition as soon as it appeared. Congressional Democrats were more or less united in opposition, and many Republicans were uncomfortable with aspects of the proposal. Moreover, senior citizens’ lobbies mobilised against it on a large scale. Resistance strengthened as it became apparent that public support for the proposal was falling. By mid-2005, it was clear that there would be no significant Social Security reform before the 2006 mid-term elections, and the Republicans’ loss of their congressional majorities at those elections effectively eliminated any prospect of legislative action on the administration’s proposal.

The Bush administration’s inability to advance its reform agenda reflected a number of factors:

- **Lack of popular acceptance of the urgency of reform.** While there was widespread awareness of the long-term threat to the solvency of Social Security, there was considerable controversy about just how urgent the system’s problems were or how drastic the measures needed to shore it up might be. Yet a sense of urgency about reform was critical to winning support for any transition to partial pre-funding of Social Security, as such a move would involve significant costs in the near term but offered benefits only over the long run.

- **Divisions among proponents of reform.** There were deep splits among Republican supporters of reform over the need for payroll tax increases or benefit cuts. Partly for this reason, the administration avoided presenting a specific reform plan – initially, the White House focused only on the proposals for personal accounts, which almost all Republicans could support, while leaving it to congressional negotiators to choose what approach to solvency to pursue. The administration hoped that congressional Republicans would move on the issue once the president and his team had built up a base of public support for reform. However, there were few incentives for Republican legislators to risk taking the lead on such a contentious issue.

- **Poor communication of the case for reform.** Divisions among advocates of reform also hindered the administration’s communications efforts. While the solvency problem was most often cited in making the case for the urgency of reform, the White House initially avoided committing itself to specific solvency proposals. When it later began to put forward proposals on solvency in response to Democratic attacks, its opponents were able to convey the impression that the
administration was unveiling the painful bits of its reform agenda only reluctantly and under pressure.

- **Strong resistance from vested interests.** The administration’s proposal triggered a large-scale mobilisation by lobbies opposed to the reform, led by the labour unions and the American Association of Retired Persons (AARP). While attempts were made to mobilise supporters of reform in the business community, these came to little. This asymmetry in lobby mobilisation was mirrored in the electorate as a whole: polls suggest that the social groups most likely to approve of the proposal were those for whom the issue was least salient, while those to whom it mattered most tended to oppose it. There was no clearly defined constituency that was supportive of the reform and for which it was a critical issue.

- **Failure to achieve a bipartisan approach.** For decades, polls have consistently shown that the Democrats enjoy a 15-20 point advantage in terms of public trust when it comes to dealing with Social Security. This meant that, when the Social Security debate heated up, the public was likely to be more receptive to Democratic messages than the administration’s. Survey data suggest that this was indeed the case. However, polls also suggested that voters feared that the Democrats would do too little to shore up Social Security, even as they feared that the Republicans would go too far – a clear indication that the electorate wanted any reform to proceed on a bipartisan basis. Yet in the absence of a crisis that might have made opposition look like obstructionism, the Democrats had little incentive to back a reform that would antagonise many of their core supporters and bring no real benefits in the near term.

**The context for reform**

**Framework conditions for reform**

The administration’s Social Security reform proposal was launched shortly after the 2004 general election, which had left the Republicans in control of the White House and with enlarged majorities in both chambers of Congress. The reform drive was thus launched as far ahead of the next national elections as possible; however, these were still only about 21 months away, so there was little opportunity to legislate in relative freedom from electoral pressures. Moreover, President George W. Bush had won with the smallest margin of any re-elected president in over a century, and he began his second term with a public approval rating of 50%, the lowest of any newly re-elected president since polling began. There was therefore no “political honeymoon” in which to pursue Social Security reform at the start of the presidential term.

The recovery from the 2001 recession had gained considerable momentum in 2003-04, and growth in early 2005, though moderating somewhat, remained strong, at an annual rate of close to 4%. Nevertheless, polls continued to show high levels of economic discontent and anxiety. Politically, voters’ subjective feelings may have mattered more than actual performance, since the economic impact of any reform would not really be felt for decades. As households grew more insecure and risk-averse, the attachment to Social Security’s guaranteed benefits tended to deepen. The other economic development that complicated the politics of Social Security reform was the disappearance of the federal surplus and the rise in the deficit to around 3.5% of GDP. Previously, some advocates of reform had hoped to use the surplus to finance a structural...
reform of Social Security; with that option no longer available, any transition to partial pre-funding would increase the system’s visible deficit significantly unless accompanied by tax increases or benefit cuts. That said, the vast majority of reform proposals emphasised the need to maintain separate financing of Social Security: under most proposals, general revenue transfers would be temporary and would be paid back by surpluses in other years.

Demographic pressures were among the factors putting Social Security reform on the political agenda: the ratio of contributors to retirees in the Social Security system in 2000 stood at around 3.4:1, down from 5:1 in 1960, and it was projected to fall to close to 2:1 by the mid-2030s. Moreover, retirees were living longer than ever – the average length of retirement in the United States had increased by 17 years during the 20th century. Nevertheless, the United States in the early 2000s was, by European standards, still a relatively “young” country, and it would be difficult to argue that the “greying” of the population was yet approaching the point at which radical reform was urgently required.

**Actors and interests**

For most of its history, the Social Security Administration (SSA) played the dominant role in Social Security policy making, particularly during the phase of Social Security’s expansion. Its influence was gradually eroded as a result of increased outside interest in Social Security – Social Security policy was much less politicised when the programme was expanding than it became in the era of retrenchment. Technological progress played a role here: until the 1980s, few outside the SSA could easily conduct the sort of large-scale analyses that it did, but the debate in recent decades has increasingly been shaped by outside analyses, not least those coming from lobbying organisations and influential conservative think-tanks and interest groups. Several of these became increasingly vocal advocates of private accounts during the 1980s.

The partisan politics of Social Security reform were characterised by a degree of asymmetry. While the administration and many congressional Republicans strongly supported personal accounts, a substantial minority of Republicans in Congress were uncomfortable with the idea. Moreover, advocates of reform in the Republican caucus were divided over whether their introduction would need to be accompanied by benefit cuts or tax increases in order to shore up the solvency of the system. Thus, while administration officials had been working on a range of proposals for addressing solvency concerns along with the introduction of private accounts, the need to hold the caucus together meant that many of these ideas could not be pushed too hard. Congressional Democrats, by contrast, were unified in their opposition to what they described as the “partial privatisation” of Social Security. Traditionally, most Democrats have seen any move away from universal public provision as a threat to the long-term future of the programme, and the vast majority of Democratic legislators rejected any sort of personal accounts financed by diverting contributions from the Social Security system itself – that is, by a so-called “carve-out”. Moreover, most of the relatively few congressional Democrats who were open to “carve-outs” either retired or failed to win re-election during 2000-04, leaving the congressional party even more cohesive on this issue by 2005 than it had been in the 1990s.

The administration’s Social Security proposals triggered a massive mobilisation of interest groups on both sides of the debate. Supporters of personal accounts, like the National Association of Manufacturers, the US Chamber of Commerce, the Club for Growth and others banded together in the Alliance for Worker Retirement Security.
(AWRS) and the Coalition for the Modernization and Protection of America’s Social Security (COMPASS). However, their public relations and lobbying efforts were more than matched by those of their opponents, led by America’s largest trade-union federation, the AFL-CIO, and the American Association of Retired Persons (AARP). While there were some pro-reform senior citizens’ groups, the AARP, with its 36 million members and 5,000 staffers, largely succeeded in establishing itself as the voice of older Americans in the Social Security debate. In 2005, polls showed that ordinary voters trusted AARP statements more than those of any other actor.

Issue history

Social Security policy making is characterised by a strong element of path-dependence. Its history from 1935 to the early 1970s was of an uninterrupted series of decisions to increase the reach and generosity of the programme, culminating in the 1972 decision to index Social Security benefits automatically to the CPI, rather than leaving it to Congress to adjust benefits periodically to inflation. There were subsequent moves towards retrenchment in 1977 and 1983, each time in response to what appeared to be imminent crisis. These reforms involved a combination of benefit cuts and tax increases; the 1983 legislation also initiated a phased increase in the normal retirement age from 65 to 67. Both packages were adopted with substantial bipartisan support, so the blame for politically painful retrenchment was shared.

Arguments for a structural reform of Social Security had been a feature of academic discussions since the 1960s, but the question took on political significance only after the funding squeezes of the late 1970s and early 1980s focused attention on the issue of financial sustainability. The next two decades witnessed a lively, though largely inconclusive, debate about the long-term future of the programme, involving proposals for everything from relatively limited tinkering with system parameters to full privatisation. However, attention was increasingly focused on two related but distinct issues: advance funding and a (full or partial) move towards privately or publicly administered personal accounts. By the early 2000s, there was broad, though by no means universal, support for moving towards partial pre-funding of Social Security, an idea unanimously endorsed by both the Clinton administration’s Social Security Advisory Council and the Bush administration’s Commission to Strengthen Social Security.

The question of personal accounts, by contrast, remained highly controversial, though support for them was growing, driven largely by a number of Washington-based think-tanks that had taken up the idea in the 1980s. Inspired by Chile’s example, they sought, unsuccessfully, to make private accounts a part of the Reagan administration’s reform agenda. While that administration never took any such action, the network of private-accounts advocates continued to grow, and it did win the support of much of the business lobby and a significant part of the Republican Party. The growth of the Trust Fund also seems to have strengthened this lobby: some economists saw diverting resources away from the Trust Fund as a means of restraining federal spending; business lobbies sought a way to hold down future payroll tax increases; and many in the financial sector were interested in the commercial opportunities offered by the emergence of private accounts. Thus, by the 2000 election, many Republicans, including the party’s main presidential contenders, openly supported the diversion of some Social Security contributions into private accounts.
Box 8.1. The Social Security system

Social Security has operated since 1939 a pay-as-you-go (PAYG), self-financing system funded largely through payroll taxes paid by beneficiaries and their employers. It is a near-universal entitlement programme, covering nearly 90% of Americans and providing the main source of income for two-thirds of the elderly; for over 30% of them, it is their sole source of income. As benefits have grown and the worker-beneficiary ratio has fallen, the combined tax rate has risen from 2% of taxable wages to 12.4% today. The system uses a progressive benefit formula, so replacement rates decline as average indexed monthly earnings rise, and a related but separate programme, Supplemental Security Income, provides benefits to low-income people who are over 65 or disabled. This helps to ensure that Social Security is relatively successful at achieving its principal objective – reducing old-age poverty – but it also weakens the link between contributions and benefits. There has long been a tension between the system’s anti-poverty objective and the need to ensure sufficient recognition of the differences in payroll taxes paid by different workers, which is seen as critical to sustaining broad public support for Social Security.

Since the 1980s, Social Security has run rising cash-flow surpluses. Current surpluses are counted as part of a Social Security Trust Fund, but the real meaning of this accounting is hotly contested, since they also enter into the consolidated federal budget and thus help reduce reported deficits. Moreover, the Trust Fund is invested exclusively in US Treasury bonds. If Social Security surpluses are used to reduce net federal debt, then they increase the federal government’s ability to pay future benefits. However, their real impact on government saving depends on their impact on the rest of the federal budget. A significant body of empirical work suggests that when Social Security is in surplus, federal spending rises – whether on Social Security (raising benefits or liberalising eligibility) or on non-Social Security expenditure – or taxes are cut. If Social Security surpluses are consumed, then they represent a political claim on future general revenues – a claim that will impose a growing drain on the federal budget at a time when other age-related spending is also expected to rise. Since there are no legally protected “acquired rights” under Social Security, some observers argue that the political feasibility of making good on those claims in full cannot be taken for granted (Cogan and Mitchell, 2003:154; OECD, 2005:45).

The system’s cash-flow surpluses are expected to turn into deficits by 2020. With the worker/beneficiary ratio falling rapidly, as the baby boomers retire, the gap between outlays and revenues is projected to rise to around 1.3% of GDP by 2035 or so under current law and to continue increasing more gradually for the rest of the century, to around 2% of GDP. As a result, the Trust Fund, which is expected to peak at around 25% of GDP in 2018 or so, is projected to be exhausted by mid-century (OECD, 2005). In 2005, the system’s unfunded obligations were estimated at USD 12 800 billion – roughly equal to 2005 GDP (Gokhale, 2005; Treasury, 2007).

1. Social Security was in fact founded as a pre-funded system, but it was transitioned to a PAYG basis just four years after its creation, in order to bring forward the date at which the first benefits might be paid.
2. There are in fact debates about the actual amount of redistribution achieved through Social Security, since wealthier retirees tend to live (and thus to draw benefits) longer and because rules governing things like survivors’ benefits tend to have a differential impact on different income groups. In any case, though progressive at lower levels of income, the system is regressive for income levels that exceed the maximum taxable earnings threshold (USD 90 000 in 2005). See Treasury (2007).
3. See, for examples, Munnell and Blais (1984); Aaron et al. (1988); Crain and Marlow (1990); Weaver (1990); Cogan (1998); Schieber and Shoven (1999); Smetters (2004); Gokhale (2005); and Becker (2005). For the argument that Trust Fund surpluses did increase national saving, see Diamond and Orszag (2004).
The reform debate: advance funding and personal accounts

Advocates of pre-funding of Social Security have long argued that it could increase national saving, thereby increasing the nation’s capital stock and productive capacity and reducing Social Security’s financial burden on future generations. Partial pre-funding has also been seen as a way to diversify risk. The worker’s rate of return on contributions to the PAYG system depends on labour-market and productivity performance, while the return on contributions to any new, funded scheme would depend on rates of return on investment portfolios. However, the impact on saving of a transition to partial pre-funding would depend on how it was financed: pension reform could only add to national saving if it involved some short-term sacrifice in public or private consumption. Prior to 2001, many advocates of reform argued for using the federal surplus to finance the move to partial pre-funding, since the surplus would otherwise be consumed, via tax cuts or higher spending, and the Social Security transition would have to be financed by increased borrowing and no increase in national savings would result.

While pre-funding would not necessarily require a shift towards personal accounts and a defined-contribution system, some experts see personal accounts as a logical concomitant to any move to pre-funding, since they would allow decentralised risk management: allowing individuals chose how their contributions were invested was seen as preferable to having one overall board allocating the hundreds of billions of dollars of common stock. Supporters argue that such accounts could be designed so as to shore up the solvency of the system as a whole, because they are likely to earn a better return on payroll tax contributions than the existing system. Moreover, creating property rights is seen as a way to make increased saving and reduced consumption in the near term more palatable to the electorate; it would make the link between contributions and benefits both stronger and more transparent to individuals, thereby creating incentives for workers to save more and retire later. Advocates of personal accounts also argue that “carve-outs” would help to check Congress’s tendency to spend the Social Security surplus – which increases the risk to long-term solvency – and that personal accounts, by creating private property rights, would protect contributing workers from subsequent ad hoc changes to benefits. The separation of retirement savings from interest-group politics is regarded by some as the strongest argument for carve-outs.

Against these arguments, opponents of personal accounts point to the high administrative charges they are likely to involve, citing what they see as the disappointing experiences of other countries with personal accounts. This raises questions of equity as well as efficiency, since administrative charges relative to assets would be considerably higher for small accounts than for large ones, creating potentially large differences in the rates of return to high and low earners. Critics of personal accounts (whether privately or publicly administered) also worry about the impact on the level of redistribution within the system of restructuring it into a PAYG/defined-benefit component and a funded/defined-contribution scheme. The latter, by definition, would not be redistributive. While it would be possible in principle to introduce a defined-contribution element without materially altering income redistribution, doing so in practice would put at risk some of the purported advantages of personal accounts. Finally, to the extent that the basic source of legislative risk to benefits stems from the overall financial position of Social Security, critics argue that the political risk to benefits would be unchanged – it would simply be spread over a smaller base of legislated benefits.
The evolution of public opinion

Social Security is one of the most popular institutions in America: polls for decades have shown broad support for the programme, with little variation by age group, and few surveys since the early 1980s have shown more than a fifth of respondents prepared to advocate any reduction in Social Security transfers, despite concerns about its financing and future prospects. Polling data for 2004-05 suggest that the on-going debate did little to change this: the public was concerned about Social Security and wanted action but did not believe that the system was “in crisis” or that there was any danger of an imminent collapse.31 Health and education were seen as far more urgent priorities than Social Security reform.32

Policy process: formulation

The search for solutions

In May 2001, the Bush administration created a commission charged with producing reform options by December. The commissioners were asked to make recommendations that would restore the financial soundness of Social Security without raising payroll taxes or changing benefits for current retirees or near-retirees. The commission’s mandate also included the requirement that its proposals provide for individually controlled, voluntary personal accounts to augment Social Security, be these “carve-outs” or “add-ons”, financed without any diversion of payroll taxes.33 Previous Social Security commissions, as well as the quadrennial advisory councils that had reported since the 1970s, had served to help the executive to negotiate with Congress, to avoid the blame for unavoidable but unpopular measures, and to engage non-governmental actors directly in the policy process. The new commission, though bipartisan in make-up, fulfilled none of these functions. Its aim was to help the administration build support for a particular approach to reform, and its staff was led by strong advocates of personal accounts.34 While the commission heard and discussed testimony from a wide range of viewpoints, including opponents of personal accounts, its mandate made it easy for critics to challenge its objectivity. The commission reported in December 2001 with three options, all involving some form of personal accounts financed by a carve-out of some portion of Social Security payroll tax and transfers from general revenues to make up the shortfall in Social Security funds.35 Two of the options also involved cutbacks in traditional Social Security benefits to future retirees (especially younger cohorts).36

The commission proposals did not lead to action. The benefit cuts envisaged were substantial, and it was not clear where the needed “general revenues” would come from, since the budget surplus was fast disappearing as a result of tax cuts and the surge in security-related spending that followed the terrorist attacks of 11 September 2001. This prompted fears that implementation of the proposals might ultimately require even deeper cuts in traditional benefits than initially envisaged. The disappearance of the surplus also made it more difficult to explain how the transition to partial pre-funding could increase national saving, since it would now have to be financed by borrowing.37 While supporters of personal accounts insisted that the state of the federal budget was irrelevant to the reform of Social Security, which is self-financing anyway, many in Congress thought otherwise, fearing any reform that led to significant larger visible deficits in the short term.38 Moreover, the administration’s focus shifted from Social Security to Afghanistan and Iraq, and, with the 2002 mid-term elections approaching, congressional Republicans were unwilling to take any of the proposals forward. Also, the Enron and WorldCom scandals reduced the appeal of proposals to invest a portion of payroll taxes in stock

markets. The administration nevertheless reinforced its commitment to personal accounts during the election campaign that followed, suffering no apparent electoral backlash for so doing. However, those elections were dominated by the issue of Iraq, so the outcome could not easily be treated as a mandate for Social Security reform. With only narrow majorities in both houses of Congress and the 2004 presidential race fast approaching, the White House did not attempt to legislate in 2003-04. The issue became prominent again only towards the end of the 2004 presidential election campaign, when it became clear that if re-elected, Mr. Bush aimed to make Social Security reform a central element of his second-term domestic agenda.

**The administration proposals**

Immediately after the election, the president signalled his determination to proceed with a Social Security reform involving personal accounts at the start of his new term. The administration set to work re-activating its network of donors and activists in support of Social Security reform, and in February 2005, it advanced a proposal to allow workers under 55 to divert four percentage points of their payroll taxes to a personal savings account. Diversion of contributions would initially be capped at USD 1 000 per year, eventually rising to 4% of total income subject to Social Security tax. The principal and interest in those accounts would subsequently be disbursed over time, as an addition to traditional Social Security benefits, to workers at retirement (or to their heirs if they died before retirement). The traditional benefit would then be reduced by the annuitised value of a notional account, equivalent to the diverted payroll taxes accrued at a real interest rate of 3% per annum. The administration continued to oppose raising payroll tax rates, citing the associated negative effects on incentives to work, but it left the door open to raising the ceiling on earnings subject to tax, as well as to the introduction of measures to discourage early retirement.

The administration did not endorse any of the 2001 Commission proposals, nor did it present a specific plan of its own to Congress, opting instead to set out some basic guidelines for personal accounts in the hope that congressional Republicans would move on the issue once the White House had built up a base of public support for reform. This was intended to leave the Republican congressional leadership a free hand in negotiating the specifics of any reform. This approach had served the administration well when adopting Medicare prescription drug benefits, but there were few incentives for Republican legislators to risk taking the lead on such a contentious issue as Social Security, which was likely to involve retrenchment rather than (as had been the case with Medicare) the extension of benefits. Moreover, there were deep divisions over the specifics of any carve-out, even among those in Congress who favoured personal accounts. The major cleavage was between those who believed that the transition to a partially pre-funded system could be financed without raising taxes or cutting benefits and those who believed that any credible reform must involve some combination of tax rises and benefit cuts. It was therefore bound to be difficult for the congressional majority to coalesce around a single set of proposals without the administration’s leadership.

Some of the administration’s chief allies in Congress tried unsuccessfully to persuade it to shift the focus to solvency rather than personal accounts, but higher taxes, lower benefits and an increase in the retirement age did not figure in the White House proposal: while declaring that “everything was on the table”, the administration actually committed itself only on the issue of personal accounts. Even the generally popular idea of raising the cap on income subject to payroll taxes was not initially taken up. In fact, the SSA and White House officials had done considerable work on solvency and were strongly
inclined to favour “progressive indexation” (see below), but the administration, fearful of dividing congressional Republicans at the start of its reform drive, initially took no public stand on the issue beyond indicating a range of options that it might consider. It insisted that Congress should be free to negotiate on solvency measures.

The other big problem the administration faced was the absence of any real consensus about how urgent the Social Security system’s problems were. Previous reforms had been adopted only when the prospect of an imminent crisis had forced Congress and the White House to act, but the Trustees’ estimates suggested that the Social Security Trust Fund would not be exhausted until 2041 (2052, according to the Congressional Budget Office’s more sanguine projections). Critics argued that the assumptions underlying the Trustees’ annual reports were exceptionally pessimistic and noted that, year after year, the reports’ shortfall predictions were improving.\textsuperscript{43} Indeed, the Trust Fund was reaching record levels even as the administration was advancing its proposals, a fact that made it hard to argue for the urgency of reform. Moreover, to the extent that solvency was the problem, the White House’s initial failure to put forward proposals to tackle the solvency issue exposed it to much criticism. Even the administration acknowledged that personal accounts did not address solvency directly, though it argued that they would make the changes needed to strengthen solvency more attractive to individual workers.\textsuperscript{44} At best, personal accounts would accompany other reforms aimed at shoring up the system’s finances; at worst, they might exacerbate the system’s solvency problems.

Policy process: adoption

“Going public” with Social Security reform

Fresh from a successful re-election campaign, the White House quickly opted for a strategy of “going public” in order to win adoption of its Social Security reform. Rather than focusing on direct dealings with Congress, where it found little readiness among Democrats to co-operate, the administration launched a carefully orchestrated 60-day, 60-stop public campaign for Social Security reform involving numerous speeches, interviews and articles by the president himself and senior administration officials.\textsuperscript{45} A great deal of attention was focused on states and districts that Mr. Bush had carried in 2004 but that had elected Democratic senators or congressmen. The aim was to generate such strong public support for reform that moderate Democrats could be won over and wavering Republicans kept on board. This effort was not successful. Polls during the 60-day campaign showed steadily falling public support for the proposal.\textsuperscript{46} By end-May, mounting public opposition to personal accounts, together with the growing divisions among congressional Republicans over the details of the reform, prompted leading Republican legislators to acknowledge that there would be no Social Security reform in 2005 and that action was unlikely in 2006, an election year.\textsuperscript{47} The administration continued to work on the issue actively through the summer, while Senate Republicans explored the scope for a possible compromise with their Democratic colleagues, but by October, even the White House had to acknowledge that its efforts to sell the reform had been stymied.\textsuperscript{48} The Republicans’ loss of their congressional majorities the next year ensured that there would be no legislation on personal accounts during the Bush administration.

A number of factors contributed to this outcome. First, congressional Democrats remained united in opposition to the reform proposals, thus depriving the administration of any bipartisan political cover. Democratic discipline in the Senate was particularly crucial, because under Senate procedures, any reform plan would need 60 votes to avoid a
filibuster, and the Republicans held only 55. In any case, Senate Republican leaders wanted 65-70 votes in favour, as evidence that the reform really was bipartisan. The Democrats’ unity reflected both the changes in the Democratic congressional caucus described above and the deep partisan divisions generated by the bruising 2004 election campaign. It was also facilitated by the Democrats’ refusal to produce any alternative reform proposal. Reaching agreement on a Democratic reform plan would have been extremely difficult, and any plan that emerged would have given the Republicans a target. Since the White House had not presented a fully fledged plan of its own and the finances of Social Security were not such as to require urgent action, the Democratic congressional leadership did not need to offer its own proposals; it could simply oppose what the White House put forward. The Democrats maintained this united front even after the White House offered “progressive indexing” in an effort to win some Democratic support with a solvency proposal that would protect the least well off. In essence, progressive indexing would have made the initial calculation of benefits less generous for more affluent workers while maintaining current-law benefits for those with more modest means: replacement rates would thus fall for those further up the income scale.

Secondly, Republican supporters of Social Security reform could not overcome their divisions on the question of solvency. Indeed, it was conflict over this issue that had prevented the White House from presenting a more comprehensive set of proposals in the first place. When, at the end of April, the president indicated his readiness to consider lifting the cap on earnings subject to Social Security contributions and endorsed progressive indexing as part of the solution to the solvency problem, many Republican supporters of personal accounts were sharply – and publicly – critical. Others worried about the impact of the reform on the budget deficit. Supporters of reform argued that it would not actually increase the debt – it would merely turn implicit liabilities into explicit ones – but this proved a hard sell politically, so the transition costs associated with the proposals remained a political liability. However, most congressional Republicans, though open to the administration’s ideas, were not strongly committed one way or another. As it became apparent that the proposal was unpopular, these waverers declined to support it. This combination of Republican divisions and Democratic discipline was precisely the opposite of the pattern seen early in the president’s first term, when congressional Republicans had been very disciplined and the White House had often been able to win some Democratic support for its major initiatives.

Thirdly, the administration proposal triggered a large-scale mobilisation by lobbies opposed to the reform, led by the AFL-CIO and the AARP. While lobby groups on both sides organised letter-writing campaigns, ran television ads, and lobbied Congressmen directly, opponents of reform proved far more active and substantially outspent their pro-reform rivals. As soon as the administration signalled its intention to act just after the election, the AARP and its allies mobilised. Ad vertisements attacking personal accounts began to appear even before the administration unveiled its proposals. The administration’s opponents were quick to accuse it of trying to “privatise” Social Security, although the meaning of this term was rarely specified. Democratic leaders worked with organised labour to coordinate the activities of around 200 organisations under the umbrella organisation Americans United to Save Social Security. The highest-spending single lobby in the first half of 2005 was the AARP; the National Committee to Preserve Social Security and Medicare, which also opposed the administration’s proposals, ranked third.

All three of these trends were accentuated by polling data showing that the administration’s communications strategy was failing. As support for the reform proposal
fell, congressional Democrats found it easier, and Republicans harder, to maintain a united front. Congressional Republicans were largely bystanders during the 60-day campaign, with few ready to risk getting ahead of public opinion on the issue. The same was true of the contest among lobbies: business groups were reluctant to put resources behind an initiative that seemed to be floundering, and most of them tended in any case to focus their lobbying efforts on other aspects of retirement reform, particularly legislation pertaining to private pension savings.\textsuperscript{58} While poll trends and the lobbies’ efforts had little impact on committed supporters of personal accounts, wavering Republicans came under intense pressure from the AARP and its allies; faced with growing negative feedback from the public, some began to distance themselves publicly from the proposal.\textsuperscript{59}

**The administration’s communications strategy**

The strong anti-reform trend in public opinion during the reform debate partly reflected the difficulties the administration encountered in trying to explain and promote the reform. The first problem was the administration’s conviction that the 2004 election had given the president a mandate for Social Security reform. The president had indeed highlighted the need for Social Security reform during both presidential campaigns, but only in general terms – no specific measures were advanced during the 2004 campaign. Post-election polls suggested that the electorate did not believe it had given the president a mandate to reform Social Security.\textsuperscript{60} Moreover, support for personal accounts in early 2005 was at or below 50\% in most polls, even when the question was posed without reference to offsetting cuts in traditional benefits. Most respondents did not initially realise that personal accounts would entail lower traditional benefits, and support for them fell sharply when questions linked their introduction to a reduction in the guaranteed benefit retirees would receive from Social Security.\textsuperscript{61} Support for personal accounts was also weaker in polls that mentioned investing diverted payroll contributions in the financial markets, suggesting that many respondents regarded such investment as too risky. Public support for reform fell further as the White House, under pressure from its critics, began to address the question of solvency more explicitly, as opponents were then able to suggest that the administration had tried to hide the real costs of its plans, focusing attention on personal accounts while only reluctantly revealing its tax and benefit proposals.\textsuperscript{62}

A further problem for the administration was that support for personal accounts was highest among younger and wealthier voters, for whom the future of Social Security was a less salient issue: for younger voters, retirement was still too far off to be the most pressing priority, while better-off older voters tended to have substantial private pension savings anyway, so the prospect of being allowed to create small personal accounts financed by a portion of their payroll taxes meant little to them. By contrast, the people most engaged in the debate, who were generally older and less well off, were also those most opposed to the plan.\textsuperscript{63} Younger workers on relatively low incomes were less likely than high earners to have private pension plans or other private retirement savings, and the administration proposal would have offered them an opportunity to diversify their retirement portfolios if they wished and to secure higher returns on their contributions than the status quo offered them. This was the group the President’s Commission had identified in 2001 as the segment of society with the most to gain from personal accounts.\textsuperscript{64} However, polls suggest that this was the group least engaged in the debate. Moreover, although the administration spoke of the reform as part of its efforts to create an “ownership society”, less well off workers seemed to place greater emphasis on
security than on higher returns and were not attracted to the idea of being able to invest some Social Security contributions in the stock market.

In any case, the president’s *ex ante* approval ratings may simply have been too low for such a strategy to be viable on an issue as hotly contested as Social Security. Recent research suggests that “going public” is a risky strategy for a president and that its success depends on already solid approval ratings. However, the deeper explanation appears to be structural rather than personal or conjunctural. Polls for decades have shown that the American public trusts Democrats more than Republicans to handle Social Security, usually by large margins. This meant that voters in 2005 were more receptive to Democratic arguments than Republican ones. In essence, the administration was effectively focusing the domestic political debate on an issue “owned” by its opponents, so it is hardly surprising that the more the reform was discussed, the lower support for the White House’s position fell. Given this public-opinion backdrop, any Republican president at any time would probably face an uphill battle when selling a Social Security reform that was not clearly bipartisan. Indeed, the electorate itself seems to have wanted a bipartisan solution. Polls found large majorities afraid that the Republicans would go too far in reforming Social Security. However, roughly equal proportions feared that the Democrats would not go far enough.

The administration also ran into problems generating a sense of urgency about reform. While aware that Social Security would have serious problems in the future, most voters were not persuaded that immediate action was required. Since 2041, the date at which the Trust Fund would be exhausted, was still so far off, reform advocates initially focused on the period around 2016-18, when the system’s cash-flow surpluses were projected to turn into deficits. While 2016 certainly sounded more urgent than 2041, the administration’s opponents denied that there was any particular economic significance to the disappearance of cash-flow surpluses, and critics of personal accounts were quick to point out that any sort of “carve-out” would merely bring forward the date at which the system’s cash-flow surpluses disappeared. More generally, proponents of reform emphasised the costs of delay, arguing that if changes were adopted only as the system approached a crisis, the tax hikes or benefit cuts needed to save it would be all the greater and those costs would be shifted overwhelmingly onto younger cohorts.

Finally, the White House also had great difficulty in explaining its progressive indexation proposal. This would have eliminated as much as 70% of the long-term deficit, but it would have cut benefits for the upper 70% of future retirees, with the wealthiest losing the most. This attracted fierce criticism from those on the right who objected to making the system more redistributive, as well as from many on the left, who saw it as the first step towards turning Social Security into a “welfare” programme, a development that would undermine long-run popular support for it. While progressive indexation was popular in principle, benefit reductions would have begun to affect people with average lifetime earnings of USD 25 000 a year or more, well below the level that most poll respondents regarded as “wealthy”. Of course, the figures referred to average earnings over a lifetime, not to final earnings, but opponents of the reform were successful in blurring the distinction between final and lifetime average earnings.

**Conclusion**

Any move towards partial pre-funding of Social Security was bound to be difficult, as such a transition would involve significant costs in the near term, while delivering no benefits for which politicians could claim credit except over the long run. The
administration clearly hoped that linking partial pre-funding to the creation of personal accounts would make the transition more attractive. However, the conditions in which the reform proposal was advanced were in many respects decidedly unfavourable. There was no strong consensus concerning either the need for reform or the solution to be adopted. While it was widely agreed that it would be less costly to reform the system before the situation was critical, legislators had few incentives to risk taking action on Social Security reform in the absence of a crisis or a very broad base of support for a reform package. Moreover, advocates of reform never fully resolved the apparent contradiction between the focus on solvency as the urgent problem and the decision to make personal accounts the centre-piece of the reform. Nor was there a clearly identifiable constituency for personal accounts. The groups which might have had the most to gain from reform – younger workers and those on lower incomes, to whom the proposal offered at least a limited opportunity to diversify their retirement savings – were either not much engaged in the debate or did not find the proposed solution attractive.

Although they controlled both Congress and the White House, the Republicans needed significant Democratic support for both procedural and political reasons: procedurally, they could not move anything through the Senate without some Democratic support, and politically, they needed any reform to be bipartisan, in order to limit the electoral risks to themselves. Congressional Democrats were thus in a position to act as veto players, as long as they remained united. Previous reforms have always been bipartisan, regardless of the administration in power, because neither party has wished to shoulder the blame for retrenchment alone. However, bipartisanship is even more important for the Republicans than the Democrats when it comes to Social Security, given the persistent asymmetry in levels of public trust on the issue. Yet given single-party control of the federal government, there is little reason for the opposition to join a compromise if loss-allocation is involved and there is no crisis so urgent as to make non-cooperation look like dangerous obstructionism. Given the difficulty of the issue and the weakness of American political parties, a suitable reform package is unlikely to emerge from “normal” congressional committee work, particularly if there is no crisis to force the issue. Yet even a special mechanism, like the President’s Commission, is unlikely to get the job done if its bipartisan status is in doubt.

The contrast between Social Security and welfare reform suggests that one of the factors that may matter is the “ripeness” of a policy regime for reform. Unlike the welfare programme Aid to Families with Dependent Children (AFDC) in the mid-1990s, Social Security in 2005 was not obviously ripe for reform. In the case of AFDC, radical reform was preceded by a shared conception that the programme was failing beyond remediation and a series of incremental decisions had already begun to weaken the policy regime. In 2005, there was no agreement that Social Security was in any sense “failing”. It remained very popular and it was, on the administration’s own estimates, still several decades away from acute financial problems. Nor was the administration proposal preceded by the kind of incremental changes that had already begun to affect AFDC. Indeed, perhaps the biggest advantage that opponents of reform enjoyed was that the alternative to reform was a popular and well institutionalised status quo that appeared to be in no danger of imminent breakdown if reforms were not implemented.
Notes

1. Composed of a 2.7% benefit offset plus 0.3% in administrative costs.
2. Pew (2005); see also Galston (2007:3).
4. Especially since headline federal deficits would have been far larger in the absence of Social Security surpluses.
5. See CSS (2001): two of the three options proposed by the President’s Commission in 2001 provided for repayment of general revenue transfers. Option 1 did not and was criticised in part on these grounds.
6. The decline would slow thereafter, but the ratio would hit around 2:1 by mid-century. CSS (2001); OECD (2005).
8. Nevertheless, the SSA has generally been the most important source of information and analysis to the various official commissions and advisory councils that have looked at Social Security. See Derthick and Teles (2003) for details.
9. These included the Cato Institute, the Heritage Foundation, the American Enterprise Institute and many others.
10. These included Calvin Dooley, Charles Stenholm, Charles Robb, Robert Kerrey, Daniel Moynihan and John Breaux.
11. The AWRS focused on lobbying Congress, while its sister organisation, COMPASS, which was backed by many of the same interest groups, targeted pro-reform messages at the general public.
12. The largest lobby organisation in America, the AARP maintains substantial policy research and lobbying divisions and is also effective in mobilising its grassroots members to contact elected officials about Social Security.
14. Ross (2007). In this, it is typical of pension regimes in most countries. A large literature on the political economy of pension reform emphasises path dependence: see, e.g., Pierson (1994); James (1998); Myles and Quadagno (1997); and Myles and Pierson (2001).
15. There were 25 such adjustments in the programme’s first 20 years, all of which tended to relax eligibility or increase benefits.
16. The 1977 reforms were adopted in response to actuarial projections showing a shortfall in financing as early as 1979, while the 1983 reform was prompted by forecasts suggesting that such a shortfall might be only months away. The 1977 package reduced benefits even for cohorts nearing retirement. See Arnold (1998).

17. See Diamond (1996) for an overview of proposals then in circulation.

18. The party’s eventual nominee, George W. Bush, had advocated allowing individuals to invest a portion of their payroll taxes in private accounts as long ago as his first race for Congress in 1978 and had raised the issue when first announcing his candidacy for president; Galston (2007:1).


20. Cf Samuelson (1958): in a stable PAYG system, the equilibrium rate of return on worker contributions equals the growth of the economy’s real wage bill – i.e. real wages times the number of workers.

21. For a detailed discussion of this issue, see Bosworth and Burtless (2004).

22. As a candidate for president, the then Governor Bush proposed to use half the surplus for this purpose.


24. CSS (2001); Cogan and Mitchell (2003).


26. If accumulating Trust Fund surpluses were spent rather than used to reduce net federal debt, the federal government’s ability to pay benefits when the system’s cash-flow turned increasingly negative would be undermined.


29. See Diamond (1996) for an overview of the arguments against private accounts.

30. Critics argued that this was what happened in Chile during the first decade after privatisation. See Vittas and Iglesias (1992).

31. Bowman (2005); Pew (2005). A Washington Post/ABC News poll in December 2004 found that only 25% of respondents thought Social Security was “in crisis”, as against 49% who felt it had major problems but was not in crisis; in a January 2005 CNN/USA Today poll, 18% opted for crisis and 53% for “major problems.”

32. See the data in Pew (2005).

33. The commission’s mandate also required that any proposed reform preserve Social Security’s disability and survivors’ benefits, and it excluded using the Social Security surplus for anything other than Social Security or allowing the government to invest Social Security funds in the stock market.
34. These included Charles Blahous, the Special Assistant to the President for Economic Policy, and Andrew Biggs, a former Cato Institute researcher who was associate director of the National Economic Council in the Bush White House.

35. Two involved voluntary accounts financed from payroll taxes, with offsetting reductions in “traditional” PAYG Social Security benefits: one left the Social Security system otherwise unchanged, while the other also provided for “progressive indexation”. The third proposal was for voluntary “add-on” accounts, with matches from payroll taxes, overlaying the traditional Social Security system, albeit with some revenue and outlay changes. Two of the three options provided for general revenue transfers to be repaid in order to preserve the self-financing character of the Social security system overall.

36. In these commission proposals, some cutbacks would have affected even those who opted not to create personal accounts.

37. The underlying argument here concerns political economy. Strictly speaking, whether Social Security is, or should be, pre-funded is independent of whether or not the rest of the federal budget is in surplus or deficit. However, as noted above, the evidence concerning the relationship between Social Security balances and federal spending suggests that the use of general revenue transfers to finance the transition to partial pre-funding would, even if those transfers were to be repaid by the Social Security system, have led to slower growth in federal spending (or reduced the scope for tax cuts).

38. See Bliss (2005); and Gleckman (2005) on the unwillingness of conservative Democrats and some Republicans in Congress to back personal accounts in the context of rising federal deficits.

39. Current-law benefits would be preserved for workers 55 and older.

40. Several leading congressional Republicans urged the administration not to submit a bill of its own; Washington Post, 12 February 2005. Others, however, pressed the White House to present detailed proposals; Edwards (2007:253).


42. Bush (2005); “Strengthening” (2005).

43. See, e.g., Weller and Rasell (2000), who note that the 2000 report was based on projected rates of GDP, wage and productivity growth for 2010-75 that were lower than those recorded over any extended period in the previous half-century. However, the SSA pointed out that for much of the 1970s and 1980s, it had been criticised for excessive optimism and argued that the long-run rate of productivity growth would have to more than double to keep the system solvent through the 21st century without any adjustment to tax or benefit rules. The argument that the SSA has been too pessimistic rests largely on its failure to anticipate the surge in wage growth of the late 1990s rather than any intrinsic bias in its methods.

According to the White House, during 3 March-1 May 2005, the president, the treasury secretary and other senior administration officials took part in 166 events in 40 states as part of the campaign. They also gave more than 500 radio interviews.

Public approval of the president’s handling of Social Security fell by 16 percentage points between February and early June. See Galston (2007); also Ross (2007:426); and the polling data in Bowman (2005).

See Ross (2007:426) for comments from Republican leaders in the House and Senate.


US Senate rules permit a senator, or a series of senators, to speak for as long as they wish and on any topic they choose (a tactic known as a filibuster), unless a supermajority of three-fifths of the Senate votes to end debate. In practice, continuous floor speeches are no longer required, so filibusters have become much more common in recent years. Budget bills are governed under special “reconciliation” rules that do not allow filibusters. Since 1996 these rules have been used for all budgetary matters except Social Security, which cannot be subject to reconciliation.

Polls showed widespread distrust of the language of “privatisation”, so the administration and its supporters referred to “personal” rather than “private” accounts, while opponents spoke of “privatisation”. For details on the use of language at this stage, see Edwards (2007:279-80).

Only one of the 201 Democrats in the House of Representatives backed the administration proposals; the four Democrats in the Senate who had previously expressed some interest in personal accounts rejected Republican overtures to win them over.

Progressive indexation would have modified the basic benefit formula except for the lowest 30% of career-average earners; their basic benefits would have remained unchanged. Beginning in 2012, there would be reductions for workers with career earnings above this threshold, as the link between initial benefits and wages would be attenuated. The reductions would be smaller for workers at medium career earnings levels and larger for workers at higher career earnings levels, reaching the equivalent of full CPI indexing of the benefit for the very highest earners. Individuals becoming eligible for benefits before 2012 were not to be affected by the change.

Some congressional Republicans objected to the fact that the administration’s proposal included offsetting benefit reductions along with the diversion of some payroll taxes into personal accounts. They insisted that such offsets were unnecessary and would increase opposition to reform.

Galston (2007:5)

Galston (2007:3-4).

Reform advocates pointed out that the term “privatisation”, which proved rather inflammatory in the Social Security reform debate, could not easily be applied to publicly managed personal accounts; its suitability was not obvious even with private management of accounts, since these would still be funded from mandatory Social Security contributions.
Security contributions and would form part of a heavily regulated, overwhelmingly public system.

57. Data from the *Congressional Quarterly’s* “MoneyLine”. By early April, the AARP reported 460 000 calls to Congress from constituents complaining about the administration plan.


59. Edwards (2007:300-02) describes the pressure brought to bear on many Republican legislators by lobbyists and constituents, as well as their increasingly open qualms about it.

60. An NBC/Wall Street Journal poll in December 2004 found that 35% of respondents agreed with the proposition that the election had given the president a mandate to “carve out” private Social Security accounts; 51% disagreed.

61. See the data in Bowman (2005:18-28); see also the *Los Angeles Times*, 15 January 2005: support for a carve-out fell from 42 to 23% of respondents when it was linked to lower guaranteed benefits.

62. See Bowman (2005) for detailed results; most striking, perhaps, was the ABC News/Washington Post poll of 21-24 April, which found public support for personal accounts at its lowest level since the question was first asked in May 2000.

63. Pew Research Center data suggested that the more someone was dependent on Social Security and concerned with its future, the less he/she supported the administration’s reform proposals. See Schneider (2005) and Edwards (2007:290-2).

64. See CSS (2001:33) for the Commission’s reasoning, which lays great stress on the argument that low-income Americans and minorities, who tend to have shorter life expectancies, are particularly disadvantaged by the absence of inheritable assets from the current Social Security system.

65. See especially Edwards (2003), who argues that “going public” fails more often than it succeeds: presidents rarely manage to arouse sufficient public support for their proposals to pressure Congress into adopting them unless it is otherwise inclined to do so.

66. Kernell (1997) suggests that unpopular presidents risk undermining support for their message. This seems to have occurred in 2005 – Bowman (2005:18) observes that support for personal accounts fell when survey questions mentioned the president’s name in conjunction with the proposals, which is likely to reflect both partisanship and declining public support for the president overall.

67. Bowman (2005:35-36) presents data from a variety of major polls from 1981-2005; the Democratic Party’s edge on Social Security issues typically exceeded 20 percentage points.

68. In a May 2005 Harris poll, for example, 49% of respondents endorsed the Democrats’ charge that the administration’s real objective was to terminate Social Security.
69. The importance of the Democrats’ “ownership” of the Social Security issue is addressed at length by Ross (2007). See also Edwards (2003) for extensive evidence on the limited ability of US presidents to lead public opinion.

70. See the data in Bowman (2005).

71. Treasury (2007) estimates that an increase in the payroll tax rate by 3.5 percentage points starting in 2007 would have much the same effect on system solvency as a 5.8 percentage-point rise starting in 2041 and would impose much lower costs on younger cohorts.


73. Furman (2005) estimates the losses at an average income level of USD 90 000 or more at 29% for a worker retiring in 2045 and 49% for one retiring in 2075.

74. See, for example, Furman (2005).

75. Indeed, the administration eventually acknowledged explicitly that personal accounts would not resolve the funding issue but they were intended to make the solution more attractive to the average worker; Edwards (2007:253).

76. On this issue, see Ross (2007).
Bibliography


Weaver, C., ed. (1990), Social Securityʼs Looming Surpluses: Prospects and Implications, AEI Press, Washington, DC.

Part III

Case studies in labour-market reform
Chapter 9

France: The contrat d’insertion professionnelle, 1994

Summary and main conclusions

In an effort to reduce France’s exceptionally high rates of youth unemployment, the government of Edouard Balladur attempted in 1993-94 to introduce a new form of employment contract with a lower minimum wage for the young. The contrat d’insertion professionnelle (professional insertion contract, CIP) would have allowed employers to pay younger workers as little as 80% of France’s statutory minimum wage, the SMIC, for up to two years, regardless of the employees’ qualifications and with no obligation to provide training. The introduction of the CIP, which quickly came to be known as the “SMIC-jeunes” (youth SMIC), triggered several weeks of sometimes violent protests that, in the end, compelled the government to scrap the measure just over a month after it had issued the decrees outlining how the new contract was to work.

The CIP proposal failed for a number of reasons, including:

- **A fractured political dialogue.** While deteriorating labour-market conditions created an understanding of the need for reform and the political position of the government looked quite strong, organised interests were largely excluded from policy making and social tensions were marked. Negotiations took place “between government and the street”.

- **Lack of clear purpose and planning.** A close study of the evolution of the CIP proposal in late 1993/early 1994 shows that there was no clear, shared understanding within the government as to whom the CIP was meant to help and how. Competing agendas influenced the detail of the proposal, so that the measure eventually legislated bore little resemblance to what had originally been conceived.

- **Lack of a communications strategy.** Until the protests erupted, the government made no real effort to explain the reform to the public, nor does it even appear to have recognised the need to do so. In particular, no attempt was made to sell the reform to its intended beneficiaries – French youth. Yet the fact that the reform appeared to “target” young people proved a major political problem for the government, which was accused of devaluing their qualifications and treating them as a source of cheap labour. Closely related to the government’s failure to formulate a clear communications strategy was its inability to respond to charges that it was trying to institute a de facto youth minimum wage. The government denied any such intention, while adopting a CIP that, in practice, was difficult to distinguish from a SMIC-jeunes.

- **Lack of political will.** When the protests began, the government almost immediately began making concessions, and within a month it had dropped the
CIP entirely. It was unwilling to face down the protests. In part, this was because the CIP as it emerged from the policy process was a rather incoherent compromise, lacking strong support within the government: no one really “owned” it. However, the government’s lack of determination in the face of the protests also reflected its fear that a major confrontation could put social stability at risk.

- **Lack of a mandate.** While the government was clearly under pressure to do something about high and rising unemployment in 1993, nothing along the lines of the CIP had been discussed in the spring election campaign, nor was there any serious attempt to prepare the public for this approach to reform. On the contrary, it seemed at odds with the government’s strong public commitment to protecting the SMIC.

**The context for reform**

*Framework conditions for reform*

France has traditionally been seen as a centralised state, with a tradition of state-led policy making. Although social partners are consulted, there is no tradition of corporatist “concertation” and the institutional preconditions for concertation are weak. Natali and Rhodes (2004) argue that the French labour movement presents something of paradox for policy makers: it is too strong to be isolated but too weak and fragmented to be a real partner in corporatist concertation. Historically, French unions have proved far more effective in mobilising protest than consent. Nor is there a strong tradition of corporatist consultation with the business community: the *patronat*, too, is relatively weak. The limited role of social partners in policy making leaves the government of the day considerable freedom of manoeuvre, but, as will be seen, it can also leave it politically exposed when things go wrong.

The economic backdrop to the CIP episode was grim. Slow growth from spring 1990 culminated in a deep recession in 1992-93. While a recovery did get under way in 1994, unemployment continued rising, to reach 12.5% by mid-year. It was already above 10% and rising when the first CIP proposals were mooted within the government. Rapidly rising unemployment in 1993 added urgency to the search for solutions but increased the political risks facing the government as it sought to tackle the issue. The government’s room for manoeuvre, moreover, was limited by the deterioration of public finances during the early 1990s, which threatened both the medium-term goal of reducing public deficits to levels that would stabilise the debt/GDP ratio and France’s prospects for meeting the criteria for Economic and Monetary Union set out in the Maastricht Treaty. Nor could monetary policy grant much relief: the determination to maintain the franc’s position in the European Monetary System despite the post-unification pressure for real appreciation of the German Deutsche Mark ensured that the monetary stance remained tight.

Despite the difficult economic situation, the political position of the centre-right Balladur government in 1993-94 looked strong. It had taken office following a convincing victory in the March 1993 parliamentary elections and it commanded heavy majorities in both houses of parliament. Moreover, the government was too new to be blamed for the recession, and opinion polls consistently showed that the prime minister himself was popular. While Balladur served under Socialist President François Mitterrand, the latter was a “lame duck” by 1993, and political “cohabitation” was not a major factor in what followed.
The real political constraints on the government’s ability and willingness to pursue reforms stemmed from its own fear of social tension in the midst of a severe recession. The government believed that a confrontational approach to reform would put social stability at risk. They therefore opted for a much more circumspect strategy. The government would cautiously advance reform proposals after limited (sometimes minimal) consultation with social partners, but would retreat if they provoked significant popular resistance. While this approach did not prevent the government from pursuing a successful reform of private-sector pensions in 1993, the CIP episode was preceded by a series of government reversals in response to protest, including one involving students, who had successfully resisted proposed changes to their access to housing benefit. As the government’s opponents came to understand how the government’s approach to reform worked, they faced ever greater incentives to mobilise opposition to reform on the street, especially in view of the weakness of the parliamentary opposition.

The approach of the 1995 presidential election constituted the other major political constraint on the government in 1993-94. Balladur, then the front-runner in the polls, wished to avoid repeating the fate of Jacques Chirac less than a decade before. Chirac had served as prime minister during the right’s first period of cohabitation with Mitterrand, and his pursuit of privatisation and other reforms proved unpopular and helped cost him the 1988 presidential contest. The emerging competition between Balladur and Chirac as rival candidates for the 1995 election was also a factor: many in Balladur’s team were convinced in early 1994 that Chirac supporters were orchestrating criticism of the prime minister from within the parliamentary majority.

**Actors and interests**

Ultimately, the key players in the CIP episode turned out to be the government, led by Balladur and Labour Minister Michel Giraud, and the student movement. It was, indeed, a negotiation “between the government and the street”. Yet the students’ active involvement in the debate really began only at the implementation phase. Numerous other players were deeply involved in the formulation and adoption of the CIP legislation. Within the government, the labour ministry clearly played the lead role, but, as will be seen, the finance ministry’s views prevailed with respect to one major issue, and the two ministries’ lack of a shared vision as to the objectives and instruments of reform was to prove a major problem. Parliament, too, was an important player. While the executive dominated the policy process, the parliamentary debates over the bill did result in significant changes, largely reflecting the centre-right majority’s determination to be seen to be doing something about rising unemployment among young graduates.

France’s major union confederations – the moderate CFDT, the harder-line CGT and the smaller Force Ouvrière – participated in the reform’s parliamentary phase. However, their role was marginal, as one would expect, given both France’s tradition of state-led policy making and the presence of a centre-right government with solid parliamentary backing. More surprising, perhaps, is that the patronat played almost no role in the CIP reform whatsoever. The contract was ostensibly intended to address business concerns, and, indeed, the left repeatedly accused the government of doing the bidding of the employers and selling out French youth to the bosses. Yet neither the General Confederation of Small and Medium Enterprises (CGPME) nor the National Council of French Employers (CNPF) was particularly active at any stage. The patronat displayed no apparent inclination to support the government once the protests began. Nor does the evidence suggest that the government expected it to do so: the government made no attempt to secure a public show of support from French business. A strong public stand...
by employers might in any case have reinforced the perception that the government was advancing business interests at the expense of young people.

**Issue history**

By 1993, France had experienced a steady secular increase in unemployment over successive cycles for 20 years. Youths, women and low-skilled workers faced especially acute problems. Between the early 1970s and the mid-1980s, the unemployment rate roughly quadrupled, but youth unemployment rose six-fold. The number of new hires of school leavers directly on completion of studies, apprenticeships or military service dropped by 48% between 1973 and 1985.\(^{13}\) While unemployment among the 15-24 age group fell after 1985, thanks to targeted employment schemes for youth, it began rising again in 1991. By 1994, it had returned to the levels of the mid-1980s, as activity slowed and the number of beneficiaries of labour market programmes fell. Altogether, only about one-third of young people were in work in 1994, and only about 25% were in jobs paying market wages (Table 9.1).

Successive governments tried a wide range of different measures in an effort to arrest the secular rise in unemployment and to reduce youth unemployment in particular.\(^{14}\) From 1989, there was increasing emphasis on active rather than passive measures, and the number of beneficiaries of active labour-market policies rose from 750 000 in 1989 to 1.25 million in 1993, as expenditure on job-creation schemes tripled. Nevertheless, France remained below the OECD average with respect to the share of unemployment-related spending devoted to active measures.\(^{15}\) Steps were taken to make unemployment insurance less generous.\(^{16}\) The policy of encouraging early retirement to ease labour-market pressures was abandoned, and successive cabinets took a series of hesitant steps towards greater labour-market flexibility.\(^{17}\)

| Table 9.1. Main characteristics of youth activity in France, 1983-1994 |
|---|---|---|---|---|---|---|---|
| Per cent of population aged 16-25, March of the corresponding year |
| In employment | 46.8 | 42.8 | 42.7 | 40.7 | 40.6 | 37.7 | 33.9 |
| Assisted by specific measures | 5.3 | 5.6 | 10.8 | 8.9 | 8.5 | 9.4 | 7.3 |
| Apprenticeships | 2.6 | 2.5 | 2.6 | 2.7 | 2.6 | 2.3 | 2.7 |
| Other (including trainees) | 2.7 | 3.1 | 8.2 | 6.2 | 5.9 | 7.1 | 4.6 |
| Unassisted | 41.5 | 37.2 | 31.9 | 31.8 | 32.1 | 28.3 | 26.6 |
| National service | 2.9 | 2.9 | 2.7 | 2.7 | 3.2 | 3.0 | 2.8 |
| Unemployed | 9.8 | 12.4 | 11.2 | 8.7 | 7.7 | 9.4 | 10.2 |
| Inactive | 36.7 | 38.6 | 40.3 | 43.3 | 48.2 | 49.3 | 50.8 |
| In education | 29.9 | 32.0 | 34.6 | 38.2 | 42.8 | 44.4 | 45.7 |
| Others | 6.8 | 6.6 | 5.7 | 5.1 | 5.4 | 4.9 | 5.1 |
| Residual | 3.8 | 3.3 | 3.1 | 4.6 | 0.3 | 0.5 | 1.4 |
| Memorandum item: Youth unemployment rate | 18.9 | 23.7 | 21.7 | 17.9 | 17.4 | 22.2 | 23.2 |

*Note: Care should be taken in comparing 1991 with earlier years, as the methodology changed.*


The main employment schemes targeted at young people included numerous special contract regimes that involved a combination of employment and training; they usually offered some exoneration from social charges for the employer. These included, *inter alia*, the *contrat d’adaptation*, aimed at young people with some formal qualification who nevertheless might need some complementary professional training in connection with their first job, and the *contrat d’orientation*, which targeted less qualified youths who faced particular difficulty entering the labour market. The young were also
among those targeted by the *contrat emploi-solidarité* (CES), which provided beneficiaries with short-term employment in the public or not-for-profit sector. Other schemes provided for the state to pay social security contributions and in some cases subsidies for apprentices and unskilled youth. In 1992, the Socialist government of Pierre Bérégovoy even considered following the Dutch example and adopting a youth minimum wage, but it backed off in the face of opposition from the unions and its own parliamentary party. The impact of all these efforts was mixed. The OECD (1995:63) estimated that net job creation from public schemes might represent up to 20% of gross hirings, after allowing for windfall and substitution effects, and that about 50% of gross employment creation was the result of active labour-market schemes, including 23% of private-sector job creation. Yet the secular trend in unemployment was still upwards.

There was a great deal of debate about the causes of France’s unemployment problem and, in particular, about the extent to which it was cyclical as opposed to structural in character. Many analyses emphasised the economic cycle and tight macroeconomic policies, as well as a hypothesised “ladder effect”, whereby the shortage of jobs arising due to tight macroeconomic policies led to the formation of a “job queue”. This queue disadvantaged the less skilled, as more qualified workers moved down the “skill ladder” in search of employment. Those inclined to focus on structural causes stressed, among other things, the comparatively high cost of unskilled labour. While French unit labour costs were not particularly high compared to those of other European OECD members, the cost of unskilled labour was – a fact that many attributed to the level of the SMIC.

The ratio of the SMIC to the average wage, at just over 50%, was the second-highest in the OECD, and some empirical work in the early 1990s pointed to an elasticity of unskilled unemployment to the SMIC of around 0.2. There was also evidence to suggest that the high SMIC created a particular barrier for young people. The ratio of the youth unemployment rate to the total unemployment rate had risen in line with the relative level of the SMIC until the mid-1980s, when the introduction of active measures aimed at young people and a rise in the school-leaving age weakened the link. Moreover, the relationship between the relative level of the SMIC and the rate of youth unemployment remained extremely close into the early 1990s. Econometric work pointed to an elasticity of youth unemployment to the SMIC of 0.1 to 0.25.

One of the major priorities of French labour-market policy from 1993 on was therefore to reduce the cost of unskilled labour. Since social charges amounted to 49% of the cost of labour at the level of the SMIC, efforts to reduce the cost of unskilled and low-skilled labour were centred largely on relieving employers of some portion of the social contributions they paid for workers earning less than 1.5-1.6 times the SMIC. Leaving take-home pay untouched meant that incentives to work were not affected. A report presented to the Socialist government just before the 1993 elections proposed addressing this issue in exactly this way, and its view was reinforced in a report commissioned by the new Balladur government and prepared by a working group headed by Jean Matteoli, then President of the Economic and Social Council. The prime minister himself stressed the need for such an approach when he first called for the preparation of a major law on labour-market issues – the law that would eventually include the CIP among its provisions. This emphasis on making less skilled workers less expensive to employ forms an important part of the backdrop to the emergence of the CIP. While the government continued to reject a SMIC-jeunes, it broadly accepted the Matteoli report’s call for “the development of measures to lighten the minimum salary during the first two years of employment, matched by exonerations from social charges in return for training provided by the enterprise.”
Policy process: formulation

The origins of the CIP

Not surprisingly, given this combination of a severe recession and a long-term secular rise in jobless rates, unemployment was a major issue in the 1993 parliamentary election campaign, and labour-market reforms were a high priority for the Balladur government from the moment it took office. However, despite the rise in joblessness among the under-25s, the issue of youth unemployment as a particular problem had not loomed large in the campaign, and nothing resembling the CIP was put forward at that time. The CIP appears to have originated after the election, in two different ideas, conceived by different actors to address different problems. Both arose in the context of the government’s preparation of a five-year law on employment issues. These two visions were never properly reconciled, a fact that contributed greatly to the political setback that followed.

Officials in the Ministry of Finance originally conceived of something simple, limited and focused on the least-qualified young people. Recognising that a youth minimum wage would be politically unacceptable, they proposed a fixed-term contrat d’insertion that would allow employers to pay 80% of the SMIC to young people without qualifications (i.e. a baccalauréat or less), up to about the age of 20. The contract would also offer employers partial exoneration from social charges for this employment. In lieu of a specific training commitment, the contract would involve a mentoring relationship (tutorat) within the enterprise. Its purpose was simply to foster labour-market attachment among the least qualified, least educated young people, particularly those who had not even completed their baccalauréats or obtained apprenticeships.

Because the gap between productivity and labour costs was large in respect of this particular group, the aim was to make the contract as simple as possible, lest the bureaucratic burden on enterprises render it unattractive. The tutorat was included in the proposal in order to justify the payment of a sub-SMIC wage. There were already many employment relationships involving wages below SMIC, but they invariably involved formal training, and officials were aware of the controversy that might arise if they were seen to be trying to create a youth minimum wage without any training obligation. Even with the tutorat requirement and an explicit focus on unqualified, very young workers, a contrat d’insertion that involved paying some workers a sub-SMIC wage and that was not linked to any formal training requirement would create an important precedent, one that would be seen by many as an attempt to introduce a youth SMIC by stealth.

The Ministry of Labour, Employment and Professional Education, by contrast, was concerned with developing a suitable replacement for the contrats d’orientation and adaptation, which were not proving popular with enterprises, and with enhancing the employment prospects of young people with qualifications beyond the baccalauréat, who were encountering increasing difficulties on the labour market. This was all but unprecedented; a tertiary diploma had traditionally been considered a de facto “passport to employment”. These difficulties were only partly a product of the economic cycle: there was a long-term rise in unemployment among young people with tertiary qualifications, and those finding work were often employed in low-skilled jobs for which they were formally over-qualified. From the beginning, therefore, Labour Minister Giraud favoured a two-level CIP, with somewhat higher pay for better educated young people. The labour ministry also came down in favour of making exoneration from social charges contingent on the provision of formal training – not just a tutorat.
When the inter-ministerial committee met on 17 August to finalise the text of the five-year law for submission to parliament, the finance ministry’s position prevailed with regard to the level of remuneration envisaged.\textsuperscript{32} In other respects, however, the CIP proposal differed considerably from what finance ministry officials seem to have had in mind. The draft proposal that went to the National Assembly provided for a contrat d’insertion\textsuperscript{33} available to young people under the age of 26 – a relatively high age threshold for any sort of youth minimum wage and a higher one than finance ministry officials had initially envisaged – with a maximum educational attainment equivalent to a baccalauréat. The level of remuneration was to be set by government decree but was clearly to be below the SMIC, and partial exoneration from social charges was made conditional on the provision of formal training for at least 10\% of the period of the contract. The contract was to be for 6-12 months and renewable once, making for a maximum duration of two years. A tutorat was required, but its parameters were not defined.

\textit{A contested diagnosis}

As is clear from the foregoing, there was no single government view as to what the contrat d’insertion was designed to achieve or whom it was intended to benefit. In the view of some participants, this reflected, at least in part, the speed at which the process unfolded: the preparation and adoption of a law 83 articles in length, introducing hundreds of modifications into France’s labour code, was executed in a little over six months, with the secondary legislation following soon thereafter. The finance ministry, for its part, was acting on the basis of a diagnosis similar to that of the Matteoli group, but the labour ministry does not seem to have shared this view. It feared that the substitution effect of a youth SMIC, in the absence of any training or other requirements, might be substantial, with young people displacing older low-skilled workers. The labour ministry remained worried about the implications of paying a youth SMIC. Despite having lost the argument within the government over the possibility of a scale of CIP remuneration levels, Mr. Giraud declared on 3 September that it was “out of the question” that sub-SMIC remuneration might be paid to young people in the absence of formal training. He publicly proposed limiting the CIP’s application to those with educational attainments below the baccalauréat level\textsuperscript{34} and making it contingent on the provision of training for at least 15\% of the duration of the contract.\textsuperscript{35} The CIP thus entered its parliamentary phase sponsored by a minister who was publicly uncomfortable with some of its principal provisions and openly in favour of revising it substantially.

\textbf{Policy process: adoption}

The adoption of the five-year labour law was, on the whole, a fairly straightforward affair. The government commanded large majorities in the National Assembly and the Senate, and the executive was charged with preparing the secondary legislation needed to implement the law’s provisions. However, the CIP proposal underwent significant changes during its legislative phase.

\textit{The discussion of the reform}

The speed with which the five-year law was prepared limited the scope for consultation with social partners. The Labour Minister discussed various issues with trade union and business leaders during the spring and summer, but more or less formal consultations did not begin until the government had prepared the text for submission to
the National Assembly. Representatives of the trade unions and business groups met the prime minister to discuss the law and subsequently presented their views to the National Assembly’s Commission on Cultural, Family and Social Affairs. In hindsight, it is striking that the proposed contrat d’insertion attracted very little attention at this stage: the representatives of CGT, FO and CFDT sharply criticised many aspects of the law but said nothing about the new contract, nor did the spokesman for the CNPF.  

The only substantial comments came from the small business federation CGPME and the white-collar Confédération Générale des Cadres (CGC). The CGPME expressed concern that the proposed contract was a “hybrid”, precisely because it aimed to create a single framework for young people ranging from teenagers who had not even completed their baccalauréat to individuals in their mid-20s with vocational qualifications. The CGPME urged that it be more clearly targeted at the former. In a similar vein, the CGC called for the creation of a contrat junior for young diplômés entering the labour market for the first time, offering employers substantial exoneration from social charges but paying a salary well above the SMIC. The Commission took up the CGC’s idea and adopted an amendment creating a contrat de perfectionnement des cadres. It also approved an amendment to make training obligatory for all contrat d’insertion.

The National Assembly considered the government draft of the five-year law in late September and early October. Article 40 of the text, introducing the contrat d’insertion, was the subject of some debate but did not emerge as a major focus of the opposition’s attack on the bill. Left-wing deputies criticised the contract’s provisions and accused the government of trying to institute a SMIC-jeunes, while deputies from both sides expressed concern about the vagueness of some provisions, which were to be resolved in the secondary legislation needed to apply the law. This was true, not least, of the level of remuneration. The government assured the Assembly that these issues would be resolved in close consultation with social partners. This promise implicitly addressed another criticism of the proposal: the contrat d’insertion had been devised entirely by the government, whereas the contracts it was to replace had – whatever their other defects – been negotiated by the social partners before they were fixed in legislation. According to Mr. Giraud, the government had prepared the CIP itself only after his attempt to reach agreement on a new contract in negotiations involving the unions and the patronat had stalled. He argued that it was precisely because the contract had been articulated by the government rather than the social partners that it was appropriate to leave many of the details open: the government would then be able to resolve these together with the unions and the employers. A number of contentious points were thus left to be resolved when the application decrees were prepared. The government promised that these decrees would be issued rapidly – a promise that implied that the time for consultation with social partners would be limited.

With its heavy majority, the government easily defeated amendments proposed by the Communists and Socialists that would have scrapped or substantially modified the contract. It also thwarted the Commission’s attempt to make training an obligatory element of the CIP. However, the government gave way to pressure from its own backbenchers to make the new contract available to young people with tertiary education, raising the educational threshold to include “level III”. Relief from 50% of social charges was linked to the provision of training for at least 15%, rather than 10%, of the duration of the contract. The more significant of these changes was the former, as it extended the coverage of the new contrat d’insertion to young people with up to two years of tertiary education. In this, it reflected the desire of the Assembly’s centre-right majority to aid the employment prospects of middle-class young people, a priority that
reflected concern about the unprecedentedly difficult conditions faced by new graduates on the labour market and perhaps also electoral calculation. The consistent concern with the plight of unemployed *jeunes diplômés* was thus entirely understandable. It also reflected the labour ministry’s priorities. This made it relatively easy to amend the legislation in response to the majority’s concerns – the ministry was regaining in the parliamentary phase of the discussion some of what it had lost when the finance ministry’s position had prevailed in the inter-ministerial committee in August.

It was in the Senate that the CIP really began to come under attack from the left. Speaker after speaker repeated the charge that Article 40 represented the creation of a *SMIC-jeunes* by stealth. Opponents of the measure cited an INSEE study commissioned by the Senate which suggested that a 20% cut in the SMIC for youth would result in 136 000 additional hires among young people but that the net effect on employment would be around 30 000, owing to a large substitution effect. While the findings of this study were contested, its appearance underscored the fact that the government had presented no serious econometric analysis of the possible consequences of the CIP proposal: it was thus difficult to refute the charge that the CIP would merely trigger some reallocation of unemployment, quite possibly at the expense of older workers. Indeed, the wider the remit of any youth SMIC (in terms of maximum age or educational attainments), the greater was the likely scale of substitution. While older workers would probably not have had much to fear from unqualified teenagers earning a lower wage, a population of educated people in their twenties available to work for a wage below the SMIC would have been more threatening.

Critics in the Senate also attacked the vagueness of the proposals, particularly the “professional project” that youth with post-*baccalauréat* education were supposed to execute under the terms of the contract (in lieu of the *tutorat* promised to their less qualified colleagues). Like the *tutorat*, the *projet professionnel* was conceived as a means of justifying a level of remuneration below the SMIC – even if formal training were not involved, the contract would involve a degree of formal training (*formation professionnelle*) and could thus be defended against the charge of introducing a youth SMIC.

In the Senate, as in the Assembly, the government’s majority made it relatively easy to resist attempts to abolish the CIP or modify it significantly. The only major innovation adopted by the Senate was the provision making formal training for at least 15% of the duration of the contract obligatory in cases where it was renewed.

**The government’s communications strategy**

Throughout the discussion of the five-year law, the government’s position with respect to the CIP was confused. The minister sponsoring the CIP was uncomfortable with some of its provisions and thus ready to acquiesce when the parliamentary majority wished to modify them. The government also finessed many of the most contentious issues – including the critical question of remuneration for young people with tertiary qualifications – by insisting that they would be resolved in discussion with social partners during the preparation of the decrees implementing the law. Moreover, the government continued to insist publicly that a “youth SMIC” was not part of its plans. All indications were that reform of the SMIC would not be seriously considered until after the 1995 presidential election. The period between December and February, when the secondary legislation required by the five-year law was being prepared, offered a brief interval in which the government could have modified its position, but there were neither
public statements on the subject nor any private discussions of the issue with social partners. In short, nothing was done to prepare public opinion for what was to come in the decrees implementing the CIP.

One extremely important issue was left out of the debate altogether: education policy. The proposal to pay a sub-minimum wage to young people with tertiary qualifications raised questions about the value of those qualifications and why French young people who had obtained their *baccalauréats* and university diplomas were apparently so ill-prepared for the job market. If the rationale for extending the CIP to level III qualifications was correct, then this was a major issue to confront, especially in view of the widespread view of the diploma as a passport to secure employment. However, neither the students nor their parents nor the teaching profession wished to question the value of these qualifications, and the government showed no interest in tackling the issue head-on. Given the history of failed attempts to reform higher education, many of which had proved politically damaging to their sponsors, this is not entirely surprising. As a result of this reticence, no serious attempt was ever made during the preparation of the CIP to explain why some *jeunes diplômés* might, for a time at least, have to accept a lower wage.

**Policy process: implementation**

The five-year law reached the statute books on 21 December and was published the following day, just ahead of the holidays. The timing of its adoption helped ensure that there was no immediate reaction to its passage from the unions or other interested parties. Further details of the CIP were worked out over the months that followed and set out in a pair of government decrees published in late February. The decrees stipulated that employers could pay CIP employees 80% of the SMIC regardless of their qualifications and with no obligation to train beyond the *tutorat* (for those with a *baccalauréat* or less) or *projet professionnel* (for those with some tertiary education). They also provided a remuneration scale for cases in which training *was* provided; this ran as low as 30% of the SMIC for 16-17 year olds, 50% for those aged 18-20, and 65% for the 21-26 age bracket. The decrees offered exoneration from 50% of social charges if at least 15% of the duration of the contract was devoted to training. They provided minimal detail concerning the *tutorat* and left enterprises considerable freedom with respect to the provision of training, which could be offered by the firm itself if it had its own training centre.

Despite the relative quiet in which passage of the five-year law had been received in December, the publication of the February decrees implementing the CIP provisions triggered a strong reaction on the part of a group whose views were not canvassed during the law’s preparation – French students. Student unions and other organisations were not consulted during the preparation of the five-year law, largely because it was seen as a labour-market reform rather than a question of educational policy. While there were regular contacts between student leaders and officials responsible for education policy, the minister responsible for higher education was not closely involved in the preparation of the five-year law, and the students themselves were primarily concerned with housing benefits in the autumn of 1993. As a result, they failed to recognise what the new legislation might mean for them. The government, for its part, made no attempt to sell the CIP reform to the public at large or to the young people who were supposed to be its principal beneficiaries.
One of the student unions, UNEF-ID, was quick to seize on the CIP issue, in part because its leadership faced a strong left-wing challenge within the union. UNEF-ID quickly called for a meeting of student organisations and the major trade union syndicates, and on 28 February, 17 organisations, including the CGT trade union, issued a joint call for protests against the contract. In any case, demonstrations were already being organised by students around the country. Students and recent graduates holding level III vocational qualifications – the *Brevet de technicien supérieur* (BTS) and the *Diplôme universitaire de technologie* (DUT) – were particularly active in opposing the measure. Although the other union confederations were initially slow to react, perhaps reflecting a traditional mistrust of student activism, they soon joined the CGT in supporting the students’ protests.

Protests erupted across the country, leading to the first serious consultations with student groups on the issue. These resulted in a series of government concessions. On 2 March, the prime minister met with student leaders and on 3 March, the CIP became the sole topic of discussion at a long-planned roundtable on social policy hosted by the prime minister. Mr. Balladur insisted that he understood the criticisms raised by opponents of the CIP and promised the preparation of a “complementary” decree in consultation with social partners. The government promised that this decree would tighten up the rules on the *tutorat* and that the reference to the SMIC would be removed; the salary would be calculated on the basis of the reference salary for each branch of industry or commerce. On 8-9 March, the labour ministry raised the prospect of excluding young people with level III qualifications from the provisions allowing payment of a discounted wage, provided their qualifications corresponded to the requirements of their jobs. The demonstrations nevertheless continued and, in a number of places, turned violent. There was growing concern within the government that it was just a matter of time before someone was killed. With polls showing a heavy majority of the public against the CIP, even the moderate unions hardened their positions. The CGC, which had previously viewed the measure rather favourably, warned of “open war” if the CIP were not withdrawn.

The government was clearly caught off-guard by the wave of protests and at a loss as to how to respond. On 18 March, the newspaper *Libération* carried an open letter from the prime minister to French youth. Mr. Balladur also addressed the business community, telling them that it was a “moral and economic duty” to hire young people without “abusing” the CIP. In respect of the students, at least, the appeal fell on deaf ears. The protests continued.

At this point, electoral politics seemed for a time to transform the situation. While the public at large appeared hostile to the CIP and Mr. Balladur’s own poll ratings had been falling, the centre-right did well in the first round of local (cantonal) elections on 20 March. Its combined score of 44.6% even represented a slight advance on its first-round score in the parliamentary elections a year earlier. The implication seemed to be that, despite the on-going protests, a “silent plurality” of French voters still supported the government. The government moved quickly to exploit this victory and put the CIP conflict behind it. On 21 March, a decree revising the 23 February application decrees was confirmed. It contained significant concessions: the requirements of the *tutorat*, and the qualifications for those who served as *tuteurs*, were made more stringent, and the salary for CIP employees was set at 80% of the reference salary for a given branch but in any case *no lower than* the SMIC. Finally, the decree stipulated that at least 20% of the duration of the contract must be devoted to “practical training” and no more than 80% to “productive activities”. At the same time, however, the government insisted that there
would be no further concessions and no withdrawal of the CIP. The unions appeared to share the government’s view that the worst was over. Their public statements now focused on their determination to discourage employers from resorting to the CIP. The implication was that the contract would remain a part of the labour code.61

The students, however, did not give up, nor were they satisfied with the concessions contained in the 21 March decree. Demonstrations continued throughout the week, culminating on 25 March, with rallies in Paris, Nantes, Lyon, Toulouse, Grenoble, Lille, Montpellier, Marseilles and Saint-Étienne. On 26 March, student leaders from across the country met in Paris to plan nation-wide protests for 31 March. The continuing demonstrations generated increasing disquiet among parliamentarians on the centre-right, some of whom feared a repetition of the events of December 1983. Some of the prime minister’s supporters were also increasingly convinced that the pro-Chirac faction was using the CIP protests to undermine the prime minister.

Against this backdrop, the government suffered an unexpected setback in the second round of local elections on 27 March, as a resurgent left captured control of the Dordogne, the Gironde and Réunion, while depriving the right of victory in a number of locales which the first-round results suggested that it would win comfortably.62 Within the government and the parliamentary majority, this second-round setback was attributed to the CIP.63 By noon the following day, the prime minister had suspended the relevant articles of the five-year labour law, and ordered the head of the National Agency for Employment to study ways of replacing the CIP. Two days later, a further CIP decree appeared, annulling the previous three,64 and the following month, the government replaced the CIP with a new monthly subsidy of 1 000 francs to be paid over nine months to any company hiring an under-26 for at least 18 months; the subsidy for hires before 1 October 1994 was set at 2 000 francs per month.65 A collective agreement signed in the summer updated provisions for apprenticeship contracts, and the provisions of the five-year labour law pertaining to the CIP were repealed in August.66

Conclusion

The CIP was seen as a threat by many of those whom the governing centre-right would ordinarily have seen as its natural supporters. By offering employers the option of paying 80% of the SMIC to workers under 26, regardless of educational attainments and in the absence of any formal training, the CIP set a potential standard for all young employees. Hence, the strength of the reaction to the CIP reflected the perception that it really did institute a SMIC-jeunes. The tutorat-projet professionnel was not sufficient to provide political cover for such a move, and the wide discretion granted to employers with respect to the provision of training in return for relief from social charges was regarded as unacceptable. A CIP paying a sub-SMIC wage to unqualified teens might have been politically viable. One that encompassed graduates in their twenties was not: they were unwilling to be treated on the same basis as youths with incomplete secondary education. In seeking to reassure middle-class parents that it was trying to address the employment needs of their children, the government upset the children and called into question the value of the qualifications for which they had worked.

The fact that the CIP was felt to target young people also seems to have presented a real political problem. Reforms that appear to target particular groups are often very sensitive in France, particularly where the young are concerned. The struggles surrounding the proposed contrat de première embauche (first employment contract, CPE) in 2006 reflected some of the same sensitivities. This is one reason why French
governments in recent decades have tended to give way in the face of strong student protests – in addition to the CIP and the CPE, governments of both right and left have backed down over such measures as the education reform proposals of 1986 and 1998-99, as well as the 1993 housing benefit reform. The perceived need to protect the young makes it difficult to adopt policies that are believed to erode their rights or infringe on their interests.

The CIP debacle was in part the result of a confused policy process. The government itself seems not to have had a clear, common understanding of what the CIP was meant to achieve and why. Some wanted it to be simple, flexible and narrowly targeted at a specific group of educationally and economically disadvantaged teens; others sought to create a framework to enhance the employment prospects of a much larger and more diverse population of young people. Given wider and more extended consultations, these differences might have come to the fore more clearly and been resolved. However, the government’s determination to proceed rapidly with the five-year labour law – and it is important to recall that the CIP was the subject of just one of the law’s 83 articles – ensured that discussion was limited and warning signals were ignored. It also meant that there was little or no time to do any serious analysis concerning the potential impact of the reform. As the policy process unfolded, the CIP proposal, which had been relatively simple to start with, became ever more complex, as the government made piecemeal concessions to address the concerns of the many different constituencies involved. Ultimately, the result was a policy that satisfied almost no one.

Lack of clarity about objectives also contributed to the government’s failure to explain the reform to the public. As noted above, the government could claim no electoral mandate for a reform along the lines of the CIP, so effective public communication was essential, particularly in view of the sensitivities surrounding anything to do with the SMIC. Yet the lead minister on the CIP was not entirely comfortable with the proposal, and at times this was publicly apparent. The failure to bring the issue of education policy and the relationship between education and labour-market outcomes into the discussion was particularly critical. While most people had little difficulty in understanding why something like the CIP might be needed to help very young people with no formal qualifications to enter the labour force, nothing was said or done to explain to French students or to the public at large why those with a full secondary education and as much as two years’ university or professional study still faced such severe problems when entering the labour market. In the end, this failure to define and communicate widely the aims and rationale of reform contributed to the government’s unpreparedness for the backlash that followed publication of the February decrees. Not for the first or last time, a measure that had been relatively smoothly adopted triggered a fierce reaction when the moment for implementation arrived, precisely because that was when those most directly affected became fully aware of it.67

More than a decade later, the problem the CIP was meant to address remains a serious issue for France. The unemployment rate for the under-25s continued to rise, reaching 30% in 1997. At 23.2% in 2006, it was roughly equal to the level of 1994 and ranked fourth highest in the European Union, despite the fact that further declines in participation rates among the young have left France with one of the Union’s lowest rates of labour-force participation among under-25s. Indeed, the proportion of under-25s in employment in 2005 stood at just over 29%, down from around 42% two decades earlier. Yet as the recent history of the CPE suggests, France is still some way from finding an economically sound and politically acceptable way to address the problem of youth unemployment.
Notes

1. *Salaire minimum interprofessionnel de croissance.*

2. Where formal training was included as part of the contract, the wage level might be lower still and the employer would gain partial exoneration from payment of social charges for the employees in question.

3. Unionisation rates were below 10% overall by the early 1990s and below 5% for the private sector, and French unions are divided among a number of competing peak-level confederations. Nevertheless, more than 90% of the workforce is covered by sectoral collective bargaining, as the vast majority of collective agreements are extended by law to all employees in a sector. Moreover, unions play an important role in administering unemployment and health insurance, as well as the pension system.

4. French budget deficits prior to the early 1990s were below the OECD average and well below Maastricht but they rose sharply in 1993-94. See OECD (1995).

5. France’s past history of devaluation may have increased the costs of its defence of the franc at this time.

6. The government was not alone in its fears: polls at the time suggested that most voters (69% in one sample) expected “serious social crisis” in the coming months; see *Oxford Analytica Daily Brief*, 24 March 1994.

7. Levy (2000:333) describes *la méthode Balladur* as “a kind of direct negotiation between the state and the streets.”

8. For details, see the case study on the 2003 pension reform.

9. The other government retreats included the provision of subsidies to Breton fishermen protesting cheap imports, the withdrawal of a plan for restructuring Air France following union protests, and the abandonment of planned changes in the financing of capital expenditure in education in the face of opposition from teachers.


11. France’s majoritarian electoral system probably reinforced Balladur’s caution: the French two-round system tends to ensure a high degree of bipolar competition, and this can make it riskier for political leaders to advance reforms, even if their methods and goals do not differ much from those of the rival bloc. See Schludi (2003).


13. OECD (1992:65). At the same time, total new hires from the ranks of the unemployed rose four-fold.

14. See OECD (1992) for a summary of the measures adopted over the preceding years. These included steps to lower the costs incurred by the least-skilled and other vulnerable groups in finding employment, to develop training and apprenticeships, and to raise incentives to look for work.

16. Eligibility requirements were tightened, and both phased reductions in benefits and shorter maximum duration of benefits were introduced. However, this was partly offset by increased take-up of other benefits, like the revenu minimum d’insertion, or RMI, an income benefit introduced in 1989 for working-age people who had no other income but were not entitled to contributions-based unemployment benefits. See Malo et al. (2000:257-58) for details.

17. The cap on working hours was reformulated as an annual rather than a weekly maximum, the burden of social contributions on part-time workers was reduced, and other measures were adopted to make less-than-full-time work easier. Employment-protection legislation remained largely untouched.

18. CES were for 3-12 months and offered little prospect of stable employment, although they could be extended to 24 months, with provision for training. Social security charges were waived and the state funded 85% of the young person’s pay. The programme was to expire in 1993 but was extended by the Balladur government, and in 1994, some 400 000 people were working on such contracts. See Symes (1995:67). The CES was also used widely to benefit women and the long-term unemployed.

19. In some schemes, the state did not bear the cost of employers’ contributions; social security receipts were simply reduced accordingly.


21. Trade union representatives, who blamed restrictive macroeconomic policies for the rise in unemployment, argued that higher wages, by increasing household purchasing power, would lead to increased consumption and thus job creation; see Bulletin (1993:1549ff).

22. See Gautié (1997) and Malo et al. (2000), who argue that other measures adopted in 1993 to reduce the relatively high cost of low-skilled labour do not appear to have had much impact on the unemployment differential of unskilled workers.


26. That is, the ratio of the SMIC to the average wage.


30. The Loi quinquennale no 93-1313 du 20 décembre 1993 relative au travail, à l’emploi et à la formation professionnelle.

31. Over 40% of 18-20-year-olds in work were earning below the SMIC in the late 1980s; see OECD (1992:75) and CERC (1990). A very small number of 16-18-year-olds could be paid 10-20% less than the SMIC even in the absence of training, provided they had been working less than six months. This general provision did not result from the statute on the SMIC but from secondary legislation. It was in fact very marginal, because of the existence of standard training contracts which provided for legal minimum wage reductions according to age, allied to specific training requirements.


33. The word professionelle was dropped from the initial draft prior to its submission to parliament and was not restored to the text until the Senate considered it in November. For convenience, it will be referred to as the CIP throughout.
34. In French parlance, he wished to limit the contracts to education levels V and VI, excluding those with level IV qualifications – baccaulærat, Capacité en droit, Diplôme d'accès aux études universitaires, Brevet Professionnel, Brevet de Technicien, Brevet des Metiers d'Art, Probatoire du Diplôme d'Etudes Comptables Supérieures or Brevet de Maîtrise.

35. See Jacquet (1993:77-78).


37. Bulletin (1993:1572). Socialist deputy Michel Berson also argued that the contract was a hybrid and that it attempted to address the needs of too many diverse constituencies within a single framework; see Débats (1993:3816).

38. The CGC representative spoke of 8-9 000 francs per month, at a time when the gross monthly SMIC (169 hours) stood at 5 886.27. See Bulletin (1993:1578-79).

39. For the texts of the amendments, see Jacquet (1993:377-81).


41. Communist deputies attempted to impose two conditions on the contract: training leading to an officially recognised qualification and the guarantee of a permanent contract at the end of the CIP.

42. Brevet de technicien supérieur, Diplôme universitaire de technologie, Diplôme des Métiers d'Arts, Diplôme d'études universitaires générales (bac+2) and Licence (bac+3). The contrat de perfectionnement, however, was not adopted.

43. The government also added an amendment that regulated the financing of training provided by the enterprises themselves, an issue that had apparently been forgotten in the hasty preparation of the bill.

44. There was also a genuine sense of shock at the sharp and unprecedented rise in unemployment among recent graduates. However, some centre-right parliamentarians also recognised that the jeunes diplômés they sought to help were to a large extent the children of the centre-right’s core constituents. A narrowly focused CIP, directed at teens with incomplete secondary education or, at best, a baccaulærat, would have targeted a constituency which was both less likely to vote and less likely to vote for parties of the right.

45. For the Senate’s discussion of the CIP, see Senate (1993:4151-71).


47. The prime minister himself had explicitly ruled out any reform of the SMIC, declaring that he would “not take responsibility under the pretext of fighting a crisis for destroying the system of social protection in our country.” Financial Times, 23 August 1993, p. 13.

48. For an overview of the 1970s-1980s, with particular emphasis on the travails of the previous centre-right government, see Weiler (1988).


50. The only requirements were that the tutor has a level of education at least equal to that of his/her tutees and that he/she oversee no more than three at a time.

51. This is not to say that such an attempt would have succeeded: in 2006, the government tried hard – and unsuccessfully – to persuade French students of the benefits of the proposed contrat de première embauche (‘first employment contract’).
52. UNEF-ID leader Philippe Campinchi recognised that a strong lead on the CIP could strengthen his position, while failure to act promptly would risk allowing his internal opponents to use the issue against him.

53. CGT leader Louis Viannet walked out of the discussions, insisting that the government was offering only cosmetic changes.

54. Some ministers had been in government in December 1986, when protests against the Chirac government’s education reform proposals resulted in the death of a student, an event that did serious damage to the government and helped to sink both the reforms and the political career of the minister responsible for higher education.


56. “We were guided by only one inspiration and that is to find jobs for 750 000 youths. Do you really believe that a government would be interested in attacking the young and seeking to hurt them deliberately?” the prime minister asked his readers. “I am searching with all my power to find possible solutions. I believe in your good faith. Believe in mine.”


58. In mid-March, a major poll for the first time showed more voters disapproving than approving of his performance, by a 47-45 margin.

59. Comparing the 1993 total of the “Union pour la France”, combining the Union pour la Démocratie Française (UDF), the Rassemblement pour la République (RPR) and other centre-right candidates aligned with the Union, against the 1994 total for the UDF, RPR and non-party candidates of the centre-right.

60. Décret no 94-228 du 21 mars 1994 modifiant le code du travail et complétant les dispositions relatives au contrat d’insertion professionnelle.

61. See the comments of the CFDT’s Jean Masson and Force Ouvrière’s Marc Blondel in Libération, 22 March 1993; see also Mouvement (1994).


63. Giraud (2001:147-48). Some of those involved insist that Balladur was prepared to withdraw the CIP even before the second-round vote, as he realised that the 21 March decree had not defused the crisis; they claim that he held off for fear that a pre-election climbdown would damage the right’s electoral chances.


67. A similar backlash occurred in 2004 over the so-called “conflit des recalculés”. An agreement signed in 2002 by the government, a number of trade unions (including the CFDT, the CGC and the CFTC) and various employers’ organisations envisaged changes in access to unemployment benefits. The backlash set in only at the end of the following year, when the changes began to come into force; at that point, protests by the unemployed and some of the unions forced the authorities to retreat.
Bibliography


Chapter 10

Germany: The Hartz reforms of the labour market, 2002-05

Summary and main conclusions

The adoption and implementation of the four laws known collectively as the Hartz reforms in 2002-05 marked perhaps the single biggest change in German labour-market policy in a generation. The reforms’ focus was on the supply side: they consisted chiefly of reductions in benefits and more effective active labour-market measures designed to increase the incentives for the unemployed to try to find jobs as quickly as possible and to improve their prospects for so doing. Adoption of the reforms followed a long period in which labour-market reform had largely been stalled and was followed by a further period of inaction and partial retraction. The favourable conditions of 2002-04 came about as a result of a number of factors:

- **Electoral pressure combined with a sharply deteriorating employment situation.** From late 2001, unemployment began rising sharply and perceptions of Germany’s economic condition worsened markedly. With elections approaching, the government had to be seen to act decisively to stem the rise in joblessness, particularly since it had been elected on a promise to cut unemployment. This strengthened the arguments of “modernisers” in the governing Social Democratic Party who favoured labour-market reform. Moreover, the Chancellor’s endorsement of reform proposals in the midst of the election campaign gave the SPD left and its trade union allies little choice but to endorse them, lest they damage the government’s electoral prospects. The government’s re-election then put it in a strong position to claim a mandate for reform and to begin legislating at the start of its new term.

- **A weakened status quo.** A scandal concerning large-scale data falsification at the Federal Labour Office (BA) in early 2002 put both the Ministry of Labour and the social partners, who oversaw the BA, on the defensive. The scandal made a reform of the BA unavoidable and persuaded many that a broader reform was necessary. It also offered the government an opportunity to bypass the existing corporatist institutions, which had proved incapable of delivering reform. Chancellor Gerhard Schröder thus appointed an *ad hoc* commission under Volkswagen Personnel Director Peter Hartz to prepare proposals for reform of the BA. With the Chancellor’s support, the commission interpreted this mandate rather broadly, introducing into the centre of public debate a range of proposals that would probably never have emerged from Germany’s established corporatist institutions. After the 2002 election, moreover, the Ministry of Labour was merged with the Ministry of Economics and Technology, a step that made it easier to turn the reform proposals into law.
• Careful selection and sequencing of reforms. In focusing on activation of the unemployed and changes to the benefit system, the reform proposals left untouched a range of labour-market issues that would have generated far more union opposition (collective bargaining, employment protection, etc). The reforms were adopted in sequence, with those that were technically and politically easiest being adopted first and the more difficult steps following later on.

Though adopted and implemented relatively smoothly, the reform laws – particularly Hartz IV – subsequently generated a backlash. This was partly a product of improved labour-market performance, which resulted in part from the reforms: as the situation improved, the sense of urgency about reform declined and politicians were more willing to restore benefits and relax some of the more rigorous elements of the reform. However, the backlash also reflected poor communication of the meaning of Hartz IV, both to the general public and to those directly affected by it. Finally, at least some of the resentment generated by the Hartz reforms seems to have resulted from their success: participation rates among older workers rose sharply after the reforms eliminated some avenues to early retirement, but such workers in many cases did not welcome extended careers.

The context for reform

Framework conditions for reform

The political challenges facing the centre-left SPD-Green coalition as it sought to press ahead with labour-market reforms in 2003-05 were substantial. First, the government commanded only a narrow majority in the lower house of parliament, the Bundestag, and for most of the period, the centre-right opposition held a majority in the upper house. While not all federal legislation requires Bundesrat approval, two of the four reform bills did, and split-party control of the two chambers was an important factor shaping them. Secondly, vocational training and employee relations fell within a constitutionally guaranteed sphere of collective labour relations in which the social partners played a central role, which they jealously guarded. Thirdly, the reform of unemployment benefits undertaken as part of the reforms involved the integration of social assistance programmes under the administrative and financial control of local governments with contribution-financed, federal/state-administered measures. This raised difficult questions about the allocation of costs across levels of government.

Proponents of reform also had to contend with the existing institutions of the German welfare state, which have tended to treat status-securing social insurance entitlements, based on dedicated contributions, as a sort of property right. From the beginning, unemployment benefits were meant to maintain a worker’s social status rather than to provide a safety net, and the principle that benefits be roughly proportional to contributions was therefore deeply ingrained in the system and strongly supported by German legal culture. It was also supported by German voters, who appear to have been more resistant to welfare state retrenchment in the early 2000s than electorates in Scandinavia or the Netherlands. The significance of these public attitudes as a barrier to reform was compounded by the bipolar character of electoral competition between centre-right and centre-left blocs; the major parties had to pursue vote-maximising strategies if they wished to take office.

The reforms were undertaken at a time of exceptionally weak economic performance, which had undermined the confidence of the political class in existing policies and
institutions; there was much concern that Germany was coming to be seen as “the sick man of Europe”. Labour-market performance, in particular, was disappointing. By 2002, the seasonally adjusted registered unemployment rate had been fluctuating between 9 and 12% for five years. The German Council of Economic Experts estimated that the unemployment rate would have reached 13.4% if it had included the more than 1.7 million unregistered unemployed, who were enrolled in active labour market programmes (ALMPs) or early retirement schemes designed to ease pressure on the labour market. Participation rates among older workers were among the lowest in the OECD. Moreover, long-term unemployment (over 12 months) accounted for around half the total. While high and rising unemployment clearly underscored the need for reform, it also made many potential reform measures particularly sensitive, especially those that might reduce employment protection or unemployment benefits. There was, finally, a less urgent but more structural source of pressure for reform: changes in technology and the organisation of work were steadily increasing the demand for “atypical” forms of employment that the German system was ill suited to accommodate: German legislation was still very restrictive with respect to part-time and temporary agency work, in particular.

**Actors and interests**

Adoption of the Hartz legislation required a number of cross-party compromises: although the reforms were undertaken by the newly re-elected SPD-Green coalition, the centre-right CDU-CSU and the liberal FDP exercised a de facto veto on many issues, owing to their control of the Bundesrat. Both the SPD and the CDU-CSU were internally divided over labour-market policy, with significant “traditionalist” social policy wings in both parties committed to generous benefits, employment protection and a status-securing welfare state. However, the traditionalists on both sides of the partisan divide were in decline, in part as a result of poor economic performance, and within the SPD, in particular, the so-called “modernisers”, including Chancellor Gerhard Schröder himself, became increasingly vocal, inspired in part by the example of the “New Labour” government in the United Kingdom. In any case, differences between the two chambers did not turn only on party politics: when the fourth Hartz law was debated, questions of fiscal federalism came to the fore, with the Bundesrat resisting what it saw as the extension of federal authority at the expense of the municipalities. At that stage in the process, the major associations representing counties and municipalities were also engaged. Finally, the issue of fiscal federalism, as well as the strong legal protection of acquired rights in the social insurance system, meant that the Federal Constitutional Court played a limited but significant role in the implementation of the reform.

Within the federal administration, the key institutions concerned with labour-market policy were the Ministry of Labour and Social Policy and the Federal Employment Office (BA). With some 90 000 employees, the BA was managed on a parity basis by the state and the social partners. While responsible for the administration of both active and passive measures, it was principally concerned with the latter: government spending on active measures was fairly limited, whereas passive benefits were based on legal entitlements and thus rose with the level of unemployment. The BA paid out roughly EUR 48 billion a year (roughly 2.3% of 2002 GDP) in unemployment benefits.

As a rule, the federal ministry in charge of a government bill invites experts, associations, unions, sub-national governments and other federal departments to comment on a first draft, which is then revised, if necessary, before submission to the federal cabinet. This consultative approach to law-making ensured the social partners,
particularly the trade unions, a voice in negotiations over the Hartz package. The leading roles were played by the two major employers’ bodies – the Federation of German Industry (BDI) and the Confederation of German Employers’ Associations (BDA) – and the main umbrella union organisation, the Confederation of German Trade Unions (DGB). Germany is not heavily unionised, but the unions’ influence was enhanced by their close links to the major political parties, the extension of collective bargaining agreements to cover non-union workers in many sectors and the unions’ role in administering elements of the social insurance system like the BA. Co-management of such institutions by the social partners enabled both unions and employers’ organisations to maintain a presence even in places where they had few members.

The reform proposals originated in the Commission on Modern Services for the Labour Market (the “Hartz Commission”), which was constituted in March 2002 under the leadership of Peter Hartz, the personnel director of Volkswagen, and instructed to present its final report in mid-August, some six weeks ahead of the federal elections. While the commission wrapped up its work after launching its final report, Dr. Hartz himself remained an active player, promoting the commission’s ideas to the public during the legislative phase of the reform.

**Issue history**

Traditionally, the main emphasis of German labour-market policies was on passive adjustment via a relatively generous benefit system, which aimed to maintain an individual’s status during periods of unemployment until he could find a job commensurate with his qualifications and experience (Box 10.1). The unemployed were under little pressure to accept much lower-paying jobs than those they had lost. In an economy that had no statutory minimum wage, the generosity of the benefit system also served to some extent as the basis for a de facto minimum. Relatively high levels of unemployment and social assistance benefits, combined with the very high tax wedge on labour, meant that labour market reintegration was unattractive to many benefit recipients.

**Box 10.1. Unemployment and social benefits prior to the Hartz reforms**

Under the pre-Hartz arrangements, a recipient of unemployment insurance (UI) was eligible for a benefit equivalent to 60-67% of his former wage for at least 12 months and, in some cases, for up to 32 months, depending on family status, age and contribution history. If he was still unemployed when his UI benefits were exhausted, an individual was entitled to an income-related but means-tested unemployment assistance (UA) benefit equivalent to 53-57% of his former wage. This was financed from the federal budget rather than from social insurance contributions and was paid for an unlimited period. Finally, social assistance (SA) provided a tax-financed, means-tested basic income benefit to those who had not worked (and were thus ineligible for unemployment benefits) or could not enter the labour market on health or other grounds. It was also used to top up the benefits of unemployed persons whose UI or UA benefits, as calculated under the rules then prevailing, were less than was needed to ensure the minimum acceptable standard of living.

ALMPs began to be introduced when unemployment started to rise in the 1970s and became much more important after unification in 1990, when they played a role alleviating the social consequences of the post-communist transition in eastern Germany; the BA to some extent became the focal point of social policy in the East after unification.
Although ALMPs played a growing role from the 1990s on, there was widespread scepticism, not least within the BA and the Ministry of Labour, about their capacity to improve labour-market conditions. The evidence suggests that this scepticism was warranted, at least in respect of those active measures administered by the BA. Extensive public subsidies for short-term work, public job creation and further training combined to create a secondary labour market characterised by strong lock-in effects, zero (or even negative) post-participation effects and substantial dead-weight losses. Yet the parties running the system arguably had little interest in reforming it, since most of these training and other programmes were run by the social partners themselves, who controlled around 60% of the further education sector. Support for direct integration into regular employment played only a modest role.

Unemployment was a central theme in the 1998 election campaign that first brought the SPD-Green coalition to power. The SPD vowed to reduce the number of registered unemployed from 4.2 million to 3.5 million within four years but remained broadly committed to its previous social policy focus on “status-adequate” employment reintegration and maintenance of a publicly financed secondary labour market. The new government therefore expanded job-creation schemes, widened the scope of jobs covered by social insurance and made access to training and job-creation schemes easier for older workers. It also revoked some of the labour-market reforms adopted by the previous government. With the economy growing reasonably well until 2000, unemployment was falling, making the government’s target of 3.5 million by 2002 look achievable without reform. Given how divisive labour-market policy was for the SPD, this trend made it easy not to act. However, when unemployment began to rise sharply in late 2001, the government – facing an election in 2002 – had to respond, and it was the SPD “modernisers” who took the lead. The so-called Job-Aktiv legislation, adopted in late 2001, put greater emphasis on reintegrating the unemployed into the regular labour market. Many elements of the Job-Aktiv act were found in the EU’s European Employment Strategy, which made them politically easier to legitimate in the eyes of SPD constituencies that were wary of reform.

Discussion of labour-market reforms before 2002 centred on the corporatist “Alliance for Jobs”. The Alliance had been created under the previous centre-right government of Helmut Kohl in the hope of creating a consensus on reform but had broken down, prompting the Kohl government to risk pursuing some reforms unilaterally in 1997-98. Both the content of those reforms and the government’s break with Germany’s traditionally consensual model of policy making in the social sphere contributed to its electoral defeat the following year. The Alliance was therefore revived by the SPD-Green government, but it bore little fruit. The government, itself divided over labour-market policy, had little to offer the unions in return for concessions on labour-market flexibility, particularly since it had reversed its predecessors’ reforms without demanding any quid pro quo. Moreover, the government’s efforts to ensure wage moderation in 2000-01 made it hard to pressure the unions on structural issues as well. The social partners, for their part, had little reason to bargain, since the SPD’s internal splits and opposition control of the Bundesrat meant that there was little prospect of unilateral government action should the Alliance fail to produce suitable proposals.

One thing that made it difficult to reach agreement on measures to tackle unemployment was that there was no commonly accepted diagnosis of the roots of Germany’s employment problems. The dominant view of the SPD’s left wing and the trade unions was that the root cause of unemployment was inadequate demand and that the solution lay in raising public expenditure to close the output gap. Others viewed the
problem as essentially structural. However, there was disagreement on the nature of the structural problems involved. Some in the trade union movement, the SPD and the Ministry of Labour and Social Affairs advanced a “lump-of-labour” argument that focused on the distribution of available work. Other structural analyses focused on the size of the tax wedge on labour (a focus of business lobbies), wage rigidities, the high reservation wages created by the benefits system, and barriers to employment growth in the labour-intensive services sector.

Policy process: formulation

The search for solutions

In early 2002, the government, facing an uphill battle at the coming general election, took advantage of a scandal surrounding the BA to break the impasse in the Alliance for Jobs and press for significant labour-market reforms. The revelation that the BA was spectacularly inflating the statistics on its success in placing unemployed workers in jobs not only discredited the agency but also reflected badly on the labour ministry and the social partners, who were responsible for overseeing it. The need for reform could no longer be denied. In response to the scandal, the government turned the volunteer administrative board of the BA, dominated by the social partners, into a supervisory board with no executive functions and installed a new BA management board, made up of professionals and operating under the same rules on the separation of executive authority and oversight that applied to listed companies. Chancellor Schröder then appointed a commission under Dr. Hartz to develop a vision for the future role and organisation of the BA, including plans to implement that vision.

The 15-member group was selected by Dr. Hartz, working closely with the Chancellor’s office. It included two academic experts (a law professor and a political scientist), two trade unionists and one representative of an employers’ organisation. The remaining members were drawn from management consultancies, company boards and political circles. There was no one from the DGB, the BDA, the BDI or the labour ministry. The composition of the commission thus marked a striking – and deliberate – departure from the tripartism of the Alliance. Also significant was the absence of representatives of either opposition parties or of the SPD’s “traditionalist” wing. Despite its technocratic composition, then, the commission was a highly political instrument. Its purpose was to introduce into the policy debate a range of measures that could substantially reduce unemployment and to do so in a manner that could win public acceptance. In a very real sense, the commission’s role was to run risks for the government, as was evident in June when Dr. Hartz himself revealed its intermediate conclusions to the press in order to test public reaction. The news that the Commission was thinking of reducing the duration of income-related UI benefits provoked strong trade-union opposition, and the proposal was dropped from the final report. Although the commission, with the backing of the Chancellery, interpreted its mandate to reform the BA rather broadly, veto-players inside and outside the commission ensured that certain taboos were observed: it did not tackle such issues as collective bargaining, employment protection or barriers to entry in handicrafts sectors. It focused instead on policy towards the unemployed, an emphasis that, while sometimes controversial, did not provoke the kind of conflicts with vested interests that an attack on the protections of employed “insiders” would have entailed.
The Hartz recommendations

The final report, structured into 13 modules, contained a mix of proposals that reflected the wide range of actors involved and their differing assessments of the causes of unemployment (Table 10.1). Many of the proposals had long been debated in Germany and were already widely accepted by the expert community, including some drawn from the European Employment Strategy and the OECD Jobs Strategy. Others, however, were newer and reflected the perspectives of the managers and consultants on the commission, particularly their interest in engineering change in complex organisations. This was especially apparent in the proposals concerning the future of the BA. A key innovation was the stipulation that there should be a systematic ex post evaluation of all the measures implemented as part of the reform.

Table 10.1. Main recommendations of the Hartz Commission

<table>
<thead>
<tr>
<th>Job creation</th>
<th>Improved placement</th>
<th>Reform of the BA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction of “Personal-Service Agencies” in each employment office district,</td>
<td>1. Creation of job centres throughout the country,</td>
<td>1. Simplification of instruments for promoting employment, and strengthening competition,</td>
</tr>
<tr>
<td>2. promotion of part-time employment in domestic services (“mini-jobs”),</td>
<td>2. simplification of labour law, use of advisory teams, and introduction of notification requirement for dismissals and resignations,</td>
<td>2. merger of unemployment benefits and welfare benefits, and</td>
</tr>
<tr>
<td>3. introduction of self-employment or family employment as a new form of labour supported by tax incentives and grants for up to three years (“Ich-AG”),</td>
<td>3. expansion of availability of child day-care facilities,</td>
<td>3. more effective organisation of workflows and control activities at the BA,</td>
</tr>
<tr>
<td>4. introduction of “Job Floater” low-interest loans to secure SMEs that invest in new jobs and employ previously unemployed individuals, and</td>
<td>4. tightening the conditions under which individuals could refuse job offers,</td>
<td></td>
</tr>
<tr>
<td>5. introduction of company employment “balance sheets” and payment of bonuses from the unemployment insurance system to those with positive employment development.</td>
<td>5. reversal of the burden of proof for cases relating to refusal of job offers and the institution of flexible sanctions,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6. promotion of training and employment for young people, and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. development of a “bridge” system of employment for older persons involving grants to older workers who accept lower-paid jobs or take early retirement.</td>
<td></td>
</tr>
</tbody>
</table>


The commission proposed restructuring the BA so as to focus its work more clearly on job placement rather than just benefit payment. Personal profiling and counselling of job-seekers was to become a core function of BA staff, while training vouchers and other quasi-market measures were to be introduced to improve the efficiency of BA services and prevent the formation of the kind of informal insider relationships that critics believed to exist between the BA and providers of training and other services. A network of new Personnel-Service Agencies (PSAs) were to function as temporary employment agencies, treating the unemployed as placement clients rather than passive benefit recipients. Anyone still unemployed after six months would be employed by a PSA and hired out to firms on a short-term basis, albeit at union-negotiated rates. The PSAs
themselves could remain in the public sector or be established as private enterprises undertaking contract work for the state.

This emphasis on activation was also reflected in proposals to increase the pressure on unemployed persons to seek re-employment. The conditions under which unemployed individuals could refuse a job offer were to be tightened, and the burden of proof would henceforth be on them to justify their refusal. The commission proposed that individuals should face benefit cuts if they refused to accept jobs without adequate grounds for doing so or, in the case of young, single people, to relocate for work. At the same time, the report proposed a number of measures, including new forms of start-up subsidies and wage subsidies of various kinds, to help recipients back into work, along with proposals to encourage various non-standard forms of employment. The Hartz commission also endorsed a benefit-reform proposal that had been under discussion for some time and was already being studied by a commission on the reform of municipal finances. The commission proposed leaving the initial UI benefit unchanged (it did not resurrect the proposal to shorten the duration of UI), while replacing the old UA with a new means-tested flat-rate benefit, which would also cover those SA recipients judged fit to work. This would dramatically reduce the duration of benefits linked to previous income.

Both the presentation and the timing of the Hartz report played a significant role in shaping its initial reception. Released just six weeks before the election, the 340-page document bore little resemblance to the prototype of a government report, in style, layout or structure: the summary read more like a manifesto, ending as it did with an appeal to the “professionals of the nation” to bring forward proposals to reduce unemployment, and it appeared to be aimed at the media and the public rather than at policy elites. Chancellor Schröder endorsed the commission’s recommendations immediately and, with the parliamentary election campaign well under way, both the trade unions and the SPD’s left wing were to some extent compelled to follow his lead, despite the fact that many on the left had misgivings about the proposals. To have done otherwise would have been to damage the SPD’s electoral prospects. In any case, the Commission’s decision to drop its proposal to reduce the duration of UI benefit made it easier for the unions to accept the rest of the package. Business groups were much less enthusiastic, arguing that the proposals had been diluted under union pressure and that they failed to address the really critical issues, such as the tax wedge on labour.

Policy process: adoption

Once re-elected, the government moved rapidly to push legislation based on the Hartz recommendations through parliament. Having won an electoral mandate for reform, SPD leaders were keen to use the immediate post-election period to press ahead with reforms, and the deteriorating economic situation added to the urgency of the issue: growth was not picking up as expected, and unemployment continued to rise, reaching almost 4.5 million in early 2003. Employment fell to its lowest level in a decade. These developments strengthened those within the coalition who favoured reform. A top-level steering group on the implementation of the Hartz recommendations was formed under Mr. Schröder himself and included Dr. Hartz, the head of the Federal Chancellery, and the new Minister of Economics and Labour, Wolfgang Clement. This reflected the importance the government attached to speedy action on the proposals. By merging the Ministry of Labour and Social Affairs, with its largely traditional social policy orientation, into the more liberal Ministry of Economics and Technology, the government put labour-market policy at the heart of its economic agenda while eliminating the labour
ministry as a potential centre of opposition to reform. It also gave Mr. Clement the prominence and authority needed to push the reform forward. In addition, much of the work involved in preparing legislation on the restructuring of UA and SA was undertaken by a working group under the Commission on the Reform of Municipal Finances, which brought together representatives of the municipalities and counties, the Länder, various federal ministries and the social partners.

The decision to package numerous recommendations contained in the report into just four complex bills (Box 10.2) was a further reflection of the government’s desire to streamline the parliamentary process and avoid endless battles over unpopular individual measures. Although the government was re-elected on a pledge to implement all the Hartz recommendations, many of them were dropped or substantially modified as government officials prepared the bills. The most significant—and controversial—departure from the Hartz recommendations concerned unemployment benefits: although the commission had dropped the idea, the government went ahead with a proposal to shorten the duration of income-linked UI benefit. While there was broad agreement on the transfer of all employable benefit recipients (i.e., those on UA and SA) onto a new flat-rate benefit, the reduction in the duration of UI was fiercely contested. The reform aimed to stimulate the reintegration of those benefit recipients best equipped to re-enter the workforce by lowering their de facto reservation wages. However, critics charged that it would lead to increased poverty, particularly among children, and the public largely accepted this, viewing Hartz IV as an attempt to cut benefits for vulnerable groups during a severe recession.

A further element of the Hartz IV reform, largely overlooked by the media and the public, concerned the reduction in benefit withdrawal rates for the long-term unemployed: previously, these had been 100% across a wide band of earned income, meaning that recipients who took up any employment forfeited their benefits. The introduction of phased withdrawal rates reduced the risk that employment would be financially unattractive.

In fact, although it was clear that more benefit recipients would experience a reduction in transfers than an increase, the distributional consequences of the reform were complex and not entirely predictable. The new proposal offered some advantages to recipients of SA, who would henceforth gain access to labour-market services and would also be included in the pension system. Moreover, the new benefit would be available to some who were unable to access SA, owing to its more stringent means test. The benefit formula implied increased benefits for many poor families, especially single-parent families. However, it hit other groups hard, particularly single people with no dependents and those who had previously had relatively high earnings. Older workers stood to lose the most, as they tended to have higher past earnings, fewer dependents, longer contribution histories and longer periods of eligibility for UB I. In effect, Hartz IV threatened a widely employed means of transitioning older workers into early retirement.

The four bills were passed within two years of the report’s appearance. The first two laws, which addressed the least controversial elements of the Hartz proposals, were speedily adopted in the autumn of 2002, reflecting the strong reform momentum that existed in the wake of the BA scandal and the government’s election victory. The third and fourth bills became centrepieces of the government’s broader “Agenda 2010” economic reform strategy, unveiled in March 2003. However, Agenda 2010 included measures that went beyond the Hartz recommendations, and it provoked considerable opposition from the unions and the SPD left. At the same time, the government was experiencing a rapid decline in its popularity, a drop that initially appears to have owed
more to the economic situation than to the reform proposals but that nevertheless made it harder to press ahead with the Hartz legislation. Over time, moreover, opposition to the government came to focus on elements of Agenda 2010, including Hartz IV in particular. Passage of Hartz III was nevertheless relatively smooth, since it did not require the approval of the opposition-controlled Bundesrat and because attention was focused on the much more controversial Hartz IV benefit reform.

Box 10.2. The “Hartz legislation”

Four new federal laws were adopted in 2002-03 pursuant to the Hartz Commission report:

1. Hartz I (December 2002) encompassed: the establishment of so-called “Personnel-Service Agencies” all over Germany; reform of the Law on Temporary Employment and Labour Leasing; the requirement of early registration even for impending unemployment; stricter rules for taking up “reasonable” employment; transfer of the burden of proof concerning rejected job offers onto the job-seeker; and the introduction of training vouchers.

2. Hartz II (December 2002) included: the provision of new benefits for business start-ups (Ich-AG); reform of the minor employment relationships (“mini-job”) legislation; promotion of household services; preparations for the organisation of Job Centres combining employment and social welfare services; and reorganisation of the BA with a view to making it a more effective service provider.

3. Hartz III (October 2003) continued the reform of the BA to focus its activities more clearly on reintegration of the unemployed and introduced some changes in the rules governing partial retirement. The most controversial element of this bill was the provision allowing the BA to cut benefits by up to 30% if recipients of unemployment benefit refused a job offer without adequate grounds for doing so.

4. Hartz IV (July 2004) restructured unemployment benefits. Entitlement to the old income-related UI benefit (henceforth to be known as unemployment benefit I or UB I) was limited to a maximum of 12 months for those under 55 and to 18 months for older workers. Thereafter, the unemployed were eligible for the new flat-rate unemployment benefit II (UB II). Hartz IV also further tightened the definition of a “reasonable” job. (The Hartz IV law itself did not curtail the duration of UB I, and, as noted above, the Hartz Commission had shied away from making such a recommendation. However, this measure was adopted at the same time as Hartz IV, in a separate Labour Market Reform Act, and it thus came to be identified with Hartz IV in public debates on the issue.)

1. The legislation raised the limit for low-paid work exempt from social security contributions and introduced a rising scale of contributions for incomes of EUR 400 to EUR 800 a month.

2. Henceforth re-named the Bundesagentur für Arbeit or Federal Labour Agency.

All four bills were modified in the course of their adoption. In respect of Hartz I, the most controversial issue was the liberalisation of temporary agency work that the creation of PSAs would entail. The unions tended to see any relaxation of the existing very severe restrictions on this form of employment as a threat to their core membership, but they had otherwise played little role in the regulation of temporary agency work, largely because the workers involved were not organised and because there was little scope for co-determination involving temps. The government won union acceptance of the removal of many restrictions on the use of temporary work by agreeing to the introduction of collective bargaining between the social partners in this sector for the first
time. Perhaps the most important change in Hartz II was the extension, at the insistence of the Bundesrat, of the provisions for so-called “mini-jobs”: the earnings threshold for these forms of marginal employment was increased and they were not confined to domestic services as originally proposed.

A number of concessions were made to the SPD’s left wing and the trade unions when Hartz III and IV made their way through the Bundestag: these included some weakening of the work availability requirements in the bills, particularly for older workers, and making the means test for UB II less stringent. Some of the reforms to the BA itself also met resistance in the Bundestag and from both trade unions and employer organisations, at least partly owing to fear that reform would lead to reduced funding for their own further education programmes. However, some of the most significant concessions in respect of Hartz IV were reversed at the insistence of the centre-right Bundesrat majority when the mediation committee resolved differences between the two chambers. On this and other matters, it became clear that split-party control over the federal parliament might actually help supporters of reform in the government, by limiting the scope for concessions to the SPD left or the unions.

Though highly controversial, the cut in the duration of UB I did not trigger any large-scale mobilisation during its parliamentary phase. The unions opposed the benefit reform, which threatened to exert downward pressure on wages, but they were focused primarily on resisting proposed changes to employment protection legislation and the system of collective bargaining that had been advanced as part of Agenda 2010. In fact, the most difficult issues to resolve during the passage of Hartz IV concerned the allocation of financial and administrative responsibilities across levels of government. Unemployment assistance had been administered by the BA, while SA was handled by the municipalities. The government favoured federal administration of the new UB II in its entirety, but the centre-right parties dominating the Bundesrat argued for administering it at municipal level. The associations of municipalities and counties themselves could not agree a common position, because different types of local authority tended to face different labour-market conditions. After months of negotiation, a compromise was reached, involving a rather complex division of administrative responsibilities between the BA and local social welfare offices. These issues were ultimately regulated in a separate Municipal Option Act. Some 69 municipalities were exempted from this and allowed to go it alone.

The government’s communications strategy

The initial presentation of the Hartz report was itself a major – and, for Germany, unusual – public relations exercise. Both the language of the report and the manner in which Dr. Hartz publicised it were without precedent for an official commission, and they succeeded in generating widespread and broadly favourable press coverage. This made it easier for the government to move ahead quickly with the first pair of bills. However, as the reform process unfolded, there were two lapses in communication, both of which damaged the government’s public standing.

- Little was done to explain the controversial Hartz IV reform either to the public at large or to those who would be directly affected by it. This omission contributed directly to the wave of protests that erupted when Hartz IV was to be implemented. Significantly, widespread demonstrations against Hartz IV began not when it was being passed but after the government circulated a 16-page questionnaire to existing benefit recipients to assess their eligibility under the new
rules. The questionnaire frightened and confused many, and gave rise to rumours and misconceptions about what the reform would entail. Serious work on a communications campaign for the reform began only after the demonstrations had started.

- Hartz IV effectively changed the measure of unemployment, because it included several hundred thousand former SA recipients in the statistics of the BA. In November 2004, the BA warned that this might push the unadjusted figure for registered unemployed above 5 million, a politically sensitive threshold. However, little effort was made to explain this statistical effect in advance to either the press or the public. As a result, when the headline figure breached the 5 million mark in February 2005, the government’s public standing suffered. The psychological impact of the figure was widely reckoned to have contributed to the SPD’s electoral setbacks in a number of state elections in early 2005.

**Policy process: implementation**

The BA was directly responsible for implementing most of the measures contained in Hartz I-III and related legislation. Other bodies played a largely subsidiary role. Despite the sometimes very short lead times, implementation of most of these measures was relatively smooth. This reflects at least to some extent the progress achieved in overhauling the management of the agency, which did indeed reorient itself substantially towards placement rather than passive assistance. The introduction of training vouchers and better screening and assignment of individuals to programmes led to a marked shift in training patterns and outcomes. There was also clear evidence that the tougher work availability requirements were being applied, though there was considerable variation in the strictness with which they were enforced by different local labour offices. However, other elements of the package ran into serious difficulties. Given the wide range of different measures encompassed in the legislation and the speed with which they were implemented, this is hardly surprising. Staffing alone emerged as a significant issue for the first few years, as the new tasks assigned to the BA initially resulted in extremely high caseloads for placement counsellors. One of the most important elements to be implemented was the requirement for systematic ex post evaluation of the reforms: several dozen leading research institutes were involved in the assessment of the reforms, providing the basis for policy makers to consider where to extend, and where perhaps to curtail, the measures implemented.

The PSAs got off to a very difficult start and in most places did not emerge as particularly important actors on the labour market. By 2006, many labour offices were reducing their reliance on PSAs. The real impact of the legislation concerning PSAs stemmed from the liberalisation of temporary agency work that it entailed. Despite their qualms about this measure, DGB-affiliated unions quickly concluded collective agreements with the two major employers’ organisations in the sector – the first corporatist interest representation in temporary agency work in Germany. This was an important implementation achievement, since, in the absence of a collective agreement, the law would have required equal treatment of temporary workers from the first day of work, effectively pricing many of them out of the market. The result was very rapid growth in temporary agency work; its share in total employment, though small, jumped by close to 70% during 2003-06. Thus, while the liberalisation of temporary agency work was initiated outside the system of corporatist bargaining, implementation of the reform depended on the cooperation of the social partners.
Hartz IV posed the greatest implementation challenge. Technically, the benefit reform was unusually complex, requiring the creation of new joint agencies, the drafting of detailed regulations governing their operations, the development of new software and the training of staff to run the new system. All this had to be done in roughly six months in order to introduce the new benefit on time. Although there were considerable problems, benefit payments continued without disruption. Strikingly, the anti-Hartz protests that had begun in the summer of 2004 quickly subsided when the new regime was implemented, in part because its impact was less draconian than many had feared. While around half of UA or SA recipients did experience a decline in benefits, perhaps one-third received higher benefits after the reform.

The elimination of the open-ended income-linked UA meant that the overall generosity of the system, as measured by the OECD summary measure of benefit entitlements, fell by more than in most OECD members during 2003-05. Most of the losers were former UA recipients in the East; the former SA recipients, who gained by moving onto UB II, were largely in the West, as were most of the large number of new claimants, who had previously been eligible for neither UA nor SA. This appears to be one reason why the SPD vote in 2005 fell so much more in East Berlin and the eastern Länder than in the West. While most former UA recipients did lose income, the reform raised the incomes of the poorest. Moreover, some who lost benefits ended up better off, because they found employment rather than remaining on benefit. The change was most striking in respect of older men, whose labour-force participation rates rose by an estimated 10 percentage points during 2005-06.

Sustaining the reforms proved difficult. The improvement in fiscal and labour-market performance after 2005 reduced the sense of urgency about reform, while the rise of the new Left Party on a strongly anti-Hartz platform and intensifying electoral competition increased the incentives to restore benefits and relax some of the more rigorous elements of the reform. The most significant change to date has been the decision to extend older workers’ eligibility for the income-related UB I to up to 24 months. Moreover, the new benefit system faces still further revision following a December 2007 Constitutional Court ruling that the administrative and financing arrangements for UB II violated provisions of the Basic Law concerning fiscal federalism. Critics argue that the system needs more than minor legal adjustments, since it also appeared to have encouraged the municipalities to declare as many SA recipients as possible to be fit for work in order to shift the burden of supporting them onto the BA. There is also growing pressure for the establishment of further sectoral minimum wages, following the adoption of a statutory minimum in the postal sector at the end of 2007. While the postal minimum wage reflected the interests of Deutsche Post as well as the sector’s unions, the broader debate on minimum wages – which Germany has never had before – seems to reflect in part the drop in the implicit minimum wage as a result of the benefit reforms. That said, the backlash also reflected poor communication of the impact of the reforms: the benefit cuts were not as harsh, nor the activation measures as strict, as they have been perceived, but there remains a widespread impression that the Hartz reforms were exceptionally severe. Moreover, they do appear to have had an impact on unemployment: OECD (2008) estimates that Hartz IV alone, by reducing average replacement rates for the long-term unemployed, could reduce the structural unemployment rate by around 0.5 percentage points. However, some of the resentment generated by the reforms seems to have resulted from that very success: the rise in participation rates among older workers included many who did not welcome longer careers.
Conclusion

The challenge of labour-market reform in Germany during the early 2000s was largely defined by the institutional environment. The number of potential “veto players” involved in the policy process was enormous, given the heterogeneity of German political parties, divided control of the federal parliament, the complexity of German federalism and the strong tradition of according the social partners a key role in policy making. Moreover, the strength of attachment to the basic elements of Germany’s welfare state meant that opponents of reform could often gain a more sympathetic hearing from the general public than did reform advocates. For most of its first term of office, the SPD-Green coalition proved unable or unwilling to overcome these constraints on reform, pursuing more or less “traditional” labour-market policies while exploring incremental change via established corporatist channels. What made it possible for the reformers within the governing coalition to break free of these constraints was in some respects the eruption of a “perfect storm” in 2002: sharply rising unemployment increased the pressure on the government to be seen to be taking action to combat joblessness, the scandal at the BA put those responsible for the labour-market status quo on the defensive, and the approach of elections made it harder for the trade unions or SPD traditionalists to challenge government policy in public.89

In these circumstances, the Chancellor and his supporters used the Hartz Commission as a vehicle to bypass the traditional veto players in the early stages of the policy process and then moved to legislate rapidly in the immediate aftermath of their election victory. While that victory appears to have been won on issues other than labour-market policy, the Chancellor’s strong pre-election endorsement of the Hartz recommendations allowed him to claim a mandate to implement them and initially made it difficult for the unions or the SPD left to oppose him. Moreover, the reforms were sequenced so as to address the easiest issues first, leaving the tougher reforms until a later stage. They were also focused very much on activation of the unemployed and thus stopped short of changes to employment protection, the system of collective bargaining or other issues that would have generated stronger opposition from the trade unions and other organised interests.

The perfect storm did not last, however, nor did it lead to a fundamental reorientation of the various interests involved in labour-market policy. This largely explains why the reforms have remained so unpopular and, in the case of Hartz IV, so difficult to sustain. As labour-market conditions improved and politics returned to “normal”, the sense of urgency about reform abated, and there was growing pressure on the government to restore some benefit cuts and “share the fruits of the recovery”. Moreover, the key figures who had driven the Hartz reforms, including Mr. Schroeder and Mr. Clement, largely passed from the political scene, while the emergence of the Left Party added to the pressures for the SPD to back away from the reforms. The failure to communicate the meaning of the reform more clearly when it was being adopted also played a role in making it harder to sustain, as did the failure to explain more effectively why and how the reforms were working or to use the period of stronger growth that followed the adoption of Hartz as an opportunity to build support for deeper reforms.
Notes

1. The other two had to be submitted to the Bundesrat for comment only, not for approval.


4. See the comparative polling data presented in Padgett (2005).

5. In a system with more complex coalition bargaining, government participation might have been less dependent on electoral success; this would have reduced the risks associated with reform.

6. Discussion of unemployment in Germany has traditionally focused on the number of registered unemployed rather than on the unemployment rate as estimated on the ILO methodology; on the differences between the two, see OECD (2008:71-72).


8. In 2002, overall labour force participation was about 2% above the OECD average, but among the 60+ age group, it was about half the OECD average for women and about two-thirds of the average for men. See OECD (2004a) for details.


10. In 2001, an SPD working group highlighted that party’s divisions on labour-market reform, which grew more acute as the employment situation deteriorated. See Clasen and Clegg (2004). The CDU-CSU, for its part, had a significant “social” wing with links to historically Catholic trade unions that was often at odds with its more pro-business wing. Cf Wood (2001:392).


12. These were: the German Association of Cities (Deutsche Städtetag), the German County Association (Deutsche Städtetag Landkreistag) and the Association of Cities and Municipalities within Counties (Deutsche Städtetag Städte- un Gemeindebund).

13. Streeck and Trampusch (2005) underscore the employment opportunities that control of the BA offers the social partners.

14. Employer and trade union representatives each constituted one-third of the BA’s administrative board, with federal Land and municipal governments sharing the remaining third of the seats.

15. In 2002, about 23% of dependent employees were union members, down from 36% in 1991. The average for Continental Europe was 32% and for the OECD as a whole, just under 34%.
16. There were numerous former trade union officials in the leadership of the SPD, as well as institutionalised channels for regular consultation between the party and the unions; on the centre-right, the Christlich-Demokratische Arbeitnehmerschaft provided an institutionalised voice for organised labour within the CDU.

17. For the unions, this was sometimes especially important in the eastern Länder.

18. For a first-hand account of the commission’s work, see Schmid (2004).

19. A typical West German family with three children on SA in 2000 had an income just 15% below that of an identical family relying on a single breadwinner earning an average unskilled worker’s wage; in East Germany, the difference was smaller still. Jacobi and Kluve (2006:6).


21. OECD (2003). Saint-Paul (1996) goes further, citing evidence that ALMPs were structured primarily to serve the needs of labour-market insiders facing short spells of unemployment and that they therefore made it harder for outsiders (chiefly the long-term unemployed) to compete on the labour market.

22. There was only limited support for wage or start-up subsidies and job-search assistance.

23. The number of registered unemployed is somewhat higher than the number unemployed on ILO definitions, which forms the basis for the reported unemployment rate. The former number was more politically salient during the period under study. This should also be borne in mind when interpreting data on long-term unemployment: roughly half the total on ILO definitions had been unemployed for a year or more but this was true of only about 33-35% of registered unemployed.


25. Job-search requirements were relaxed and employment protection in small firms strengthened. See Busch and Manow (2001); Streeck (2003); and Streeck and Trampusch (2005).

26. Silvia (2002) finds that the fall in unemployment at this time was chiefly the result of strong export demand and favourable demographic trends and that government policies actually made reducing unemployment more difficult.

27. For details, see OECD (2003:104).

28. Büchs and Friedrichs (2005) and Clasen (2005) find that the Strategy did not so much shape German policy as provide a useful tool for policy makers trying to reorient the SPD away from its traditional approach to labour-market policy.

29. On the limited achievements of the Alliance, see Streeck (2003); Vitols (2004); and Clasen (2005).


33. Steiner (2004) suggests that the cyclical component of German unemployment in 2003 was as little as 20% of the total.
34. This view underlay calls for shorter hours, early retirement, job-rotation and similar measures.

35. An investigation of the BA’s activities carried out by the Federal Audit Court revealed that 70% of reported job placements had been recorded incorrectly. About a third of the job placements listed in the official statistics were controversial because of different interpretations of the term “job placement”, and at least 36% had simply been falsely recorded.


38. From the North Rhine-Westphalia branch of IG Metall and the services union ver.di.

39. The leak, which most observers believed to have been authorised by the Chancellery, led to considerable conflict within the commission and prompted its two academic members to write an open letter complaining about political pressure from outside; see Rehfeldt (2002); and Schmid (2004).

40. The matter was reportedly resolved in talks at the Chancellery on 5 July, involving Chancellor Schröder, Dr. Hartz, Labour Minister Walter Riester and key union leaders; Rehfeldt (2002:138).

41. The sole employers’ organisation represented on the commission represented employers in the handicrafts sectors.

42. Special arrangements were proposed for older unemployed persons, who were often particularly difficult to re-employ.

43. Both start-up and wage subsidies existed prior to Hartz, most notably the so-called Überbrückungsgeld and various integration subsidies.

44. For more detail on the old and new benefit systems, see OECD (2004b).


46. Thus, the DGB promptly expressed support for the Commission’s conclusions, as did IG Metall. Both unions were later very critical of aspects of the Hartz legislation. On the electoral pressure to support the Hartz proposals, see Rehfeldt (2002); Clasen and Clegg (2004); and Dyson (2005).

47. Rehfeldt (2002:138) suggests that during the 5 July talks, the unions may have agreed to accept the Commission’s other proposals if the curtailment of unemployment benefit was dropped.

48. For union and business lobby reactions, see BDA (2002); and Kaiser (2002).


50. Buhr and Schmid (2007). The merger also simplified both the drafting process and the committee stage in parliament – since the Bundestag’s standing committee structure parallels that of the federal executive, the number of committees involved was reduced.
For example, the “bridge” system for older workers was quickly dropped, as it ran counter to the broader policy of trying to promote longer careers. The “job floater” subsidies were found to involve excessive deadweight losses, and the proposals for paying “bonuses” from the unemployment insurance system to “employment-friendly” companies was never given serious consideration within the ministry. Neither was the proposal for “family-friendly quick placement” of breadwinners.

The Joint Welfare Alliance claimed that a new unemployment benefit at the level of SA would lead to more poverty and would increase the number of children and young people living on benefits at the level of SA by 1.0-1.5 million. See Dribbusch (2004).

See OECD (2008) for details.

Recipients of the new benefit would be covered both by statutory health insurance and by the public pension scheme, to which payments would be made on their behalf.

The implications of the means test for UB II were complex: it was more stringent than that for UA but less strict than that for SA. As a result, some former recipients of UA were disqualified, while others who had not previously been eligible for either UA or SA were able to claim the new UB II.

For an analysis of the distributional consequences of the new benefit, see Eichhorst et al. (2006); and Clauss and Schnabel (2008).

An older worker could, by agreement with his employer, accept redundancy in his late 50s, claiming some compensation from the employer as well as unemployment benefit worth up to 67% of his previous wage for as much as 32 months; he would then accept the somewhat lower, but still income-related, UA for an indefinite period, until he could claim his pension.

Hartz I and II were presented as separate bills because the former consisted of measures that did not require Bundesrat assent, whereas the latter was made up of those that did.

By early 2003, polls showed support for the Chancellor down around 15 percentage points since the election, and the coalition was suffering humiliating defeats in a series of Länder elections, resulting in a CDU-CSU majority in the Bundesrat from February of that year.

Support for reductions in unemployment benefits and social assistance hovered around 35-40% while the reform was being adopted, with support for the “overall reform course” at about 35%. See Padgett (2005), Table 4.

This included the requirement that “standard regional wages” must be accepted even if they are below union wages.

Vitols (2004:384). The prohibition of temporary work in the construction industry was lifted, as were limits on the duration of temporary assignments and the re-employment of temps by the same agency. The ban on “synchronisation” was dropped: previously, agencies could not employ temps only when they were able to lease them; temps were regular permanent employees of temping agencies.

Union density among such workers was under 5%.


67. The opposition parties had presented bills aimed at giving companies more freedom to depart from collectively agreed standards. Since the government also favoured more flexibility in bargaining, the unions feared that some legislative compromise might be adopted. See Dribbusch (2004).

68. For example, prior to reform, SA recipients tended to be more numerous in western cities, while UA was more common in the East. This pattern was even replicated in Berlin, with SA more common in the West and UA in the East.

69. These concerned, in particular, the severity of the means test and the availability of supplementary payments for things like housing and heating. See EIRO (2004).

70. In Germany, political and media attention has traditionally focused on the number of registered unemployed rather than on survey-based measures of unemployment.

71. See the interview with Chancellor Schröder in Die Zeit, 31 March 2005, for a discussion of this issue.


73. See, e.g., OECD (2004a:83); Jacobi and Kluve (2006); Eichhorst and Zimmermann (2007) and Rinne et al. (2008).


75. OECD (2006:102-03); Eichhorst et al. (2006).


77. Tergeist and Grubb (2006); OECD (2006); Eichhorst and Zimmermann (2007); Jahn and Ochel (2007).

78. The lowest wage levels in the agreements were below those of all sectors covered by collective agreements except hotels, restaurants and agriculture. This was largely because the DGB faced competition from a small union, the Christlichen Gewerkschaft Zeitarbeit PSA, which was keen to sign a collective agreement for the sector and ready to agree relatively low wages. See Vitols (2004).

79. Excluding mini-jobs and other forms of employment not subject to social security charges.


82. As noted above, key elements of this reform were embodied in the Labour Market Reform Act.


84. The appearance of an unexpectedly large number of new claimants led to substantial cost overruns; Tergeist and Grubb (2006:16).

85. See King (2005) and Helms (2007).


87. Under the December 2007 agreement, the duration of UB I was extended from 12 to 15 months for workers in the 50-54 age group, from 15 to 18 months for those 55-57, and from 18 to 24 months for those 58 and over. Burda (2007) cites polling data showing 80% support for this change. The impact of the change may be offset in part by a new voucher scheme that will significantly subsidise employers’ wage costs when employing older workers in certain circumstances. See OECD (2008:82).

88. It has been criticised as an attempt to restrict entry as the postal sector is opened up.

89. In this, the Schröder government ultimately followed the pattern of its predecessor: after attempting to operate through traditional tripartitism, it abandoned this approach and pursued reforms unilaterally.

90. See King (2002); and Pulzer (2003).
Bibliography


Tergeist, P. and D. Grubb (2006), “Activation Strategies and the Performance of Employment Services in Germany, the Netherlands and the United Kingdom”, *OECD...*


Chapter 11

Italy: The Treu (1997) and Biagi (2002) reforms

Summary and main conclusions

Labour-market reform in Italy has been an incremental process over more than a decade, but the two principal landmarks have been the Treu and Biagi reforms of 1997 and 2003, respectively:

- The September 1996 “Treu package”, named for Labour and Social Protection Minister Tiziano Treu, aimed to increase employment, particularly among the young, with special provisions for the economically depressed south (the Mezzogiorno). It eased regulation of new apprenticeships and work-training contracts, and created incentives for on-the-job training, temporary work via private agencies and intra-regional labour mobility. It also legalised worker-dispatching services for the first time and reduced disincentives to the use of fixed-term contracts.

- The 2003 “Biagi law”, named after labour lawyer Marco Biagi, who drafted the white paper on which it was based, was the most controversial labour-market reform adopted under the centre-right government of Silvio Berlusconi that held office from 2001 to 2006. It aimed to take some of the Treu reforms further, in order to increase employment among youth, women, older workers and job-seekers, particularly in the Mezzogiorno. The new measures included allowing private employment agencies to compete in the full range of services with public ones; promoting apprenticeships; improving conditions for the use by firms of part-time work; and offering greater opportunities to use other non-standard forms of employment.1

If previous labour-market policies tended to reflect an overriding concern with the protection of employed “insiders”, these reforms sought to create greater opportunities for new entrants and other labour-market “outsiders”. To a greater or lesser extent, both these reforms (as well as a number of other changes to labour legislation adopted in between) were driven by EU-related policy considerations. The drive to get Italy ready for Economic and Monetary Union (EMU) formed the background to the Treu reform, which was also partly a response to a European Court of Justice decision against Italy’s restrictions on worker-dispatching services. The Biagi reform was based in part on the European Employment Strategy and also involved changes dictated by European law. In other respects, however, the politics behind the two reforms were very different: the centre-left cabinets of the late 1990s and the Berlusconi government of the early 2000s adopted contrasting strategies for securing adoption and implementation of the reforms.

A number of factors have facilitated the progress of reform:
Political change and economic pressure. The transition to the second republic in the early 1990s created important openings for economic reform in Italy, triggering not only a re-configuration of the party system but also the entry into politics of a large number of new actors, including economists, labour lawyers and other experts, many of whom shared a common understanding of Italy’s labour-market problems. Their arrival on the political scene, moreover, coincided with a deterioration of domestic labour-market conditions that created a greater sense of urgency about reform.

Increasing external pressures. European integration, in particular, generated considerable pressure for reform. This pressure made itself felt via a range of channels, including the Maastricht criteria for Economic and Monetary Union, the direct effects of European Union rules and policies, international competitive pressures, and the learning that occurred as Italian firms came into increasing contact with the labour markets and regulatory regimes of other countries.

An effective consultation and communication process. The Treu reform was largely the result of an extended and very complex negotiation involving the government, employers’ organisations and the trade unions. Such trilateral concertation underlay a number of Italian reforms in the 1990s. The Biagi reform was only partly “concertative”, but it, too, involved extensive consultation and negotiation with the social partners. The preparation of a white paper prior to the Biagi reform – an innovation in Italian policy making – helped shape the public debate on labour-market issues.

Propitiation of vested interests. In both instances, the reformers secured adoption of proposals to open up labour market entry only after they had accepted that there would be no relaxation of employment protection for permanent employees. They also made “side payments” to key actors in the form of subsidies to the Mezzogiorno and legislative concessions on other issues, such as taxation and working time. This enabled the bulk of the reform proposals to go through, but the fact that liberalisation of entry was possible only if liberalisation of exit was abandoned did risk reinforcing the dualism of the Italian labour market. The history of the Treu and Biagi reforms thus highlights the limits of concertation, as well as its potential.

While the Treu reform was agreed with all the social partners, the adoption of the Biagi law followed a long confrontation between the government and the unions, and the largest union confederation, the CGIL, never accepted the reform. The government’s strategy of “partial concertation” made it possible to adopt measures that the CGIL would otherwise have blocked, but it made implementation harder, since many of the Biagi law’s provisions require union involvement in order to become effective.

The context for reform

Framework conditions for reform

Both the Treu and Biagi reforms were adopted against the backdrop of Italy’s evolution towards a bipolar model of political competition, in which national elections are contested chiefly between two large blocs, one centre-left and the other centre-right. Competition between the two blocs was intense: they alternated in government at
successive elections, with nation-wide margins of victory of 2% or less. Such close bipolar competition can increase the short-term political risks facing a government that attempts to implement reforms, even if its methods and goals differ little from those of the rival bloc, because a shift in the voting preferences of a very small number of people can mean the difference between winning and losing office. These risks were all the greater in the late 1990s and early 2000s, because Italy’s two major political “camps” had yet to coalesce into stable alliances: government coalitions tended to be fractious and difficult to manage.

The history of labour-market policy in Italy highlights these risks. The differences between centre-right and centre-left on labour-market issues are not dramatic, but each camp has an incentive to emphasise those differences that do exist for political gain. This is easily done when highly sensitive questions of “acquired rights” are involved. It is therefore fairly easy to put a reforming government on the defensive. However, the political upheavals of the early 1990s also created important openings for economic reform. The discrediting of the old professional political class facilitated the entry into national politics of a number of economists, labour lawyers and other specialists, who were recruited into the technocratic governments of 1992-95 and subsequently remained active in national politics. Their arrival on the political scene, moreover, coincided with a deterioration of domestic labour-market conditions and a tightening of external constraints that created conditions favourable to reform.

As noted in the case study of the 1994 pension reform, Italy in the 1990s saw a great deal of policy-making characterised by the practice of concertation, involving the government, the major employers’ organisations and the trade unions. Beginning with the 1992 agreement on incomes policy, successive governments elaborated increasingly precise and formalised rules governing concertative policy processes in fields like pension and labour-market policies. The Treu reform was one of the major outcomes of this process. However, the extent to which concertation involved real negotiation – rather than consultation – with social partners varied considerably, and the centre-right government elected in 2001 initially tried to challenge this model of policy making, not least in its pursuit of labour-market reform. The Berlusconi cabinet’s efforts to free itself from the constraints of concertative policy making and to legitimate its policies on an electoral rather than a corporatist basis were one of the reasons for the conflict surrounding the Biagi law. In the end, however, the law was adopted following a “partial concertation” process that resulted in a tripartite pact involving the government, employers and two of the three main union confederations.

Coalition politics

Though it had won the April 1996 election, the centre-left government of Romano Prodi did not command a majority in either chamber of parliament. In the Senate, where it was just two seats short, it could rely on ten life-senators, but in the Chamber of Deputies, it needed the support of the hard-left Rifondazione Comunista (RC). However, the right was in disarray following its electoral defeat, and no other national elections were scheduled before June 1999. The Prodi government thus had a brief opportunity to govern without any very immediate electoral pressures. Moreover, the popularity of the government’s main policy goal – securing Italy’s entry into the first wave of EMU – gave it a degree of strength and cohesion greater than its minority status might have suggested. Nevertheless, its first months in office had been difficult, as the RC used its position to exercise near-veto power over the government’s programme. Securing the RC’s support for labour-market reform was therefore a key challenge for the government.
By contrast, the centre-right Casa delle Libertà coalition that took office under Silvio Berlusconi in 2001 commanded comfortable majorities in both chambers and survived a full parliamentary term. Like the Prodi government in 1996, the Berlusconi cabinet initially benefited from both an extended period free of major elections (there were none due until mid-2003) and an opposition still trying to cope with the fall-out from its electoral defeat. Nevertheless, from the start there were conflicts within the centre-right coalition over a range of issues, including spending caps and judicial reform. During the labour-market reform battles of 2002, the government’s unity was severely tested and its internal divisions sometimes broke into the open.

**Macroeconomic situation**

Both the Treu and the Biagi reforms were undertaken in the context of slow economic growth and high unemployment. While poor labour-market performance may have strengthened the case for labour-market reform overall, it made the relaxation of employment protection legislation (EPL) politically more difficult. The value of such protection to incumbent workers would have risen in a period of high unemployment. Yet while cyclical factors may have added to resistance to reform, longer-term structural changes were adding to the pressure for change. In Italy as elsewhere, increasing international economic integration – and, in particular, the opening up of many markets necessitated by the creation of the European single market – was reducing product-market rents. The disappearance of the rents arising from past formal and informal protection reduced the ability of producers to carry on living with the costs of existing labour-market policies and institutions. Italy was not by any means unique in experiencing such pressures, but its exceptionally severe labour-market problems made both the pressure for reform and the resistance to it particularly acute.

**Actors and interests**

Apart from the government and the opposition, the principal actors in the labour-market policy debates of 1997-2003 were the social partners – the employers’ organisations led by Confindustria and the three big trade union confederations: the Confederazione Generale Italiana dei Lavoro (CGIL), the Confederazione Italiana dei Sindacati dei Lavoratori (CISL), the Unione Italiana del Lavoro (UIL). In the end, both reform laws were adopted pursuant to “social pacts” between the government and the social partners. However, there were differences in the roles played by the social partners in the two episodes:

- Though involved in the negotiations leading to the Treu reform, Confindustria did not play a particularly active role. At that stage, the organisation was dominated by Italy’s largest companies and was chiefly concerned with wage restraint and the implementation of the commitments on incomes policy agreed in the social pacts of 1992-93. It was therefore broadly happy with the way concertation was working and eager to take part. Although it regarded the Treu reform as insufficient in scope, Confindustria welcomed its major provisions and signed the Pact for Labour. In the early 2000s, by contrast, Confindustria’s stance was both more visible and more confrontational. The organisation pushed hard and publicly for faster and more extensive reforms. Labour-market reform was particularly important for Confindustria at this time, since its leader, Antonio D’Amato, had been elected with considerable support from medium-sized firms, many of which were strongly in favour of greater labour-market flexibility.
• Elements of the Biagi proposals brought out differences among the three confederations. In particular, the emphasis on decentralised negotiations on many issues created tensions between the CGIL, with its traditional commitment to national bargaining, and the CISL, which has tended to be much more open to local and firm-level bargaining. The CISL was also keen to secure a role in shaping training policies, another issue in play during the discussion of the Biagi reform. In short, there were issues on which both the CISL and the smaller UIL could reach agreement with the government, whereas the left-wing CGIL disagreed with them and was in any case never likely to accept any agreement with the Berlusconi government. Yet it was not easy for the UIL and the CISL to cooperate with the government, as they came under strong pressure from their own members to join the CGIL in resisting any change to the dismissal protection provisions of the 1970 Workers’ Statute, which emerged as a central focus of conflict.

On one major issue, elements of the labour and business lobbies approached the question of reform with a shared interest: a desire to preserve the sectoral-corporatist system of wage-replacement benefits that had evolved in large-scale industry during the post-war period. This was one of the obstacles to a broader reform of unemployment benefits.

**Issue history**

**Italy’s labour market problems**

Italy had long stood out among OECD and EU members for its exceptionally low participation and employment rates – both of which were, moreover, falling steadily during the economic turmoil of the early 1990s. Labour force participation rates for working-age Italians, at 57-58% in the mid-1990s, were about ten percentage points below the levels typical of the European Union as a whole, and the country’s unemployment rate was consistently above the EU average (Table 11.1). In fact, actual employment levels were probably somewhat higher, owing to the size of Italy’s black economy, estimated by Istat at 20-25% of GDP. However, there was no shortage of other indicators suggesting that the labour market was functioning very poorly. By the mid 1990s, youth unemployment had surpassed 32%, and the share of long-term unemployment reached 63.2% in 1995, the highest in Europe.10

The composition of employment in Italy also differed markedly from what was found in most other developed economies. Historically, “regular dependent employment” has accounted for a relatively low share of total employment (around 57% in 1997). By contrast, Italy records unusually high levels of self-employment – around 32% at the time of the Treu reform, as against an EU average of 17%.11 Many forms of non-standard employment, including temporary agency work and intermittent work, were all but unknown prior to the 1990s. Yet demand for such forms of employment was growing, driven by increasing internationalisation, technological change and changes in the way many businesses operated. Much of this demand for non-standard employment was being met in the underground economy or via the circumvention of existing regulations in the formal sector. One of the aims of both reforms was to formalise and regulate underlying – but hidden – trends in the labour market.
### Table 11.1. Labour-market indicators, 1995, 2000

<table>
<thead>
<tr>
<th></th>
<th>Employment rate</th>
<th>Unemployment rate</th>
<th>% of 15-24 age group looking for a job</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men¹</td>
<td>Women¹</td>
<td>Total</td>
</tr>
<tr>
<td><strong>1995</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>65.9</td>
<td>35.4</td>
<td>50.6</td>
</tr>
<tr>
<td>Centre-North</td>
<td>70.0</td>
<td>42.3</td>
<td>56.2</td>
</tr>
<tr>
<td>South</td>
<td>58.4</td>
<td>23.1</td>
<td>40.6</td>
</tr>
<tr>
<td>EU</td>
<td>70.2</td>
<td>49.7</td>
<td>60.0</td>
</tr>
<tr>
<td>OECD</td>
<td>75.4</td>
<td>53.2</td>
<td>64.2</td>
</tr>
<tr>
<td><strong>2000</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>67.5</td>
<td>39.6</td>
<td>53.5</td>
</tr>
<tr>
<td>Centre-North</td>
<td>71.9</td>
<td>48.0</td>
<td>59.9</td>
</tr>
<tr>
<td>South</td>
<td>59.5</td>
<td>24.6</td>
<td>42.0</td>
</tr>
<tr>
<td>EU</td>
<td>72.5</td>
<td>54.0</td>
<td>63.3</td>
</tr>
<tr>
<td>OECD</td>
<td>76.2</td>
<td>55.2</td>
<td>65.6</td>
</tr>
</tbody>
</table>

¹ 15-64 years of age.


The Italian labour market has traditionally been very heavily regulated, particularly with respect to the forms of labour contract that employers may offer, and by the late 1980s, a growing number of observers had concluded that this rigidity was one of the major reasons for poor and deteriorating labour-market performance. On OECD indicators, Italy’s employment protection legislation (EPL) was restrictive even by the standards of Continental Europe, but its impact was highly variable. Some aspects of EPL for employees with permanent contracts were not exceptionally stringent, but firing restrictions were about the toughest in the OECD, owing in part to the protections enshrined in Article 18 of the 1970 Workers’ Statute. Article 18 stipulated that employers with more than 15 employees must rehire any worker found by a labour tribunal to have been fired “without just cause” and must pay back wages and social security contributions, as well as a large fine.¹²

Such severe restrictions on firing were believed by critics to deter job creation and reduce workers’ motivation: a study of 858 white-collar employees found a doubling of the absentee rate after a 3-month probationary period.¹³ The size restriction in Article 18 meant that it applied to just about 15% of all firms but covered 44% of all workers.¹⁴ Geography was also important: dismissal cases in the depressed regions of the Mezzogiorno tended to last longer and result in higher levels of compensation to workers. Some evidence suggests that the south’s high unemployment made judges keener to reinstate workers.¹⁵ Thus, poor labour-market conditions contributed to stricter enforcement of EPL, which, in turn, increased the adverse impact of strict EPL on labour-market outcomes.¹⁶ Article 18 was also of enormous symbolic value, which meant that any moves to alter it were likely to generate significant controversy.

Another peculiar feature of Italy’s labour market was the limited scope of the Italian safety net. Unemployment benefits have traditionally been very low and unavailable to many workers.¹⁷ The standard benefit in the 1990s amounted to just 30% of average gross earnings received over the last three months of employment, payable for six months; in 2000, this was raised to 40% and the maximum duration for workers over 50 was extended to 9 months.¹⁸ Workers in large manufacturing enterprises, however, have long been protected by a sectoral-corporatist system of short-term wage replacement benefits known as the *cassa integrazione guadagni* (CIG). Introduced on the basis of a private
agreement between employers’ associations and trade unions, the CIG are based on private contributions but also involve complementary contributions from the state. They provide benefits proportional to wages (usually 80% in the first instance) in the event of a temporary reduction of working time or suspension of activity. If redundancies occur, the affected workers are entitled to a “mobility payment”, which provides similar but declining benefits.  

The CIG system is characterised by two peculiarities. First, coverage is limited to large manufacturing enterprises. It thus provides exceptionally generous arrangements for the regular industrial workers who constitute the core of the trade-union movement. Secondly, provision of CIG benefits is not automatic but rather depends on the agreement of the state, the unions and the employers. It therefore provides an important role for political mediation in the management of enterprise adjustment and in the distribution of the costs of that adjustment. This discretion also means that, in combination, the short-term CIG benefits and mobility funds sometimes pay rather generous benefits for far longer than the twelve months regarded as “normal” in the scheme.  

The search for solutions

Until the mid-1980s, the peculiarities of Italy’s labour market were not widely regarded as a cause for concern. The protection of core dependent workers and family heads was a generally accepted priority, and this favoured strict regulation and high levels of job protection. This reflected not only the interests of unionised labour, whose core constituents enjoyed the greatest protection, but also of Italian firms, which relied on high internal flexibility – the system allowed for a great deal of de-centralised bargaining over issues like working time and conditions. The state was not very keen on external labour-market adjustment, as it would also have necessitated the extension of a social safety net at public expense. The practice of negotiating collective dismissals in large firms among the social partners was also valued by those involved, as it reduced adjustment costs for workers and firms during cyclical downturns. Moreover, the public monopoly on placement was an important political instrument, since for much of the post-war period, placement offices had considerable scope to tell firms which workers to hire. Altogether, this “breadwinner model” of employment protection made for a highly segmented labour market, characterised by difficult access to employment for women and young people, in particular. Regulatory rigidities also contributed to the development of the large underground economy.

The first tentative steps towards labour-market reform were taken in the mid-1980s, as successive governments sought to check the steady rise in unemployment that followed the second oil shock. While the major focus was on incomes policies and on stimulating activity in the Mezzogiorno, increasing attention was paid to active labour-market policies, and the authorities did introduce flexible employment-and-training contracts for young workers (contratti di formazione i lavoro), as well as job-sharing, part-time and “solidarity” contracts, all with a view to creating new opportunities for labour-market entry for disadvantaged groups, particularly women and the young. The impact of most of these measures was limited: the contratti di formazione i lavoro were popular on account of the fiscal incentives involved, but the other new instruments were rarely used, often because they imposed high non-wage costs.  

The real impulse to reform came with the economic and political crisis of the early 1990s. The technocratic cabinets of 1992-93 negotiated agreements with the social partners covering incomes policy and the reform of collective bargaining. In the
July 1993 pact, the unions accepted in principle a reduction in constraints on apprenticeships, employment-and-training contracts, fixed-term contracts and part-time contracts, as well as the introduction of temporary or agency work. Little was done to implement these measures in 1993-94, but labour-market reform was among the three tasks included in the mandate of the government of Lamberto Dini, which took office in January 1995, and the July pact constituted the *acquis* on which subsequent labour-market reform was to be built.

**Policy process: formulation**

*The preparation of the Treu and Biagi reforms*

**The Treu law**

Dini appointed Tiziano Treu, a professor of labour law from Milan and a newcomer to national politics, as Minister of Labour and Social Protection. Treu gathered a small group of experts (mainly labour lawyers) with a view to preparing a reform package. They spent 1995 and early 1996 working on the reform and consulting extensively with the social partners and representatives of the left-of-centre parties in parliament. This work continued almost without interruption through 1996 and into 1997, as Treu was retained in his post by Romano Prodi after the latter came to power following the 1996 general election. However, the government did not prepare draft legislation, present it to the social partners and then seek to secure its adoption; instead, the outline of the reform package was developed through a process of intense concertation. Only when the basic elements of the reform had been agreed and confirmed by the government and the social partners in the Pact for Labour did they become the basis for specific legislative proposals. To that extent, the politics of adoption were largely played out while the measures were being formulated: the draft legislation reflected what had been negotiated in advance.

Given the Prodi cabinet’s reliance on the far left RC in the Chamber of Deputies, the government operated under serious political constraints. It quickly became clear that any relaxation of employment protection for permanent workers would be politically impossible and might sink the entire reform. Unable to do much about exit, therefore, reform efforts were concentrated on liberalising entry. Here, too, the traditional Italian attachment to protective labour-market regulation meant that many of the innovations had to be very tightly regulated in order to win the acquiescence of the unions and the parliamentary left. For example, the legislation legalising temporary/agency work was initially rather restrictive. Interim agencies were subject to substantial entry barriers. The law allowed only a limited amount of temporary employment, which was to be determined by collective agreements, and forbade its use for unskilled workers. In addition, agency workers were to be guaranteed at least the minimum contractual wage of their sector of employment. Even then, “side payments” had to be included in the package in order to ensure RC support for the reform: the package included some 1 000 billion lire (equivalent to 0.05% of GDP in 1997) for “labour grants” for young people, particularly in the *Mezzogiorno*. In addition, the government promised the RC that it would prepare legislation on the transition to a 35-hour work-week. The government did indeed prepare a draft bill on this issue, but the text was rather weak and no serious attempt was made to pass it. However, the promise of such a law gave political cover to the RC as it voted in favour of the Treu law. The government’s commitment to transcribe into Italian law EC Directive 93/104, reducing the standard working week to
40 hours, was another symbolic gesture designed to ease acceptance of the reform. Although the legal standard of 48 hours, set in 1923, remained on the statute books, the 40-hour week had already formed the basis for Italian labour contracts for decades.

The Pact for Labour’s principal elements included an easing of regulations governing new apprenticeships; the creation of incentives to increase on-the-job training; an end to the ban on the organisation of temporary work via private agencies; and the creation of incentives for greater labour mobility. Labour grants were made available to young people who had never held jobs. The law also provided for special training and apprenticeship contracts for 16-24 year-olds (16-26 in the south) and for personalised counselling for young unemployed people in public employment offices in the south. The fiscal treatment of part-time work was brought into line with EU directives. The package put an end to the automatic conversion of fixed-term contracts into indefinite ones; this was one of the most controversial elements of the reform, owing to a fear that some individuals might spend decades working on successive short-term contracts without securing a permanent job. Regions and local authorities were given responsibility for implementation of most of the new measures. Decentralisation was attractive, given the wide variations in labour-market conditions across Italy and the apparent inefficiency of much national-level regulation, but the entry of more than 20 regions as new players in the field of labour-market regulation did entail significant implementation challenges of its own.

A number of considerations limited the government’s ability to move very far on the issue of unemployment benefits. A reform of unemployment benefits might have made a reform of EPL for permanent workers much easier to sell: some argued strongly for a compromise involving reduced EPL and increased expenditure on unemployment benefits and active labour-market measures. Both the July 1993 agreement and the Pact for Labour had included commitments to such a reform. The Prodi government prepared a bill on benefit reform, but the cabinet fell before parliament could consider it. However, the fiscal situation meant that there was only limited scope for any extension of the safety net. Moreover, the politics of safety-net reform were rather complicated. Some in the labour movement remained suspicious of an expanded safety net, viewing it as a “Trojan horse” that would ultimately lead to much more radical liberalisation of the labour market. Both large firms and the unions, as well as some politicians, were keen to retain the CIG, which could be managed with considerable discretion and which therefore presented considerable opportunities for local political bargaining. Any extension of the universal unemployment insurance system would probably have involved significant cuts to CIG schemes. There were also concerns about the purely administrative side: the difficulty of monitoring beneficiaries and avoiding fraud, given the scale of second-economy activity, and the weakness of local public employment services, which were being transferred to regional and provincial control.

The Biagi reform

Mindful of the risks of confrontation with the unions, the Berlusconi government in 2001 initially seemed ready to negotiate with them over labour-market issues, pension reforms and other matters. Centre-right politicians remembered the pension battles of 1994 and were aware that the centre-left’s success in engineering fiscal consolidation in the late 1990s had owed much to a good working relationship with the unions. However, the perception that the unions – and, in particular, the CGIL – were stalling soon led to a hardening of attitudes inside the government and among employers. In the late autumn,
the government began moving ahead with legislative proposals that had not been agreed with the unions.

The basis for these proposals was a white paper prepared at the request of Labour and Social Protection Minister Roberto Maroni by Professor Marco Biagi, a specialist in labour law from Bologna who had previously worked as an advisor to Labour Minister Treu. This in itself was a remarkable procedural innovation: Italy has no tradition of white papers. The white paper began with an overview of both EU recommendations in the field of employment policy and the state of the Italian labour market. It then proceeded to assess then-current labour-market policies, both active and passive, before presenting recommendations concerning future labour-market policy. The unions were divided in their reaction to the white paper: the CGIL promptly rejected it but the CISL accepted it as a “basis for discussion”. The government made clear its readiness to proceed, if necessary, without the unanimous agreement of the unions, a statement that provoked a sharp response from the CGIL. The CGIL saw itself defending not merely its position with respect to specific questions of policy but also the principle of concertation, which it regarded as the only legitimate way to make policy when acquired social rights were at issue. Yet despite its tough line, the government remained keen to avoid an unnecessary confrontation with the unions, and decisions about what would be included in the new legislative proposals and what would be left out reflected in part a desire to achieve some consensus with organised labour.

The draft legislation prepared on the basis of the white paper included: allowing private employment agencies to compete with public ones in the full range of services, including the provision of training (this relaxed some of the restrictions imposed on private agencies when they were legalised by the Treu reform); permitting open-ended staff-supply contracts in certain circumstances (the 1997 reforms had permitted staff-leasing only on a limited, fixed-term basis); a wider range of apprenticeships; improved conditions for the use by firms of part-time work; and greater opportunities for – but also better regulation of – non-standard forms of employment, such as on-call work, job sharing and project-based work (known in Italian parlance as “co-co-co”). Some of the most important innovations concerned training and employment contracts. Apprenticeship contracts were revised, and entrance contracts were proposed as a replacement for the old contracts, which had been found to contravene EU competition rules. Far from representing a radical de-regulation of the labour market, the proposals in many respects amounted to a re-regulation of employment relationships. The provisions concerning the “co-co-co” contracts, for example, were tightened in order to limit the scope for using such contracts for purposes of tax evasion.

In many respects, then, the proposals were clearly extensions of the Treu package. However, the first stages of debate over the reform of labour law were dominated by another issue entirely: the government’s attempt to suspend, at least partially, the application of Article 18 of the 1970s Workers’ Statute. The ultimate aim was to replace the right of reinstatement with a right to financial compensation, something the unions rejected. Initially, however, the government proposed only a four-year experiment in lifting the application of Article 18 in three cases: when small firms fell under the provisions of the article because they had raised employment above the 15-worker threshold, when businesses emerged from the shadow economy, and when short-term contracts were converted to indefinite ones. This meant that the changes would not affect the rights of anyone already protected by Article 18.
Despite the bitter controversies that surrounded the Biagi reform, there was in fact a high level of consensus concerning labour-market reform. Even many opponents of the Biagi reform praised the white paper, arguing instead that the government proposals did not properly reflect its recommendations. There was a broad body of opinion on the centre-left and centre-right that recognised the white paper’s overall assessment as being in line with the thinking that underlay the Treu reform. Indeed, some of the white paper’s proposals had first been included in the Treu package but had been withdrawn during negotiations with the social partners. Both aimed to increase the flexibility of labour-market entry not via de-regulation but via re-regulation – re-specifying and where necessary expanding the scope of regulation to meet new circumstances. Such an approach reflected in part the extent to which the reform process was dominated not by economists but by labour lawyers. This, in turn, reflected a number of factors, including the traditionally heavily prescriptive regulation of the non-economic content of contracts in Italy, the recognition that any reform would touch on highly sensitive acquired rights, and the belief held by some of the reformers that only careful revision of regulatory norms could limit the ability of Italy’s sometimes interventionist judges to subvert policy makers’ intentions.

The European dimension of Italian labour-market reform

European pressures account for a good deal of the continuity in Italy’s approach to labour-market reforms. As noted above, the single market itself generates pressures for reform. Moreover, during the 1990s, the drive to meet the Maastricht criteria loomed in the background as a sort of “conditional variable”: while the Treu reform did not have the same sort of direct relevance to EMU as pension reform or fiscal consolidation, there was an awareness among policy makers that higher employment rates would relieve some fiscal pressures. There was also an awareness that greater labour-market flexibility could help Italy adjust to shocks after the introduction of the euro and the consequent loss of national control over monetary policy. Moreover, elements of both the Treu and Biagi reforms were also necessitated by EU-level legal rulings. Over time, the European Union itself was becoming increasingly active in employment policy, and its initiatives may have had some impact as well – the Biagi reform was based in part on the European Employment Strategy. The Lisbon Strategy and related European employment initiatives are based largely on peer pressure and are not as binding as Maastricht or the ECJ’s rulings, but they have helped to disseminate reform ideas, increasing awareness in Italy of both the exceptional rigidity of some of its labour-market institutions and the successes achieved as a result of labour-market reforms elsewhere in the Union.

Policy process: adoption

The Treu reform

As noted above, the major conflicts surrounding the Treu reform were resolved as the package itself was articulated. Once the relevant concessions were made to the RC, its adoption was relatively straightforward, apart from the reform of unemployment benefits, which never took place. Labour-market reform had not been a major issue in the 1996 elections, but the Prodi government found it relatively easy to proceed on the basis of the July 1993 agreement and the mandate given to the Dini government. Two statutes passed in early 1997 gave effect to those elements of the reform that required parliamentary approval (many required only changes in secondary legislation).
The adoption of the Biagi reform was a different matter entirely. Despite the CGIL’s extremely hostile reaction to the October 2001 white paper, the coalition was determined to proceed. While it had presented no specific labour-market reform proposals during the spring campaign, it had won the elections on the basis of a broad commitment to economic liberalisation, and it needed to demonstrate a readiness to make good on that commitment. In mid-November, therefore, the government approved for submission to parliament a bill authorising the executive to reform the labour market, at the same time signalling its readiness to talk to the unions and even to modify the draft bill. However, the unions were unimpressed by what was on offer and the CGIL, in particular, took an absolutist line on Article 18, fearing that any revision to this highly symbolic provision would open the way to further erosion of its protections. Indeed, to a great extent, the CGIL’s opposition to the government’s proposals was driven by the logic of the “slippery slope” – a fear not of what was in the bill but of what would come next if it were adopted. The unions – particularly the CGIL – also objected strongly to the government’s determination to proceed with reform even in the absence of a tripartite agreement involving the unions.

While the Article 18 proposals were not the most important measures in the bill, they were of enormous symbolic importance and soon became the focus of intense conflict. In early 2002, the unions mobilised against the revision of Article 18, with the CGIL in the lead; the other two confederations were initially hesitant but grew more active as the depth of feeling of their members became apparent. Tensions rose sharply on 19 March, when Mr. Biagi was assassinated in Bologna by Red Brigade terrorists. While all sides condemned the killing, the heated rhetoric on both sides of the political divide that followed the tragedy served to widen the gulf between the government and the unions, strengthening the resolve of the reform’s opponents and further reducing the chances for a resumption of dialogue. Indeed, only after Mr. Biagi’s death did the three confederations finally agree to strike together, bringing some 13 million workers out in opposition to the Biagi proposals on 16 April. Altogether, the seven nation-wide strikes called by the unions resulted in an estimated 26.3 million working hours lost – almost three times the number of hours lost to strikes in all of 2001.

The government’s unity came under strain as the protests escalated. Some in the coalition openly criticised the reform, while others refrained from publicly defending it. At the end of May, the government began to give way on Article 18, offering to drop the proposed modifications of the article from its framework bill and deal with them separately. This opened the way to a compromise with the CISL and the UIL concerning the other elements of the reform, but the CGIL continued to boycott talks with the government. In June, the government offered EUR 700 million to fund the first phase in a gradual expansion of unemployment benefits as part of a package deal involving the key elements of the Biagi proposals. This allowed the government and some three dozen employers’ organisations and trade unions to sign the “Pact for Italy”; the CGIL, however, refused to sign. In addition to labour-market reforms, the pact included new resources to promote the development of the Mezzogiorno, a reduction in company taxation and a cut in income tax for families earning up to EUR 25 000 per year. This cleared the way for passage of the “Biagi law” and the relevant secondary legislation. However, the CGIL’s refusal to sign remained a problem, because the unions’ participation in the implementation of the reform was needed: details of many of the new contracts were to be hammered out in negotiations between employers and the unions.
The struggle over Article 18, however, continued. The government remained committed to revision, although it had to proceed cautiously, as it wanted the unions’ agreement on pension reform. The far left launched a counterattack in the form of a referendum on the extension of the article to all Italian enterprises. Neither succeeded in the end. Initiated by the RC, the referendum was held in June 2003. With the CISL and UIL calling for abstention and only the CGIL, after some hesitation, calling for a “yes” vote, the proposal failed. However, the government’s attempts to suspend Article 18 for three years in respect of small enterprises that grew beyond the 15-worker threshold also stalled. In negotiating the Pact for Italy with the CISL and UIL, the government had linked the partial revision of Article 18 to the reform of unemployment benefits, but it ultimately failed to deliver on this promise.52

Government communications strategies

Because there was more or less continuous discussion of labour-market reform in Italy throughout the period under study, one cannot clearly separate the communications efforts related to the two reforms. The Treu and Biagi reforms were both part of an on-going reform process, and the debates surrounding the former undoubtedly influenced the latter. Nevertheless, the two governments’ approaches to public communication were very different. The Prodi government made comparatively little effort to sell the Treu reform to the general public. Its attention was focused on the social partners, on the understanding that their acceptance of the reform would be sufficient to ensure its acceptance by the electorate. The centre-left only began to “market” the Treu reform at a later stage, when its benefits were becoming apparent. To the extent that the reformers addressed public opinion at all in 1996-97, they exhibited a tendency to present greater labour-market flexibility as something that was necessary rather than as a positively desirable good. Reform was something the government was obliged to do in view of the specific requirements of some pieces of EU legislation, as well as the general pressure for adjustment generated by Italy’s path towards EMU and increasing internationalisation. This approach to public relations arguably influenced the way flexibility was increased: liberalisation proceeded in piecemeal fashion, with numerous transitional and temporary measures, and new measures were often introduced with restrictions on their applicability, which were subsequently relaxed.

The centre-right government of 2001-06 initially sought to present a strong, public case for reform: that was in large measure the motivation for commissioning the Biagi white paper. However, when the battle over Article 18 erupted in early 2002, the government was unprepared, and for several months Labour Minister Maroni and his colleagues were largely isolated during the backlash against the reform: press coverage was generally hostile and they received little public support from other members of the government. However, as the debate over the Biagi reform continued, the government came to adopt a much more forthright posture. Far from under-selling the benefits of reform, as the Prodi government had done, the Berlusconi cabinet arguably oversold them. Although the Biagi law was in many respects an extension of the Treu reform, the government presented its plans as a break with past policy and exaggerated their potential impact. This strategy meant, however, that the government deprived itself of one of its strongest potential arguments in the battle for public opinion: the beneficial effects of the Treu reform – which were becoming increasingly apparent by 200153 – were not highlighted as a part of the government’s case for further reform.

Ironically, the government line was reflected to some extent in the response of the unions and the parties of the left: since they had accepted the Treu package but were
rejecting the Biagi proposals, the government’s opponents also had an incentive to present the Biagi law as a sharp break with the Treu reform and to pretend that it was far more radical than it was. More generally, the finely balanced competition between the two broad coalitions then alternating in power meant that each side faced incentives to emphasise discontinuities in policy, even against an underlying reality characterised by a high degree of continuity.

In essence, the centre-left seems to have tried to pursue labour-market reform by stealth, consulting widely with social partners and quietly introducing changes, without publicly questioning the basic elements of Italy’s labour-market institutions and policies. The centre-right, by contrast, openly challenged the system and bid for public support, while advancing measures that were, for the most part, simply further incremental steps in the direction taken by its predecessors. Moreover, its case for reform was arguably weakened by the failure to pursue liberalisation in other fields: the labour-market reforms were not presented as part of any broader reform agenda.

Policy process: implementation

The Treu and Biagi reforms have differed considerably with respect to implementation. Watanabe (2007:5) argues that the Treu measures were less extensive but more effective, precisely because they were formulated and adopted in concertation with the unions, which therefore co-operated in their implementation. As noted above, the Biagi reform depended in part on the willingness of social partners to implement many of its provisions in practice, and union resistance meant that many elements of the Biagi law were never implemented. The CGIL, having rejected the legislation, was not interested in participating in such bargaining. In theory, this need not have been a major obstacle, as the law provided for local or even plant-level negotiations. In practice, however, the issue was difficult: the norm in the past had been for such issues to be resolved via national-level bargaining. The CIGL rejected decentralised bargaining in these circumstances, and the CISL took the view that any decentralised bargaining under the law must be undertaken by consensus. In any case, the readiness of the more moderate unions to cooperate was undermined by their disillusionment with the Pact for Italy and their fear of being left politically exposed if they worked with employers to implement the law while the CGIL maintained its “rejectionist” stance. Some negotiations among the social partners nevertheless took place, but they have been focused on training and other relatively uncontentious issues.

Some observers see this as the main reason why many of the new labour contracts are still not used at all, while others remain extremely rare. However, Confindustria surveys of firms’ use of the Biagi innovations cast doubt on this conclusion. While difficulties with the unions were cited as a problem by some respondent enterprises, the principal obstacle cited in explaining why they do or do not use the new contracts were problems with the complex secondary legislation implementing the Biagi law and difficulties in dealing with the workers affected. Overall, the survey data show a very positive employer assessment of the Biagi reform, particularly in respect of measures designed to improve the matching of labour supply and demand.

Uncertainty may also have limited the application of some measures. The centre-left, as well as the CGIL, remained hostile to the Biagi reform, and the centre-left L’Unione coalition promised to overturn much of it. Its 2006 election platform included promises to restrict non-standard contracts to four main types; to limit temporary employment to pre-determined areas; and to restrict the duration (or number of renewals allowed) in
cases of temporary employment. *L’Unione* also called for the introduction of some flexibility into indefinite contracts by gradually extending employment protection in proportion to seniority and for the restoration of tax credits for permanent job creation. To date, however, the *L’Unione*-led government has made no move to overturn Biagi, which now looks set to remain on the statute books. This has led some observers to suggest that the social partners may now begin to negotiate on some of those innovations in the law that remain unimplemented.

**Conclusion**

While the Treu and Biagi packages have helped to move the focus of labour-market policy somewhat from an overriding concern with security and income-protection for employed insiders towards a greater emphasis on creating opportunities for outsiders, the reforms had their limitations. The two key issues that the Treu package did not address have largely remained unresolved since 1997. The first was the relaxation of employment protection for permanent workers. The second was the comprehensive reform of unemployment benefits. While some progress was made on this front, it was limited, and the perceived weakness of Italy’s safety net remains an obstacle to reforming indefinite labour contracts.

The evolution of this reform process suggests that concertation need not be considered a *sine qua non* for adopting reforms in Italy, but a government may pay a price for opting to proceed with reforms in the absence of concertation. In the case of the Biagi reform, the mere fact that the government signalled its determination to proceed without going down the path of tripartite negotiations seems to have strengthened union opposition to the reform. To a great extent, the unions were simply defending what they had come to regard as their rightful role in the process. Nor did the confrontation over the Article 18 proposals in the spring of 2002 help matters: the government’s initial determination to defend these relatively minor proposals proved costly in political terms and made it harder to secure agreement on the remaining elements of the reform. Moreover, subsequent problems with the implementation of certain elements of the Biagi law do appear to reflect, in part, the unions’ resistance to some of its innovations. That suggests that there may be a trade-off facing reformers: those willing to override union opposition may be able to adopt measures that labour opposes, but they will find it hard to implement anything that requires the unions’ active cooperation. The choice of whether, and how much, to pursue negotiation rather than confrontation may thus depend to a considerable degree on the nature of the measures under consideration and not merely the relative strength of the government and its opponents: the degree to which stakeholder cooperation is needed for effective implementation may also be a factor.57

The Biagi reform highlights another challenge that would-be reformers may confront: stakeholders’ fear of the “thin end of the wedge”. Even a relatively modest reform may engender fierce resistance if forward-looking agents see it as setting an important precedent or otherwise reducing the obstacles to further reform. In the case of Article 18, the Berlusconi government suffered a politically costly defeat over a measure that would have had very little impact if implemented. Yet the CGIL’s position from the start reflected the conviction that even a minor modification of the article would represent the beginning a sustained effort to weaken it. Such “slippery slope” logic is not without foundation. The elimination of the rents or protections enjoyed by some agents may well make those retained by others both more apparent and more difficult to defend.58 The growing dualism found in some partially reformed labour markets exemplifies this
problem: as the regime under which labour market “outsiders” operate is opened up, the privileged position of “insiders” becomes harder to justify. If, moreover, first-stage reforms are seen to generate economic benefits, then the case for further reform will be strengthened. At a more general level, there appear to be interlinkages among structural policy settings in different domains (e.g. trade, product markets, labour markets) that suggest that reforms in one field may generate pressure for reform in others.\textsuperscript{59} As increasing international – and, in particular, European – integration has reduced rents in Italian product markets, pressure for labour-market reform has grown.

Notes

1. Including on-call working, job sharing, accessory working, work experience programmes and staff leasing on an open-ended basis.
3. This is one reason why the advent of Italy’s “second republic” in 1993 has not brought the kind of cabinet stability many had hoped for: to date, only one of the nine governments to take office since the 1994 election has run a full term.
5. The RC was the more left-wing of the two successor parties to the old Italian Communist Party; the other, the Partito Democratico della Sinistra (Party of the Democratic Left, PDS), was actually the largest single party of the Ulivo coalition.
6. It may be symptomatic that the coalition broke down soon after this goal was achieved.
7. OECD (1999); and OECD (2003a).
8. See IMF (2004); Brandt et al. (2005); Berger and Danninger (2006); and Høj et al. (2006).
9. See the case study of the 1994 pension reform for more background on the trade union confederations.
10. Long-term unemployment accounted for 69.2\% of the unemployed in the under-developed Mezzogiorno and 55.5\% in the north. Most of the unemployed were “secondary” workers with family support (i.e. young people and housewives), many of whom preferred to work in the black economy while waiting for a job in the primary sector. Ferrera and Gualmini (2000a:380-1).
11. All data from ISTAT, cited in Samek Lodovici (2000), who argues that, in the south, self-employment was mainly a way to absorb excess labour and circumvent fiscal and employment regulations. In the north, by contrast, she argues that it reflected a tradition of craft work and small family firms, which are largely exempt from heavy employment regulation, and also the expansion of professional groups in business and personal services. Self-employment thus tended to be highly cyclical in the south of Italy but fairly stable in the north.
12. Within the OECD, only Austria and Portugal also retained a right to re-employment; other OECD members provided only for monetary compensation.

14. Many labour-market regulations applied only to firms above a certain size. Stricter regulation of larger firms constituted an incentive for firms to remain small. This is probably one reason why average firm size in Italy is relatively small by EU standards; see OECD (2007) on the implications of this tendency.


16. Such judicial behaviour also raises equity concerns: the law should, in principle, be applied uniformly and consistently across the country.

17. Even in 2001, after a significant increase in the level of benefit, Italy spent only about 0.3% of GDP on unemployment benefits – a lower level than in the United States, where the unemployment rate was less than half the Italian level.

18. Further increases in benefits were adopted in 2005 and 2007.

19. Traditionally, such workers have been placed in a pool that enjoys a certain priority when new hiring takes place.


21. See Libro Bianco (2001), Section 3.1. Critics argue that the system may be prone to manipulation by local patronage networks.

22. According to the OECD’s employment protection indicators, Italy has long had the strictest regulation of collective dismissals in the OECD area.

23. Thus, if a large-scale redundancy had to be negotiated, the state might ensure affected workers would be given priority over other job-seekers by the placement offices.

24. For an overview of these early reforms, see Graziano (2004).

25. For example, by offering young people financial incentives to create new firms and co-operatives there.


27. The 1992 agreement put an end to the scala mobile cost-of-living indexation mechanism, and the 1993 pact included a commitment to concerted monitoring of wages, prices and tariffs with a view to controlling inflation. The partners agreed to a flexible adjustment to expected inflation, resolved in tripartite negotiations set up to coincide with the macroeconomic and budgetary planning cycle. This structure was later completed and strengthened by the “Social Pact for Economic Development” in 1998.

28. A new collective bargaining structure was organised on two levels: national, for the regulation of wages in conjunction with inflation dynamics and firm/territorial level for the regulation of wage differentials and compensation of productivity increases. See Rhodes (2001) for details.

29. The other two were pension reform and fiscal consolidation.

30. They had to be authorised by the Ministry of Labour and Social Protection, to operate in at least four Italian regions, and to deposit a large cautionary sum for their first two years of activity. In addition, they had to pay 5% of workers’ gross wages into a special Vocational Training Fund.

31. Italy has no national statutory minimum wage.

32. Apprenticeships were redefined to last 18–48 months, their applicability was widened to include large firms, and the maximum qualifying age was raised from 20 to 24 (26 in the south).
33. Up to 24 hours a week.
34. See Ferrera and Gualmini (2000b) for details. The tasks which remained the responsibility of the central state pertained to control, certification and delegation, whereas two fundamental groups of functions – job placement and the promotion of pro-active (employment) policies – were instead devolved to the regions.
35. In part, this was because of the postponement of savings from pension reform.
36. During the negotiation of the Treu reform, some representatives of the employers’ associations stated explicitly that they would not agree to replace the CIG with what they regarded as an expensive and unmanageable universal system that would be fraught with moral hazard and inefficiency. This position also enjoyed support among some trade union officials.
38. CGIL leader Sergio Cofferati described the white paper as “a photocopy” of Confindustria’s own programme, while admitting that he had not actually read Biagi’s work.
39. For example, porter duties, cleaning work, transportation services, storage, managerial consultancy, call-centre management.
40. Collaboratori coordinati e continuativi.
41. Eligibility for the old contracts varied over time and across regions, but they entailed payroll tax cuts for people up to the age of 32, even where the beneficiaries had tertiary qualifications. The training content of these contracts was in any case very limited. (See Bertola and Garibaldi, 2003.) The proposed new entrance contracts were to be based on individual projects specifically aimed at adjusting the individual worker’s skills to a working context. They were to be open to young people aged 18-29, as well as to some other categories of disadvantaged workers. Workers employed on such contracts could receive a remuneration two levels below that payable to a regular employee with the same professional qualification, and they were not to be included in employee headcounts for the application of regulatory measures that were size-dependent.
42. “Co-co-co” contracts were meant to be for project work and similar arrangements that fell in between dependent employment and self-employment. However, the tax benefits that they conferred often made them attractive in situations devoid of any genuine element of self-employment. The draft law thus included tighter eligibility requirements for such contracts and measures to bring the tax position of “co-co-co” employees closer into line with that of dependent employees. See Maggi and Tozzoli (2004).
43. There was plenty of scope for bargaining over economic issues – Italy has no national minimum wage, its working-time regulations are fairly relaxed, etc. – but the typology of permitted contracts has tended to be rigid and constraining.
44. The Treu reform’s opening up of worker-dispatching services and the Biagi law’s replacement of the contrati di formazione e lavoro.
45. Beginning with the Essen European Council in 1994, which identified five priorities for improving labour-market performance, the title on employment included in the Treaty of Amsterdam (1997) and the Lisbon Strategy.
48. The fact that some elements of the draft legislation would relax restrictions that had been negotiated in the Treu reform merely reinforced these concerns.
49. Based on ISTAT data reported by ANSA, 16 December 2003.

50. The government insisted that it could not immediately provide the EUR 4 billion or so needed for a complete overhaul of unemployment benefits.

51. The pact called for an increase in the Mezzogiorno’s share of capital expenditure from the state budget to 45% over 2002-08, as well as a requirement that 30% of public spending on areas such as railways, road and other infrastructure go to the Mezzogiorno.

52. The government proposed legislation to increase the ordinary unemployment benefit to 60% of previous wage for the first six months, falling to 40% for months 7-9 and 30% for the final three months. However, the legislation was never adopted. In 2005, the maximum duration was extended by one month, to seven months for people under 50 and 10 months for those aged 50 or more. The initial benefit level was raised to 50% for the first six months, falling to 40% thereafter and to 30% in the tenth month (for the over-50s only). However, this was a temporary measure, covering the period from 1 April 2005 through 31 December 2007.

53. See OECD (2001) for a summary of the evidence suggesting that improved labour-market performance was in part a result of the greater flexibility brought about by Treu.

54. See, e.g., Watanabe (2007); and Accornero (2005). This view is shared by many policy makers and union officials as well. Lamelas and Rodano (2005) argue that “…the reform ‘leans against the wind’ by imposing innovations to both firms and unions…. the previous reforms worked on the same line of the experiences gathered during the collective bargaining, possibly improving them.”

55. See Confindustria (2006); it should be noted that the sample consists chiefly of industrial firms: many of the contracts introduced in the Biagi law are more relevant in service sectors, and some Confindustria data suggest that some of the non-standard contracts may indeed be more widely used in services sectors (Confindustria, 2007); see also Accornero (2005) on the use of the new contracts. See also Cipollone and Guelfi (2006); and Destefanis and Fonseca (2007).

56. Temporary, part-time, interim, apprenticeship.

57. For example, relatively simple liberalising measures – such as the removal of price controls – may be easier to impose over the resistance of trade unions or other interest groups than reforms in fields like education and healthcare, where stakeholders adversely affected by the reform will nevertheless be involved in its implementation.

58. The progress of pension reform in France illustrates the way in which the groups initially best placed to defend their privileges may be isolated over time, as the schemes for other segments of the population undergo reforms.

59. As an empirical regularity, structural reforms in domestic product markets tend to be preceded by external liberalisation (usually manufacturing trade liberalisation, followed by financial market/FDI liberalisation), while labour market reforms follow.
Bibliography


Summary and main conclusions

In July 2001, the newly elected administration of President Vicente Fox invited organised labour and private-sector employers to begin tripartite negotiations on the reform of Mexico’s Federal Labour Law (LFT). This process resulted in the introduction of two initiatives in Congress in late 2002, one associated with – though not formally submitted by – the Fox administration and the other advanced by its opponents. The administration proposals included a number of needed measures, such as the introduction of probation periods, some easing of the regulations concerning permanent contracts, modest steps to modernise industrial relations, some simplification of bureaucratic requirements for SMEs, and measures to promote labour force training. The proposals did not address all of the important issues, but they would have marked a step forward, by reducing both legal uncertainty and the scope for corruption, while increasing labour-market flexibility. The opposition draft, by contrast, focused on reforming the system of labour relations but did little to make the market more flexible. Neither bill was approved. Further negotiations in Congress in 2003 resulted in a compromise proposal, but it, too, failed to win adoption.

Among the most important reasons for this disappointing outcome were the following:

- The design of the process. The government created a special forum in which to negotiate a reform acceptable to employers and unions alike. It was decided at the outset that there would be no discussion of changes to the constitution, which would have been extremely difficult in any case and might have made reform impossible. It was also decided that the administration would take forward only reform proposals that emerged from the process with unanimous support. Taken together, these principles greatly reduced the chances of agreement: the latter was intended to ensure that no group could be marginalised from the process, but it made every participant a de facto veto player, while the former ensured that a fundamental overhaul of labour relations – the top priority of some of the trade unions involved – could not be addressed. These “independent” unions wanted to break the collusive pseudo-corporatist arrangements that had evolved between employers and the so-called “official” trade unions during the decades of rule by the Institutional Revolutionary Party (PRI). In the end, the “independent” unions rejected the proposals on the table and advanced a separate legislative initiative of their own, with the support of the left-wing Democratic Revolutionary Party (PRD).
• **Lack of a strong desire for reform among the participants in the negotiations.** While there was very wide agreement on the need for reform, none of the principal actors, except perhaps the “independent” unions, was prepared to risk serious conflict over it in the early 2000s. Employers welcomed some reform proposals, but most did not regard labour-market regulation as among their main problems, and many did not want greater labour-market flexibility at the price of satisfying the “independent” unions’ demands for a major restructuring of industrial relations. The major PRI-affiliated unions were concerned to prevent changes in the system of labour relations that would undermine their position.

• **The lack of a strong, organised constituency of potential winners from reform.** The existing regime tended to protect incumbent workers, firms and trade unions. This made it difficult to identify, let alone mobilise, a pro-reform constituency, since those with the most to gain from a successful reform were either disorganised and politically marginal (such as informal-sector workers) or did not exist (firms that had not entered the market). Many of the real costs of the status quo were opportunity costs, which were invisible to the public and largely irrelevant to the interest groups participating in the tripartite discussions.

• **Lack of government leadership.** During the roundtable phase, the STPS chose to remain neutral, arguing that it ought to remain impartial until the parties reached a consensus. The ministry’s neutrality may well have helped bring some of the parties to the negotiating table. However, when combined with an awareness that the administration would not act unilaterally, it meant that there was no pressure on the other participants to make concessions. Moreover, despite a pre-election commitment to labour-market reform, the Fox administration subsequently failed to make a priority of winning adoption of those proposals that almost all of the roundtable participants endorsed. Without presidential leadership, the parties in Congress had no incentive to risk backing reform ahead of the 2003 congressional mid-terms.

The context for reform

*Framework conditions for reform*

The Fox administration’s major labour-law reform initiative was undertaken in the years immediately following the end of the Institutional Revolutionary Party’s (PRI) 70-year monopoly on power. The PRI’s defeat in 2000 shook up many previously stable political configurations and alliances, creating new opportunities for coalition-building and reform. However, the election of an opposition candidate to the presidency also exposed the limits of presidential power in Mexico, especially as Mr. Fox’s National Action Party (PAN) did not command a majority in either chamber of Congress. Moreover, disappointing economic performance in 2001-03 added to the administration’s challenges. While there is little evidence that declining formal-sector employment in 2001-02 had much direct impact on the politics of labour law reform, the economic slowdown did rapidly undermine popular support for the new administration.

*Labour law and labour relations in Mexico*

There was and is broad agreement on the need for labour-law reform. The Mexican labour market is characterised by the coexistence of a large, and largely tolerated, informal sector, where workers have only rudimentary legal protection (and do not pay
taxes), and a formal sector where jobs are generally more skilled and protected by strict legislation. Estimates of the size of the informal sector vary, but it is widely reckoned to account for as much as 40-50% of total employment. The formal labour market is, by international standards, exceptionally rigid, at least on paper. A highly prescriptive constitutional and statutory regime governs everything from working time to trade union operations, granting extensive protection to workers in virtually every area. However, numerous analyses have argued over the years that in practice, this regime has historically served to enable employers, the unions and the state to control workers. The specific features of this system of labour relations largely defined the politics of labour-law reform, since the usual labour-management conflicts over questions of labour-market flexibility were complicated by sharp divisions among the unions themselves about the provisions governing trade union formation and operations.

Article 123 of the Constitution promises an extensive array of rights and protections to workers, although it also curtails many of these rights (particularly the right to strike) in respect of federal employees. The Constitution is quite prescriptive regarding many matters that are left to collective bargaining or regulated by statute in other countries, such as working time, maternity leave, vacation, dismissals and severance pay, promotion, and so forth. The practical significance of Article 123 is largely defined by the Federal Labour Law (LFT). The LFT is highly prescriptive with respect to things like work contracts, and it also regulates the creation and status of trade unions, collective bargaining procedures and dispute resolution. As administered by successive PRI administrations, the LFT formed the basis for a system of state-dominated corporatist interest representation. Critics argued that these arrangements made union loyalty relatively easy to secure, as it was often possible in practice for officials to deny union registration or to refuse recognition of unions’ executive officers, both of which were prerequisites for negotiating collective contracts. Moreover, unions allied to the PRI enjoyed significant benefits, while “independent” unions suffered discrimination or even outright repression. Loyal union leaders were elected to seats in state and federal legislatures on PRI lists, and union members enjoyed access to superior healthcare, state-sponsored housing programmes and stores offering subsidised prices on consumer goods.

Although the labour-relations situation began to change significantly in the 1990s, many features of this system survived the end of PRI rule. Unions that were historically affiliated to the PRI – often referred to as “official” unions, although they have no particular official standing in law – still predominate and retain substantial control over personnel policies within the enterprise, since “exclusion clauses” in collective contracts typically stipulate that firms must hire union members and must fire employees who leave, or are expelled from, the union. Exclusion clauses in cases of separation from the union were declared unconstitutional by the Supreme Court of Justice in 1999, but the relevant provisions of the LFT were not amended. Where vacancies arise, the union often has the right to recruit new workers and present them to the enterprise.

Labour conflicts are resolved by tripartite Juntas de Conciliación y Arbitraje (Boards of Conciliation and Arbitration), which exist at both federal and regional levels. Workers may only strike with junta approval. If a junta declares a strike non-existent, the workers involved must return to work within a day or face dismissal. State officials preside over the juntas, which are responsible to state or federal governments rather than functioning as part of the judiciary, and some analyses have concluded that the juntas have often been very responsive to the views of the executive organs under which they served. Historically, the juntas have been used to curb labour militancy, as state and local
authorities sought to protect local businesses in order to ensure that they remained in the region. Representatives of “independent” trade unions argue that they still do so. Moreover, where “protective” union arrangements are in place (see below), a worker’s official representative before the junta may represent the interests of the employer rather than the employee. Junta decisions may be appealed in the courts.

This system creates opportunities for numerous informal practices neither intended by nor envisaged in Mexican legislation, including corruption and abuse. First, management can simply co-opt union leaders, offering bribes in exchange for inexpensive collective contracts. Alternatively, union bosses can extract bribes from entrepreneurs, threatening to mobilise workers unless they are appeased. Such possibilities have given rise to what is called “protective” unionism: a “protection contract” is an informal arrangement reached between a union and an employer to protect the employer from the claims of the workers and unions. Some union representatives claim that the majority of labour contracts in Mexico are protective, including up to 90% of those concluded in the SME sector. Some long-established “official” unions have been accused of “protective” unionism, and there are also unions created by the enterprises themselves, known as “white unions.” In some cases, employers rely on “phantom” unions, which are not even known to the workers and which may exist only on paper. The workers in such cases are not affiliated to the union, but their work conditions are determined by a collective agreement between it and the enterprise. When forming a business, an entrepreneur often forms, or reaches agreement with, a protective union even before beginning to hire. The collective contract is put in place ahead of time to protect the entrepreneur against unwelcome pressure from the workforce. White unions also make it easier to keep active unions out of the workplace, since it is harder for workers to change unions once one is established in the enterprise. Consequently, real trade union activity is non-existent in many firms.

**Actors and interests**

Since no party commanded a majority in either chamber of Congress, passage of any reform of the LFT would have required the support of at least two of the three major parties. The Fox administration and the PAN as a whole faced conflicting incentives when it came to labour-market reform. The PAN had long been committed to a democratising agenda that would have favoured dismantling the authoritarian-corporatist structures created by the PRI, but it was also committed to further economic liberalisation, including greater labour-market flexibility, which would be harder to pursue in an environment in which autonomous unions were proliferating and growing stronger. At the same time, the PAN, once in power, needed the cooperation of the historically PRI-affiliated “official” unions to govern; conflict with them might threaten both wage moderation and labour peace. The PRI, for its part, was divided over issues like labour-market flexibility but generally united on labour relations: PRI leaders were keen to protect the prerogatives of the “official” unions, who were strongly represented in the party’s congressional delegations. PRI control over these unions had long rested on control of the state in order to provide the legal, political and economic incentives for union loyalty, and the elections of 1997 and 2000 had disrupted this state of affairs at federal level, reducing the PRI’s leverage over “its” unions, which had in any case been gradually declining as a result of the gradual political opening that preceded those elections. This made PRI support for labour-market reform even riskier than it would have been in earlier years, when a wider range of sanctions and rewards could be used to secure union acquiescence. The left-wing Democratic Revolutionary Party (PRD) was the
only major party that was more or less united on labour-law reform: it supported an end to the corporatist model of labour relations but strongly opposed moves towards greater flexibility.\(^{23}\)

Despite sharply falling unionisation rates,\(^{24}\) Mexican trade unions remained influential in the early 2000s, partly due to the legal prerogatives they enjoyed and partly because of their ability to mobilise protest in opposition to reform. The basic cleavage in the trade union movement pitted the PRI-affiliated “official” unions against the new “independent” trade unions.\(^{25}\)

- Most of the “official” unions came under the umbrella of the Congress of Labour (CT), the institution created by the PRI to coordinate its relationships with the main confederations.\(^{26}\) The CT unions altogether had about 3.8 million members in the early 2000s. During the negotiations over labour-law reform, they worked to preserve legal provisions of Article 123 and the LFT that, critics claimed, allowed them to control their members and to prevent challenges from rival unions. At the same time, they sought to change those provisions of the LFT that were felt to make it relatively easy for employers to use “white unions”.\(^{27}\)

- The National Union of Workers (UNT), with around 1.5 million members, was the largest confederation of autonomous labour organisations in Mexico. Created in 1997, the UNT brought together both unions with a long history of “oppositionist” trade unionism and formerly PRI-affiliated unions that had left the CT. The UNT was particularly strong in non-tradable sectors, including the public sector and network industries.\(^{28}\) These sectors were not subject to the same competitive pressures as others, and their trade unions were, perhaps as a consequence, less willing to compromise than those representing workers in more exposed manufacturing sectors. The UNT’s declared goals included a reduction in state control over union formation and operation, the elimination of legal provisions that made it harder for them to challenge the hegemony of the “official” unions, and the replacement of the juntas by independent tribunals.

The two most influential multi-sectoral employers’ organisations in Mexico have tended to be the Employers’ Confederation of the Mexican Republic (Coparmex) and the Entrepreneurial Coordinating Council (CCE). While their constituencies and political orientations differ somewhat,\(^{29}\) the CCE and Coparmex approached questions of labour-market reform with much the same agenda. They wanted flexibility and simplicity in contracting, including authorisation of such basic practices as payment of wages by electronic transfer. In addition to further steps towards labour-market flexibility and the removal of obsolete but potentially problematic provisions, employers’ bodies were interested in the revision of the LFT to confirm the limited de facto flexibility that they had acquired under the administrations in office from 1982 to 2000. These administrations had already made some progress in liberalising the labour market de facto, if not de jure, via more flexible application of the LFT’s provisions and the use of state power in the juntas. Employers had long sought to incorporate these changes in the LFT itself. The law as it stood contained ambiguities that created legal uncertainty and opportunities for corruption.\(^{30}\)

Nevertheless, labour-market reform was not the top priority for Mexican employers: most had already found ways of adapting to rigid formal regulations, and many were quite comfortable with “protective” trade unionism and the restraints imposed on real labour militancy by the corporatism of the PRI era. Though heavy regulation probably did constrain job-creation in existing firms at the margin, the CCE did not see the labour-law
regime as a major day-to-day handicap for its members. Given the costs of putting such adaptive strategies in place, the over-regulation of the labour market constituted a barrier to new entrants, deterring both domestic and foreign investors. Business surveys suggested that labour-market reform was one of the reforms that could do most to stimulate entry by both domestic and foreign firms. Incumbent enterprises, having adapted to local conditions, may thus have benefited from the deterrent effect of labour regulation on potential entrants.

The real losers from Mexico’s over-regulation, then, may well have been the foreign and domestic firms that did not enter the market and the workers who might have found employment with them if they had. Since neither the non-entering firms nor their workers could be identified, they carried little political weight: to the extent that the costs of Mexico’s labour law were opportunity costs, they were largely invisible to the public and irrelevant to the calculations of many of the actors involved in the policy process. The other potential beneficiaries of labour-market reform were informal-sector workers, who enjoyed virtually no legal protection and limited access to social services, training and other opportunities to better their positions. Advocates of labour-law reform argued that a better balance between flexibility and worker protection was a critical element of any effort to reduce the disincentives to formal-sector employment and thus draw more workers out of the informal economy. Such a shift would have allowed more workers to find jobs with better legal protection. However, the interests of informal-sector workers seem to have carried little political weight in the labour-law reform debate of the early 2000s. While there was an awareness of the potential for reform to draw activity out of the informal economy, the political bargaining over the LFT was dominated by the interests of system insiders – incumbent firms and trade unions representing formal-sector workers.

**Issue history**

During the 1980s and 1990s, successive PRI administrations implemented reforms aimed at liberalising trade and capital flows, reducing and reorienting the direct participation of the state in the economy via privatisation and deregulation, and allowing the emergence of freer markets in general. However, the basic labour law remained untouched. Instead, Mexico underwent a process of gradual flexibilisation, as management gained greater control over production-related decisions via adjustments to collective agreements and changes in the application of the LFT. While this made life easier for firms, advocates of reform continued to press for a real overhaul of the LFT. They pointed to the need for more adaptability in the formal sector as a result of growing competition within NAFTA and from low-income countries, as well as lower inflation. Successive OECD and World Bank analyses, in particular, highlighted the economic and social costs of Mexico’s dual labour market and argued that reform of formal labour-market institutions constituted a key element of any attempt to draw more employment into the formal sector: both equity and efficiency concerns pointed to the need for change. There was also growing attention to the issue of labour-market reform within the country, as employers, unions and the government began to look for a way forward. While there was a consensus on the need for change, however, there was no consensus on its direction: the diversity of views and approaches was apparent in the more than 250 labour-law reform bills submitted to Congress during 1997-2008.

Broadly speaking, two different agendas for labour-law reform evolved in parallel during the 1990s. The clash between them played an important role in thwarting attempts to revise the LFT in the early 2000s.
The first reform agenda took shape in discussions involving Coparmex and the CT-affiliated Congress of Mexican Workers (CTM). In 1989, Coparmex proposed formalising many labour relations practices that were already in place, as well as measures to increase flexibility even further in some spheres. However, the proposal was quite cautious, reflecting employers’ reluctance to risk any radical changes in state regulation of union activity and labour conflicts. \(^{40}\) Piecemeal change via collective agreements therefore continued, supported by the government’s active encouragement. In 1995-96, in the aftermath of the Tequila crisis, Coparmex and the CMT, with the PRI’s approval, reached agreement on a “New Labour Culture” (NCL). The NCL saw greater workplace flexibility as critical to enhancing the efficiency and competitiveness of the Mexican economy, and rejected the concept of class struggle as the defining feature of labour relations. \(^{41}\) In place of class conflict, the NCL presented a view of the enterprise as a community of stakeholders who cooperate in the struggle for competitiveness. The basic bargain underlying the NCL was clear: the unions were ready to concede increased flexibility in return for the preservation of the status quo with respect to internal union procedures and the existing corporatist system of labour relations, about which the NCL was silent. \(^{42}\)

An alternative vision of labour-market reform was meanwhile emerging among opposition parties and “independent” unions. In 1995, the PAN presented a reform proposal drafted with the assistance of two prominent labour lawyers. While it included some measures aimed at greater flexibility, the core elements of the PAN proposal included, inter alia, an end to exclusion clauses, greater transparency in union elections, the creation of workplace committees apart from trade union structures, and the elimination of the executive branch’s role in dispute resolution via the replacement of the tripartite juntas with tribunals in the judicial branch. A few years later, the left-wing PRD put forward a similar initiative. This convergence of left and right opposition parties in the early 1990s reflected the desire of both to weaken the corporatist structures created by the PRI. \(^{43}\)

During the 2000 election campaign, a group of prominent labour lawyers and representatives of the “independent” unions issued a document called “Twenty Commitments for Union Freedom and Democracy”, which advocated steps to guarantee workers’ freedom of association and organisation more effectively, as well as genuine collective bargaining, the right to strike and the replacement of the juntas by independent tribunals. PAN candidate Vicente Fox expressed his agreement with the “Twenty Commitments” in an open letter shortly before the election, thereby giving rise to expectations that he would use it as the basis for any overhaul of the LFT. \(^{44}\) The new administration’s National Development Plan for 2000-2006, which appeared early in 2001, made it clear that labour-law reform would be a priority and that its aims would include not only greater labour-market flexibility but also the promotion of human capital development and the improvement of the position of marginalised segments of the labour force. \(^{45}\)

**Policy process: formulation**

In July 2001, the STPS formed a tripartite body called the *Mesa Central de Decisión* (“Central Decisional Board”, MCD) to consider labour-law reform. The MCD brought together twenty-two individuals, comprising equal numbers of union and employers’
The union representatives included three representatives of the UNT, the first time that it had been invited to participate in an official policy consultation; the other eight represented CT-affiliated organisations. In addition, the MCD included members of the Labour and Social Welfare Committees of the two houses of Congress, and eight representatives of the STPS, including the Minister. However, the MCD was intended to be a largely bilateral forum:

- The congressional representatives participated as observers only; this was meant to de-politicise the work of the MCD. STPS officials and some other Mesa participants expressed the view that previous reform initiatives had failed precisely because political parties and their electoral agendas had “contaminated” the process. Though welcomed by the social partners, the Fox administration’s solution generated some resentment in Congress, where some saw it as an attempt by the executive branch to marginalise the legislature, a perception that did nothing to facilitate adoption of LFT reform when it reached Congress.

- The STPS insisted that the executive would not impose its views on the MCD, acting instead as a neutral facilitator of dialogue. Given some unions’ distrust of the minister, a former head of Coparmex, this was intended to ensure that the legitimacy of the process was not felt to be compromised by his participation. However, a number of participants in the MCD, including some UNT representatives, argued that the president’s failure to get directly involved in labour law reform himself undermined the credibility of the process.

Two key principles were articulated at the outset to govern the MCD’s work. First, there would be no discussion of any revision of the Constitution. Constitutional change would have been very difficult given the requirement that such a change be ratified by two-thirds majorities in both houses of Congress and by a majority of the states, and a focus on revising Article 123 might well have made it impossible to adopt any reform whatsoever. Secondly, the Mesa would reach final decisions only through consensus. This was intended to ensure that no participating group could easily be marginalised by the others. Taken together, however, these stipulations made the challenge of reaching agreement on a substantial reform far more difficult. The first principle ensured that the autonomous trade unions’ desire for a fundamental overhaul of labour relations would not be satisfied, while the second made every participant a de facto veto player. This was a very problematic combination, especially given the administration’s reluctance to lead the process with its own proposals.

The exclusion from the agenda of the issues of greatest concern to the UNT meant that it had little incentive to do anything but exercise its veto. Some concessions to the UNT on non-constitutional issues might still have been possible via amendments to the LFT, but the UNT was also opposed to the other participants’ insistence that the “New
Labour Culture” form the basic framework for the MCD’s work. In the UNT’s view, this effectively precluded any serious discussion of corporatism in labour relations. 53 The UNT representatives argued that the talks should be based on the “Twenty Commitments”, citing Mr. Fox’s endorsement of them, but the STPS backed the CT and the employers. 54 The ministry would have found it difficult to do otherwise, since it could hardly allow a minority of the MCD participants to dictate the agenda to the majority; however, this decision compromised its neutrality in the eyes of the UNT and its allies, who were increasingly at odds with the other parties in the MCD as the process unfolded. 55 The dispute exposed the limited scope for an agreement involving both sets of unions and the employers.

The UNT’s belief that its major concerns had been marginalised led to a split in the process and the emergence of rival proposals. 56 While the UNT never formally walked out of the MCD, it increasingly undermined it in public. Convinced that the reform proposals taking shape in the Mesa would do nothing to reduce the “official” unions’ control over their members and fearing that a proposal backed by both the CCE and the CT could pass Congress, the independent confederation and its allies opted for what one UNT leader called a “pre-emptive legislative strike against the CCE-CT proposal.” 57 A UNT-backed LFT reform bill was introduced by the PRD at the end of October 2002. 58 This led to the final breakdown of cooperation between the UNT and the other partners within the framework of the MCD, and a second initiative, based on the points agreed between the CCE and the CT in the Mesa, was introduced into Congress in December. 59 Though portrayed in the media as the Fox administration’s version of labour-law reform, it was not presented by the administration; the minister had undertaken not to send anything to Congress that had not been unanimously endorsed by the MCD. Instead, the bill was introduced as an initiative of PAN, PRI and Green (PVEM) deputies. Congress thus found itself confronted with two different labour-law reform bills (Box 12.1) rather than the single consensus proposal that the MCD was supposed to produce. The two bills were not wholly contradictory in principle – a compromise combining greater flexibility and the reform of labour relations was theoretically possible – but politically, they were all but mutually exclusive, because the agendas of their supporters were so different.

Policy process: adoption

In the event, Congress failed to approve either bill. Despite past PRD-PAN cooperation on labour-law issues and Mr. Fox’s endorsement of the “Twenty Commitments”, it was never likely that the administration would back the PRD-UNT bill, which had originated outside the MCD. To have done so would have been to disavow both the Mesa and the labour minister. However, the administration did not endorse the CCE-CT initiative either, nor did it initially seek to broker a compromise between the two bills. This lack of engagement helped seal the CCE-CT bill’s fate: some trade unionists and members of Congress expressed the view that strong public support from the president would have greatly enhanced its prospects. 60 With the executive unwilling to commit itself, the various congressional parties were reluctant to approve a reform of the LFT just before the 2003 mid-term elections, especially as the proposals in the CCE-CT draft were likely to provoke opposition. The reasons for the administration’s reluctance to support the CCE-CT proposal are not entirely clear, but it seems to have reflected the tensions between the PAN’s dislike of old-style “PRI-ist” corporatism and its commitment to labour market flexibility, which was likely to be easier to pursue within the existing framework of industrial relations. The CT-affiliated unions had, after all, accepted the NCL, with its emphasis on raising competitiveness, and had cooperated for
many years with economic policies that the “independent” trade unions vociferously opposed, including the negotiation of NAFTA and the development of the *maquiladora* sector. Nevertheless, some PAN Congressmen remained attached to the anti-corporatist agenda and objected to its absence from the CCE-CT draft.

---

**Box 12.1. Key features of the labour-law reform proposals**

The UNT-PRD bill’s major provisions included the abolition of the restrictions on union activity in the public sector in Article 123 of the Constitution; the liquidation of the *juntas* and transfer of their responsibilities to independent labour tribunals constituted as part of the judiciary; an end to state control over union registration; the creation of a National Public Registry of Unions and Collective Labour Contracts accessible to the general public; the establishment of a single national minimum wage and the creation of a National Minimum Wage Institute to replace the tripartite system of minimum-wage determination; the elimination of the “exclusion clause”; free, direct and secret1 voting by workers in union elections;2 and greater transparency of union finances. It largely failed to address issues of labour-market flexibility; indeed, certain of its provisions would have reduced that flexibility further.3

Given the prevailing balance of forces in Congress, the CCE-CT bill had far better prospects for adoption. Its main emphasis was on flexibility. It proposed, among other things: greater freedom for employers to use part-time and short-term contracts; greater flexibility with respect to working time for regular employees; the introduction of a maximum 30-day probationary period for most new employees;4 stronger protection against discrimination and sexual harassment; the introduction of new training contracts;5 and the modernisation of provisions related to training and skill development, in particular by allowing training and skill development to be provided by the regular educational system, and not only by trainers recognised by the STPS. The bill’s provisions for reforming the corporatist system of labour relations included provision for some streamlining of *junta* procedures and authorised secret ballots in union elections as a *possibility* rather than a requirement. It also contained provisions intended to curtail the use of white unions, something the CT wanted no less than the UNT. There were some measures aimed at increasing the transparency of union processes, particularly in regard to strikes, but these were largely concerned with new documentation and data requirements, including clear and complete lists of workers supporting the creation of new unions or calls for strikes. Critics argued that this would have made the identities of such workers known to both management and, in the case of unionisation drives, the official trade unions, exposing the workers to possible retaliation.6

---

1. The LFT allows for voting by a show of hands, a public procedure that union officials and employers have sometimes used to identify actual or potential opponents.
2. This provision encompassed union leadership elections, as well as votes on strike proposals and elections to determine which union would represent workers in collective bargaining.
3. The UNT-PRD proposal would have expanded regulations governing hiring and dismissal, in an effort to restrict “unilateral” employer activity, and would also have created an independent government organisation to regulate minimum wages and profit-sharing.
4. Up to 180 days for managers and directors, as well as for professional specialised technical positions
5. Workers hired under these contracts would receive an income even if they did not have previous work experience. The use of these contracts was to be regulated and the period worked under these types of contracts was to be considered as seniority if the labour relation continued.
6. For a more detailed look at this bill, see OECD (2004:94-95).

The PRI’s attitude to the CCE-CT bill may have been influenced by shifts in the position of the “official” unions. As the depth of opposition to reform within the labour movement became apparent, a number of the CT unions began to back away from the MCD process and to come out against the emerging proposals. The CT itself called for the elimination of probation periods, flexible working-time requirements and the right of
employers to insist that strikes be postponed pending negotiations. Both CT- and UNT-affiliated unions began organising protests, including a union-organised sit-in in the Chamber of Deputies on 1 December. In addition, the international NGO Human Rights Watch (HRW) issued a letter condemning the reforms that received extensive coverage in the Mexican press. This initial labour mobilisation prompted the PRI’s leaders in Congress to postpone consideration of the bill. The unions used this interval to organise further opposition. They enlisted the assistance of their US and Canadian counterparts, who sent letters of protest and filed a complaint under the North American Agreement on Labor Cooperation (the labour side agreement to NAFTA) that challenged the proposals as a violation of the terms of the trade agreement, as well as Mexican and international law. While the claim had little chance of success, it generated further negative publicity, as did a second HRW letter to Mexican legislators in February, and the bill, with none of the three major parties ready to back it, was shelved.63

A second round of negotiations followed in the early summer of 2003.64 At this point, the talks between the social partners were renewed, and this time the Congressional Committee on Labour and Social Welfare was included as well. In order to allow candid discussion, it was decided that the contents of these discussions should not be made public.65 The aim of these negotiations was to bring the autonomous union representatives back into the fold and outline a consensual, if minimal, LFT reform. Despite their relatively small size, the autonomous unions were regarded as important for the legitimacy of the process.66 The working document for these talks thus addressed two of the UNT’s core demands: secrecy in voting and the creation of a public registry of unions and collective agreements.67 In the end, however, the parties could not reach agreement on a compromise text: the CT publicly criticised what it saw as the disproportionate influence over the process that the Fox administration had granted to the UNT, and both Coparmex and the CT resisted any attempt to modify the essential elements of the CCE-CT draft. The compromise project was effectively dropped after the July 2003 mid-term elections, when large losses by the PAN and gains by both the PRD and the PRI brought about an even deeper cleavage in Congress on the question of labour-relations reform.

Conclusion

In hindsight, the failure of the MCD initiative to produce a reform had more to do with procedural issues than with the substance of the administration’s LFT reform proposals. The very design of the MCD process made it very difficult to secure agreement on anything of substance. While it would have been desirable to seek the broadest possible consensus on reform in such a contentious domain, complete unanimity among the parties was never likely, because they were not trying to solve the same problems. In other circumstances, such widely divergent agendas might have formed the basis for a complex bargain, but since some parties’ core concerns were effectively taboo for others, there could be no common understanding even on what the agenda for a discussion of labour-law reform might look like, let alone on the form that such a grand bargain might take. Either the UNT or the CT might have reached agreement with employers on some additional flexibility. However, the UNT was prepared to negotiate flexibility only in return for the elimination of what it regarded as the “official” unions’ corporatist privileges, while the CT was prepared to accept a degree of flexibility only in return for the preservation of the very legislative provisions the UNT sought to change.
These tensions found reflection in the political parties’ positions in Congress: the PRI would not support any weakening of the “official” unions, while the PRD would not accept any reform that did not undermine corporatism. The PRD had little interest in greater labour-market flexibility, and the PRI was divided over the issue. The PAN, for its part, confronted a tension between its anti-corporatist agenda and its economic policy priorities, including both wage moderation and structural reform. In the end, the latter were given priority, so the administration proved unwilling to compromise with the “independent” unions. Employers, too, seem to have been more comfortable with traditional corporatism than with the more uncertain labour-relations environment sought by the UNT. Yet the administration was unwilling to push for the CCE-CT bill, partly because of its conviction that UNT participation was important to the legitimacy of reform and partly because some within the PAN were more anti-corporatist than pro-flexibility.

The administration’s refusal to take the lead in the MCD process was particularly important because none of the other protagonists was strongly committed to reform. Apart from the UNT’s desire to break the “official” unions’ hegemony, the trade unions’ attitudes to labour-market reform ranged from suspicion to outright hostility, and it was not a top priority for employers. Administration leadership in this sphere was critical precisely because the costs of non-reform were largely opportunity costs and were thus politically “invisible”. There was no strong pro-reform constituency, because the major potential beneficiaries of reform were either disorganised and politically marginal (informal-sector workers) or did not exist (firms that had not entered the market). If the administration did not push for reform, it was never very likely than anyone else would. Yet the balance of forces in Congress meant that any pressure on legislators to adopt the reform proposals would have to be very deftly handled, lest it antagonise the congressional parties and generate a backlash against the executive.

Nevertheless, discussion of the question of labour-law reform has continued since the MCD process was wound up, and early 2009 witnessed a further attempt to advance a comprehensive reform package, this time in the context of the global economic crisis. As part of Mexico’s response to the crisis, a national accord concerning employment and the domestic economy was concluded at the beginning of the year by the federal executive, Congress, the governors of Mexico’s states and the social partners, supported by the political parties, the Bank of Mexico and other institutions. Among its major elements was a commitment to renewed efforts to update Mexico’s labour law in such a way as to promote both efficiency and equity, by encouraging employment creation in the formal sector, stimulating productivity and improving mechanisms for dispute resolution.68

Notes

1. These aims for the reform are set out in section 5.3.2 of Programa (2001); and the preface to Ruiz Angeles (2003).
2. No single party commanded a majority in either chamber.

4. Much depends on the definition of “informal”, as well as the measures used; see OECD (1997:72-73), (2004:31-32); (2005:25) and (2008:79-125). OECD research based on comparisons of actual social security contributions with what would be expected on the basis of wages and salaries in the national accounts suggest that about one-third of wage and salary payments escape social security contributions, including more than half in the private sector.

5. A 2001 Bancomer study ranks it 26th out of 36 in its sample. Botero et al (2003) rank it 75th among 85 countries studied (the bottom ten consist mainly of transition countries that had yet to re-write their labour legislation, even if a good deal of de facto liberalisation had taken place) and estimated that the cost of compliance with labour regulations was equivalent to 31% of wages. The World Bank in 2004 ranked Mexico 129th of 133 countries on the labour-market flexibility indicator of the Bank’s *Doing Business* rankings (Mexico was trailed only by Angola, Brazil, Panama and Portugal). With respect to employment protection legislation, in particular, Mexico in the early 2000s was third within the OECD, behind Portugal and Turkey.

6. For discussion of the corporatist system of labour relations that evolved under PRI rule, see Cook (1995, 1998); Middlebrook (1995); Patroni (2001); Mayer (2003); Cardoso (2004); Weyzig (2007); and Kohout (2008).

7. It regulates, inter alia, the nature and duration of work, holiday entitlement, compensation, termination of employment, child labour, harassment and discrimination.

8. See, for details, Cook (1995); Patroni (2001); Mayer (2003); and Burgess (2003). The notion of corporatism used here is that outlined by Schmitter (1974:93-94): “a system of interest representation in which the constituent units are organized into a limited number of singular, compulsory, non-competitive, hierarchically ordered and functionally differentiated categories, recognised or licensed (if not created) by the state and granted a deliberate representational monopoly within their respective categories in exchange for observing certain controls on their selection of leaders and articulation of demands and supports.”


10. Patroni (2001); Mayer (2003); and de la Vega (2004). In recent times, union representation has depended to some extent on the fortunes of the PRI: the number of union officials in the federal Congress fell from 52 in the mid-1990s (11 senators and 42 deputies) to 17 in 2000 (5 and 12, respectively), rising back to 23 (4 and 19) after the 2003 mid-terms.


12. Labour activists claim that employers are often able to take actions that leave the juntas little choice but to declare a strike illegal.


15. Few workers have the resources to retain counsel for such proceedings, but the LFT (Articles 530-36) provide for a Public Labour Defendant to grant free legal assistance to workers.

16. Ruiz Angeles (2003); Programa (2001), section 5.3.2.
17. In some cases, enterprises pay union officials a share of their profits or a share of the money saved under the contract. Weyzig (2007:63).

18. Known colloquially as *sindicatos charros* or “cowboy unions”.

19. On the phenomena of white unions, protection contracts, etc, see OECD (2004); Mayer (2003); Cardoso (2004); Fairris and Levine (2004); FIDH (2006); and Weyzig (2007).


23. Some analysts argued that the PRD was opposed not to corporatism *per se* but to the corporatist institutions created by the PRI, and that the PRD leadership had its own vision of an “alternative corporatism”; see Grayson (2007).

24. Estimates of Mexican unionisation rates vary widely, depending on whether they are limited to formal sector workers or to manufacturing. Fairris and Levine (2004) estimate that union density for the formal-sector workforce as a whole fell from just over 30% in the mid-1980s, to just under 20% in 2000. There is wide variation among sectors: some public sectors are almost wholly unionised, and unionisation rates has generally been high in medium-sized and large firms in both the private and public sectors. See also OECD (1997:97).

25. For many years, membership of official unions automatically meant being affiliated to the PRI.

26. The largest of these was the Confederation of Mexican Workers. Other major constituent unions are the Revolutionary Confederation of Workers and Peasants (CROC), the Federation of Public Service Workers (FSTSE), the Regional Confederation of Mexican Labour (CROM), and sectoral unions for rail transport, mining and metallurgy, petroleum and education.

27. See Bellin (2000); and Mayer (2003:87).

28. Among the major unions to affiliate with the UNT are the Union of Telephone Workers of the Mexican Republic (STRM), the Union of Workers of the National Autonomous University of Mexico (STUNAM), the National Union of Social Security Workers (SNTSS); the National Federation of University Unions (FNSU), and the Metropolitan Union of Public Transport Workers (SMTSTC).

29. In general, Coparmex has been more aggressive in staking out economically liberal, socially conservative positions. For more on the differences in structure and outlook between the two organisations, see Schneider (2002).


34. These included such sensitive steps as cutting subsidies for basic foodstuffs and liberalising land sales.
35. Indeed, the failure to undertake any significant reform of the LFT was seen by some scholars as an increasingly striking anomaly. See, e.g. Cook (1998) and Patroni (2001).
37. Real wage flexibility becomes more difficult in a low-inflation environment.
38. See especially OECD (1997) and (1999).
39. The STPS reports 264 such initiatives over 11 years.
43. On the PAN and PRD proposals of the late 1990s, see Cook (1998:16-17).
44. La Botz and Alexander (2005); FIDH (2006).
46. For a full description of representation on the MCD, see García Pineda (2006).
50. Four from the CT and two from the UNT.
53. A number of observers commented on the centrality of the NCL to the MCD discussions; see, inter alia, Cervantes (2002a, 2002b); Mayer (2003); Kohout (2008). The emphasis on the NCL was also reflected in official statements on the matter; see Segundo Informe (2002) and Tercer Informe (2003).
54. NCL official discourse.
55. See the interviews with union officials cited in Mayer (2003).
56. On the basis of interviews with 20 Mesa participants, Mayer (2003) notes that this assessment was shared by virtually all the members of the MCD, even if they differed sharply in their evaluations of UNT behaviour. UNT representatives argued that they had been invited into the process as window-dressing, to enhance its legitimacy, without any serious prospect of being given a real hearing. Their opponents argued that it had played a “double game”, participating in the talks inside the Mesa and the working group, while seeking to undermine them from without.
63. On the mobilisation against the bill and the interventions by HRW and US/Canadian unions, see La Botz and Alexander (2005).

64. Muñoz Rios (2003).


66. A number of participants in the renewed process expressed the opinion that UNT support for any reform proposal was critical to its legitimacy. See the interviews quoted in Mayer (2003:94).

67. Interviews with PAN Senator Francisco Fraile García and PRD Senator José Castro Cervantes, cited in Mayer (2003:94). This was particularly striking – and perhaps unrealistic – because the issue of a public registry of unions and collective agreements would have required a constitutional reform.

68. See Acuerdo (2009).
Bibliography


Chapter 13

The Netherlands: Reform of disability insurance, 2002-06

Summary and main conclusions

The Netherlands’ disability insurance (DI) system has long been recognised as among the most generous in the world. By 2000, around 11% of the country’s working-age population was drawing DI benefits. A major reform of the system was agreed by the government and the social partners in the early 2000s and took effect on 1 January 2006. The reform, which applied only to persons who suffered disability after 1 January 2004, was intended to reduce annual inflows into the DI scheme from 70-100 000 over the preceding decade to just 25 000 in an effort to focus benefits more effectively on those who could not work while strengthening the work incentives of those who could. Requirements for partial disability benefits were made more stringent, and some benefits were linked to work requirements and subjected to time limits. Those already receiving DI benefits before 2004 kept the benefits awarded under the old rules, although in some cases their disability status was reassessed under the new criteria introduced as part of the reform. Among the factors that made DI reform possible are:

- **Creation of a consensus on the need for reform.** By 2000, the scale and persistence of the DI problem meant that no one could seriously argue that the system was not broken, particularly with DI numbers headed towards the politically sensitive threshold of one million. A parliamentary inquiry in the mid-1990s had exposed the extent to which the problem was exacerbated by the social partners’ management of the system and had fundamentally changed the terms of the debate about DI reform. Throughout the 1990s and early 2000s, moreover, a growing body of research into DI issues created the basis for consensus on a number of aspects of the problem. The work of the cross-party “Donner committee” in 2000-01 then provided the basis for a set of proposals that could command wide acceptance. Finally, the economic slowdown of the early 2000s generated fiscal pressure for reform.

- **Lessons learned from previous reform attempts.** The first steps to slow the growth of DI benefits were taken in the early 1990s, when the government sought to curb incentives to use DI to facilitate dismissals. However, the improvement wrought by these reforms was temporary, as they were partially reversed by collective agreements between the social partners, as well as by a softening of government policy once DI inflow numbers began to fall. In the early 2000s, the government took special care to ensure that collective agreements would not subvert the reform and to reduce the prospects for a reversal of policy once things began to improve. The government also reduced the social partners’ role in administering the system, thereby reducing the risk that the reform would be diluted or undermined in implementation.
• Early reform measures that changed interest-groups’ incentives. A number of incremental steps taken in the late 1990s and early 2000s combined to change the incentives facing both employers and employees. Employers, in particular, became more open to fundamental reform because the early steps had forced them to bear more of the costs of the status quo directly.

• Government leadership. The basic outlines of the reform were defined by the government-appointed Donner committee and then debated in the Social and Economic Council (SER), the traditional forum for consultations involving business, labour and the state. However, government leadership was crucial throughout, and the reform ultimately went beyond the SER proposals.

The context for reform

Framework conditions for reform

The reform process began in 1999-2000 under a so-called “purple” coalition, bringing together the social democratic Labour Party (PvdA) and the economically liberal People’s Party for Freedom and Democracy (VVD), but the reform was adopted under a centre-right cabinet, comprising the centrist Christian-Democratic Appeal (CDA), the VVD and the “progressive liberal” Democrats 66 (D66). Despite differences between the parties over DI reform, there was a high degree of policy continuity throughout the period. The major effect of the change of government was to increase government cohesion on the issue: the purple government had been divided over DI reform, whereas the CDA-led cabinet was united on most of the major questions pertaining to the reform.3

The new DI reform proposals were presented at a time of slow growth and poor labour market performance, conditions which reinforced resistance to a reform that many believed would increase open unemployment and reduce benefits to a large number of inactive persons. However, the government pressed ahead with reform proposals, and their adoption coincided with the start of an upswing that appears to have made implementation of the reform easier.4 Though adopted in the context of a tight fiscal situation – the general government balance swung from a 1.5% surplus in 2000 to a 3.2% deficit in 2003 – the reform was not primarily motivated by fiscal concerns, nor was its design dictated by fiscal imperatives.

DI reform was not pursued in isolation. The Netherlands has introduced a number of other labour-market and social insurance reforms in recent years, many of them aimed at reducing the incentives for withdrawal from the labour force at a time of population ageing and falling labour-force growth.5 The new DI arrangements must be analysed in this context, particularly because negotiations with the trade unions in late 2004 involved the bundling of DI reform with a number of other measures. Moreover, the interaction of DI with other social insurances was an issue: the various measures for compensating income loss as a result of inactivity functioned at times as “communicating vessels”, with changes in each leading to unintended effects on the others.6

Dutch policy making tends to be highly inclusive and consultative, with considerable importance attached to reaching consensus on major issues.7 This can make for slow decision-making at times, but the emphasis on consensus helps reduce the risk that policy reforms will subsequently be reversed. Discussions involving organised labour and the employers’ organisations are largely centred on two institutions:
• The Social and Economic Council (SER) is a tripartite body, comprising an equal number of representatives of the crown, the employers’ bodies and the trade unions. Crown representatives are independent experts: though appointed by the government, they are not accountable to it. Dutch governments tend to ask the advice of the SER on all important planned socio-economic measures. The government is not obliged to heed this advice, but it is generally reluctant to go against a unanimous SER recommendation.

• The Labour Foundation, by contrast, is a bipartite body involving employers and trade unions only: there is no government representation on the foundation and it has no statutory authority. Nevertheless, it is influential, and the agreements reached during the twice-yearly consultations between the government and the foundation provide the basic framework for bargaining between the social partners as well as guidelines for policy.

In general, the SER is more of an advisory body, whereas the Labour Foundation is a bargaining forum in which deals are negotiated. Discussions in both fora tend to reflect a strong commitment to evidence-based policy and are often informed by the reports of special commissions and expert bodies, as well as research provided by public and private sector research bodies. The most influential by far is CPB Netherlands Bureau for Economic Policy Analysis, the government’s main economic analysis unit. The CPB’s analyses produced a large body of research on the Netherlands’ DI problems, which informed both public opinion and the discussions within the SER and other bodies.

Actors and interests

The ministry most concerned with DI reform was the Ministry of Social Affairs and Employment (MSZW), though the Ministry of Economic Affairs, the Ministry of Finance and others were also involved. The finance ministry was primarily concerned with the impact of rising DI expenditure on the fiscal balance, and most major proposals emanated from the MSZW. Here there was a high degree of policy continuity in the early 2000s, despite changes of government, owing largely to a general awareness of the need for change. All the major political parties recognised that DI reform was returning to the political agenda, and none wished to defend the status quo. However, the issue was particularly sensitive for the PvdA, given its close relationship to organised labour and its commitment to retaining an extensive welfare state. The liberal VVD, by contrast, was strongly committed to reform, which its business supporters wanted and which was likely to have less impact on the VVD’s largely middle class constituency than on the working-class voters of the PvdA. The CDA, too, recognised the need for reform, but the issue was quite delicate for the Christian Democrats, who had a working-class constituency of their own, as well as an affiliated labour-union federation. Moreover, the CDA had led the government that had implemented the last major DI reform, in the early 1990s, and had lost the ensuing elections in part because of hostility to the reform.

The principal union confederation, the Federation of Dutch Trade Unions (FNV), viewed DI reform with considerable suspicion, as did the smaller National Federation of Christian Trade Unions in the Netherlands (CNV). By 2000, both confederations recognised that reform was inevitable, but they were anxious to maintain key features of the system, including high replacement rates and indefinite duration of benefits, and they opposed what they saw as excessively harsh pressures for activation of the partially disabled. Both the VNO-NCW, the largest employers’ association in the Netherlands, and the Federation of Small and Medium-Sized Businesses (MKB) strongly supported DI
reform in the early 2000s, largely because a series of incremental changes in the preceding years had imposed new DI-related costs and responsibilities on them. Both wished to reverse some of those changes in return for backing new reform measures, and the MKB, in particular, wished to secure lower contributions for its members, which it believed were subject to an unreasonable burden under the arrangements prevailing at the end of the 1990s.

**Issue history**

The Netherlands’ 1967 Occupational Disability Insurance Act (WAO) was exceptionally generous by international standards and it was long used as a tool for labour-force adjustment. For most claimants, the WAO provided better benefits than the unemployment insurance (UI) system, and for employers it offered a way around restrictions on individual dismissals. In addition, DI offered older workers a guaranteed income until they reached pension age, which UI did not. Although empirical estimates of the extent of substitution varied widely, a large body of research suggested that DI was being used in an attempt to conceal unemployment on a significant scale. The number of DI benefit recipients thus rose dramatically from the late 1970s (Figure 13.1). This rising trend continued despite a drop in the official replacement rate and a reduction in the earnings base used to calculate DI benefits during the early 1980s. The result was a 25% cut in the real purchasing power of public DI benefits in the early 1980s. Although the evidence suggests that changes in replacement rates have a substantial effect on DI benefit enrolments, this was insufficient to check the growth of DI benefits, because the cuts were largely offset by collective agreements which provided for “top-up insurances” that supplemented benefits provided by the public scheme. The number of DI recipients passed the 900 000 mark in 1991, equivalent to around 11% of the working-age population. The possibility that this figure might top one million led to an upsurge in concern about the programme. While this threshold was somewhat arbitrary, it focused voters’ and policy-makers’ attention on the problem. Public discussion of DI reform tended to increase whenever the number of DI benefit recipients approached this level.

Figure 13.1. **Disability benefit recipients, 1970-2004**

![Graph showing disability benefit recipients from 1970 to 2004](source: OECD (2004), OECD Economic Surveys: Netherlands, OECD, Paris.)
Despite concern about the increasing incidence of DI claims in the early 1990s, there was strong resistance to further cuts in benefits. Proposals for further substantial retrenchment in 1991 triggered widespread protests that nearly brought down the government. In the end, rather more modest changes were adopted in 1993 and implemented from 1994 (Box 13.1). Policy reform thereafter focused on the less contentious issues of screening, reactivation and financing, rather than benefit adjustments. A high-profile parliamentary inquiry in 1993 exposed not only the extent to which the system was open to fraud and abuse but also the lack of interest in reform on the part of the social partners, who administered the system, or the government, which was focused on other policy objectives in its dealings with the social partners. The inquiry helped create a climate of public opinion more receptive to tighter screening. While the commission’s recommendations were not immediately enacted, its findings changed the tone of the debate.\(^{18}\) The years that followed saw the implementation of tougher entry criteria and of measures requiring employers to bear more of the burden of sickness insurance and DI. The introduction of the PEMBA system was particularly important. Though unpopular with Dutch business, PEMBA appealed to both parties in the “purple” coalition: its economic rationale appealed to the liberal VVD and its emphasis on employer responsibility was attractive to the social democratic PvdA. New incentives for employers to recruit and retain partially disabled workers were also introduced, along with stricter requirements for reintegration plans for workers qualifying for DI.

Convinced that legislative changes would have limited impact if the administration of the system did not change, the government also moved to reduce the role of the social partners in managing the system and to strengthen government oversight.\(^{19}\) New legislation also made employers responsible for sick pay during the first six weeks of a sickness spell and required them to do more to improve occupational health and rehabilitation/reintegration policies.\(^{20}\) In addition, some 335 000 disabled employees between the ages of 35 and 45 were reassessed during 1994-98 on the grounds of a revision in access criteria. Around half of those reassessed in the early stages of this effort lost benefits or had them reduced, but this prompted intervention by parliamentarians in an effort to soften the reform. As a result, the incidence of benefit adjustment fell rapidly over time, and many of those who lost benefits re-applied successfully later on.\(^{21}\)

The reforms of 1993-94 had a significant impact: the number of DI recipients fell from a peak of 951 000 in 1993 to 855 000 in 1996, as inflows dropped sharply. However, the number of beneficiaries began rising again in 1997, reaching 932 000 in early 2000. In part, the reversal of previous progress again reflected the use of collective agreements to offset changes in public benefits.\(^{22}\) Also important was the growth in partial disability benefits: partial disability had accounted for 10-20% of new cases at the beginning of the decade, but by the end of the 1990s this figure had reached around one-third, of whom about half were working.\(^{23}\) Partial DI was intended to help partially incapacitated individuals shift into part-time employment rather than exiting the labour market altogether. The evidence suggests that partial DI benefits did have such an effect, but they may also have led to an opposite labour-destruction effect in cases where the nature and degree of disability were not exactly observable, leading to self-selection. For large numbers of older workers, partial DI seems to have been a path to early retirement.\(^{24}\) However, sharply rising inflows among women and younger people accounted for most of the growth in DI cases. The incidence of DI inflows among women during 1998-2002 was almost double that of men, and a rising proportion of new DI benefit recipients were under 50.\(^{25}\)
Box 13.1. Occupational disability insurance in the Netherlands prior to 2000

Employees were eligible for benefits under the 1967 Occupational Disability Insurance Act (WAO) if they were deemed to be unable – fully or partially – to earn with customary labour the income of a comparable healthy person. The disability had to reduce earning capacity by only 15% for a benefit to be awarded – the lowest such threshold in the OECD.\(^1\) No distinction was made between work injury and general disability,\(^2\) nor was the decision to grant a benefit contingent on any assessment of likely recovery or ability to return to (part-time or adjusted) work. For individuals with full disability (loss of 80% or more of earning capacity), the act guaranteed replacement of 80% of their last earned income (subject to a maximum benefit level) for an indefinite period.\(^3\) Sector- or firm-level collective agreements meant that replacement rates were higher in most cases, often close to 100%. In the 1980s, the replacement rate for the public scheme was reduced to 70%, but collective agreements largely offset this cut, at least during the early stages of benefit receipt.\(^4\) One’s employment or contribution history was not a factor – a feature of the Dutch system that was unique in Europe. In 1994, the basic WAO benefit was split into two: a benefit based on previous income was maintained for a fixed period that was positively related to age and then replaced by a follow-on benefit based on the minimum wage.

Employees who received sickness benefits for over one year became eligible for DI benefits under the WAO. The benefit was reviewed after one year and every five years thereafter. Legislation adopted in 1996 required employers to pay 70% of an employee’s salary during the first year of sickness absence.\(^5\) From 1998, a firm’s contributions to the WAO scheme were linked to the number of its employees who moved onto DI benefits (the so-called “PEMBA” system, also known as “experience rating”).\(^6\) The 1998 legislation also allowed firms to opt out of the public scheme by taking on private insurance or assuming responsibility for statutory DI benefits themselves (self-insurance).\(^7\) All employers were required to enter into contracts with health and safety services.

DI was often used to facilitate labour-force adjustment. At the end of the 1990s, around 35% of new DI benefit claimants, including half of all new claimants under the age of 35, suffered from mental health problems (including stress). Back problems were the second most prevalent basis for disability by the early 2000s; together they accounted for around 60% of new awards.\(^8\) Both were difficult to diagnose and thus open to manipulation.\(^9\) Despite some benefit cuts and numerous partial reforms in the 1980s and 1990s, the WAO system in 2000 remained one of the most generous DI schemes in the world in terms of coverage and benefit levels, two of the most important predictors of the incidence of DI benefit expenditure across countries.\(^10\)

2. Van der Ploeg (2003) and de Gier et al. (2003:4) claim that this was unique among developed countries.
3. Both the indefinite term of the benefit and the use of an earnings-based calculation (rather than a flat-rate benefit or a calculation based on contributions) were unusual internationally.
4. The scope and generosity of top-up schemes tended to decline the longer an individual was on benefit. In 2002, collective agreements covered the first year of sickness for almost all sectors, bringing replacement rates up to 100%. For the second year, 86% of collective agreements provided for top-ups, bringing replacement rates in the affected sectors to 90% on average. However, only about one-third of collective agreements covered the third year and just under 6% covered the fourth. See Brenninkmeijer et al. (2003).
5. This had been increased from two weeks to six only two years earlier, in 1994. Employers were not obliged to pay wages if the employee had deliberately caused his illness, hindered his recovery or refused to accept an adjusted or alternative job.
6. See Brenninkmeijer et al. (2003) for details.
7. Deelen (2005) finds evidence of self-selection at work: high-risk firms, which received significant cross-subsidies under flat-rated DI contributions, tended to remain in the public system, while firms with better disability records tended to opt out. This put upward pressure on the public premium.
8. Brenninkmeijer et al. (2003); it is not clear to what extent these trends reflect changes in real health
status as opposed to changes in classification, since the late 1990s also see sharp drops in the categories “other” and “unknown”.
9. Van der Ploeg (2003) notes that many countries only acknowledge psychiatric problems as grounds for disability if they are the result of traumatic events.
10. See OECD (2003) and OECD (2004), Chapter 4, for details. Exploiting the variation in replacement rates arising from firm- and sector-level collective agreements, van Vuren and van Vuuren (2005) estimate that a 1 percentage point rise in the replacement rate implies an increase in the DI enrolment probability of 2.5%, even after controlling for the possibility of reverse causation (i.e. that workers with a high risk of disability will prefer high DI replacement rates).

These trends affected public perceptions of DI reform: the typical benefit recipient was increasingly a relatively young and apparently healthy person with a complaint that was not readily apparent, rather than – as in the past – an older manual worker with a clearly diagnosed work-related disability. Also important were changing attitudes among medical practitioners, who increasingly recognised the medical, social and psychological problems associated with labour-force withdrawal and hence the benefits of reintegrating the disabled.

Pressure for change thus began to build up again in the late 1990s, not least from the employers’ bodies. Dutch employers had previously tended to use DI to facilitate dismissals, and they had found collective agreements to offset benefit cuts a convenient device in negotiations with the unions, since the employer in such a situation effectively financed only a fraction of the benefit realised by the employee. Since 1996, however, they had been required to pay for the first year of sickness absence, and from 1998, they also felt the costs of the “experience-rated” disability insurance premia they paid under the PEMBA system. These changes, which significantly altered employer incentives in respect of DI, were especially a problem for SMEs, for whom the relative impact of just one DI benefit recipient on their experience rating was all the greater. Since the Dutch system did not distinguish between professional and social risks, employers felt they were increasingly liable for things that were not their fault and that they could not control. Experience-rated contributions do seem to have exerted a significant downward pressure on flows into DI, which would have risen even faster otherwise, but the PEMBA system was phased in over five years, so its full impact was not yet apparent when the reform debate resumed again in earnest in 2000.26

Policy process: formulation

The search for solutions

The re-emergence of DI as an issue generated tensions within the “purple” coalition.27 The coalition agreement contained no agreed policy on DI, since it had not been a particularly salient problem when the government was formed.28 In January 1999, the VVD State Secretary at the MSZW, who was responsible for the disability portfolio, put forward proposals for more frequent and more stringent assessments. The proposals triggered renewed debate, with government and opposition parties expressing a range of views concerning the relative emphasis to be placed on changes in benefits, activation measures and employer responsibility. The VVD and the PvdA, though partners in government, adopted very different stances, with the parliamentary wing of the VVD calling for even more radical changes than the MSZW had proposed and the PvdA rejecting the State Secretary’s plans. The VVD also proposed including work history as an element of the benefit calculation, a last idea that the PvdA rejected as a break with the fundamental character of DI as risk insurance.29
The social partners, too, adopted opposing positions, with employers’ organisations calling for cuts in both the level and duration of benefits and the labour unions rejecting both tougher screening and changes in benefits. The FNV trade union confederation emphasised instead the need for more reintegration measures and improved working conditions. The PvdA took a position close to that of the FNV, insisting that neither the level nor the duration of benefits was up for debate. However, the unions recognised that the status quo was increasingly difficult to defend and that some sort of change was in the offing. They therefore began to press for a discussion of the issue in the Labour Foundation. Some in the government worried about what sort of compromise the social partners might reach in a bilateral negotiation and what it might cost the state – the VVD, in particular, viewed such a prospect with concern. It was in part to forestall this that the government in June 2000 appointed a cross-party expert committee under Piet Hein Donner, a Christian Democratic lawyer and member of the Council of State with a mandate to make recommendations for DI reform. This set in train a complex policy process that, over the next four years, led to a series of significant changes in DI policy (Figure 13.2).

Figure 13.2. Reform of disability benefit, 2000-04

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Donner Committee appointed, reports</td>
<td>Donner proposals considered in the SER</td>
<td>Tripartite Agreement in Labour Foundation</td>
<td>Government proposes modified version of SER proposals</td>
<td>Autumn protests end in tripartite package</td>
</tr>
</tbody>
</table>

Major policy changes running in parallel

- 2002: Reorganisation of DB administration, Gatekeeper Act, reintegration contracted out to private firms
- 2003: Small firms exempted from PEMBA system, second year of sickness absence prior to DB award, introduction of "no-risk" system for reintegrating long-term sick or disabled
- 2004: Employer sick-pay liability extended from 12 to 24 months
- October 2004: stricter assessment rules for DB eligibility

The so-called “Donner II” committee included representatives of the major political parties – though the coalition parties were not represented by members of the government – and of the social partners. All those appointed had experience of dealing with disability issues. Nevertheless, the committee’s deliberations were largely guided by Professor Donner and the committee secretariat, which brought together a small number of leading Dutch specialists on the issue. The committee worked in secret, on the understanding that any recommendations must be unanimous and details of internal deliberations confidential. This was meant to shield individual members from outside political pressure. The committee’s main recommendations were announced in April 2001. They included limiting DI benefits under the WAO to fully and permanently disabled individuals, while making employers responsible for creating appropriate functions for temporarily or partially disabled employees. The committee also proposed imposing greater responsibility on the individual to assist his own reintegration and adopting a more flexible schedule for the assessment of disability, so that it would depend
on the nature of the problem rather than simply being mandatory after one year of sickness absence. These ideas implied big challenges for implementation agencies: research showed that implementing authorities, including examining physicians, often worked under considerable time pressure and with insufficient information. The social partners’ reactions to the Donner recommendations were broadly positive, as were those of the government and opposition parties. The most contentious issues concerned the level of benefit for lifelong disability – the unions called for an increase from 70 to at least 75% – and the treatment of the partially disabled.

Although critics saw the appointment of the Donner committee as an attempt by a divided government to avoid taking action before the elections, the coalition did proceed with changes to the administration of the system, merging the five industrial insurance institutions into a single organisation charged with running the system, the Institute for Employee Benefit Schemes (UWV), in January 2002. This further distanced the social partners from administration of the scheme and made subsequent reforms easier to implement. The other major innovation adopted in this period was the so-called “Gatekeeper Act”, which aimed to increase the incentives for both employers and employees, working with occupational health agencies, to activate and reintegrate employees in the first year of sickness if at all possible, since reintegration tended to grow more difficult the longer the employee was out of the workforce. The act laid out a series of steps that had to be taken to achieve this end and the deadlines for taking them. If employers failed in their obligations, they could be required to carry on paying sickness benefits for a further full year. Employees who did not cooperate in reintegration efforts could ultimately be dismissed. There was much scepticism about the likely impact of the act, even among its supporters: it was seen as a necessary measure, clarifying the roles and duties of employers and employees during periods of sickness absence, but it was not expected to have a dramatic impact. However, the evidence suggests that more intensive screening and monitoring did significantly reduce long-term sickness absenteeism and hence DI applications.

The SER proposals

The Donner proposals were passed to the Social and Economic Council (SER) for consideration in early 2001, and the SER reached agreement in January 2002. It followed Donner II on many points but departed from it on others. The SER plan aimed at reducing the annual inflow into the DI scheme from 70-100 000 over the previous decade to just 25 000. While existing WAO recipients (then numbering about 950 000) were not to be affected, new beneficiaries were to be split into three groups based on assessments by the staff doctors of the new Institute for Employee Benefit Schemes:

- People with an assessed disability of less than 35% would no longer receive benefits. Instead, the employer would have to adapt the workplace if necessary.
- People judged to be fully (80% or more) disabled would receive full disability benefit, equivalent to 75% of their average salary in the last three years of employment. During the first five years, the disability rate of this group was to be re-evaluated through an annual medical examination. Those falling below 80% would be integrated into the third group with an income guarantee of one year.
- People with a substantial capacity loss (35-79%) would be compensated for lost earnings on the condition of remaining in employment at a level corresponding to their assessed ability percentage (“sufficient remunerative work”). Partially
incapacitated individuals engaged in sufficient remunerative work were to be eligible for a wage-related benefit equal to 70% of the difference between their (maximum) daily wage and their work-related income. Both employer and employee would be required to do everything possible to find appropriate work for the employee. Should the employer fail to fulfil this obligation, the firm could be required to pay benefits for an extended period. If the employee fell short of meeting his commitment, dismissal might result. This wage-related benefit would last from 3 to 38 months, depending on the individual’s employment history. Thereafter, the individual would be entitled to a follow-on benefit or to a wage supplement, depending on his circumstances.

The second group would thus benefit from a public scheme while employers were to carry the costs for the first group and would either take private insurance or choose the UWV for the third group. Employers who succeeded in keeping a partially disabled employee at work were to be exempted from paying sick leave if the individual became ill again within five years (the so-called “no-risk” policy) and they were to receive a partial disability discount on social security contributions for such employees or for any partially disabled persons they hired. Employers whose employees had entered the WAO system would no longer be subject to higher contribution rates – the PEMBA system was to be scrapped.

Although the SER deliberations largely followed Donner II, a number of contentious elements of the Donner report were revised at the insistence of the social partners. The unions fought to preserve the system of partial disability benefits, in particular. The Donner committee had envisaged in-work benefits for the partially disabled and reliance on UI when they were out of work. This approach was informed by evidence that the conflict-avoidance behaviour of “gatekeepers” – both physicians and officials – tended to bias assessments in favour of higher-categories of disability. The unions rejected this, and the treatment of the 35-79% disability group in the SER proposals differed significantly from the Donner report. The FNV failed, however, to secure permanent compensation for loss for this group: benefits were to be time-limited in order to maintain reintegration incentives for the partially disabled. The unions’ other major success was in defending an earnings capacity definition of disability rather than a strictly medical “work capacity” definition, as proposed by the Donner committee. Employers were distinctly unenthusiastic about the proposal to make them pay for a second year of sick pay before workers could transition into DI. The unions criticised the prohibition on top-up insurances in the second-year of sickness, but this remained in the SER proposals, prompting the unions to call for revision of the agreement in the Labour Foundation.

Agreement in the SER was facilitated by the existence of a large body of research and a high degree of consensus among experts: there was a much higher level of agreement among the experts in the SER than among the constituencies they represented. Over 250 studies of the WAO and its implementation were issued during 1976-2000, the great majority of them appearing towards the end of the period, when around 20 studies per year were appearing. Most of this work was commissioned by one or more of the institutional stakeholders, a fact that often raised its profile in policy debates but that also sometimes ensured that the specification of research projects reflected the interests of particular stakeholders. In the early 2000s, the CPB, in particular, produced a steady stream of work on aspects of the disability insurance problem. The existence of such a large body of evidence did not produce automatic agreement among the parties: policy failures can reflect clashes of interest no less than lack of knowledge, and there were
indeed significant conflicts of interest over burden-sharing and the allocation of costs and responsibilities in relation to DI reform. The reception of research findings, which were often contested, depended on the political context. Nevertheless, the growing body of WAO research did mean that the negotiations were based on a set of commonly available and commonly accepted facts. Over time, such research appears to have been instrumental in changing agents’ views of the problem.43

Policy process: adoption

The compromises agreed within the SER were generally welcomed by the social partners and the political parties, but elements of the package did come in for criticism in different quarters. Among political parties the main governing party, the PvdA, and the main opposition party, the CDA, were the most positive; the VVD criticised both increased benefits for employees deemed fully disabled and the abolition of the PEMBA system. They argued that experience rating under PEMBA had helped to reduce DI inflows and that its elimination would undermine employer incentives. The support of the CDA for the SER plans proved important, because elections were held in May 2002, before the government could act on the recommendations, and the CDA emerged as the largest party in the new coalition, with the PvdA in opposition. The CDA had been elected on a platform of support for DI reform, and Prime Minister Jan-Peter Balkenende appointed a former official of the CNV trade union federation, Aart Jan de Geus, as Minister of Social Affairs and Employment with a clear mandate to pursue DI reform as his chief priority: with the CPB projecting WAO enrolments above 1 million in the coming years, the issue took on renewed urgency. Although the new coalition collapsed within three months and fresh elections were held in January 2003, the CDA and its liberal allies remained in government and preparatory work on the reform continued throughout the period. The new government moved rapidly on the proposal to extend employer responsibility for sick pay from one to two years, a move that dramatically cut DI inflows while the reform was being prepared.

Negotiations with the social partners led to some compromises on aspects of reform. The tripartite Autumn Accord negotiated in the Labour Foundation in late 2003 committed the social partners to refrain from topping up sickness benefits in the second year of absence44 and to working to reduce DI inflows to below 25 000 per annum. The government promised that if these conditions were met, then the DI replacement rate would be increased from 70 to 75%, which pleased the unions, and the experience rating of DI contributions would be abolished, something employers had long demanded.45 The government had already scrapped experience rating in respect of firms with fewer than 25 employees, for whom the PEMBA system was a particular burden.46

Nevertheless, new conflicts erupted when the government’s legislative proposals were presented to the social partners in March 2004. Though based on the SER agreement, they departed from it on a number of issues:

- *Occupational disability criteria.* The SER had proposed the concept of “long-term occupational disability” to cover both people for whom recovery was excluded on medical grounds and those for whom there was a limited chance of recovery in the long-term. The government opted to limit the provisions concerning “long-term occupational disability” to those with no prospect of recovery. The government’s proposed criterion also represented a step back towards the work-capacity approach of the Donner report, which the SER had dropped in favour of the earnings-capacity definition that had long been used.47
• **Re-evaluation of current DI benefit recipients.** The SER had agreed that current DI recipients under the age of 55 should be re-evaluated based on rules devised for assigning them to the three groups envisaged in the new system. However, the government plans included much more stringent assessment of remaining earnings capacity, which was expected to lead to partial or total loss of benefit for many individuals. With unemployment still rising, the unions questioned whether there would be jobs available for re-evaluated DI recipients. The government, in order to allay these concerns, introduced a temporary benefit (TRI) to replace the lost disability benefit for up to 12 months for those who were not eligible for unemployment benefit.48

• **Financing.** The old WAO system was to be split into two parts – one covering the long-term fully disabled (IVA) and one for temporary/partial disability (WGA). Employer contributions for the IVA were to be flat-rated, but the experience rating for WGA for those employers relying on the Institute for Employee Benefit Schemes (UWV) rather than private insurers was to increase. The extent of experience rating for employers opting to carry the risk of temporary or partial disability themselves would depend on the terms of their agreements with private insurers.

The conflicts over financing were largely inside the coalition. Employers found experience rating more acceptable for partial or temporary disability, about which they could in many cases do something (especially under the Gatekeeper Act), than for permanent and severe disability, where the costs were enormous and employers felt that there was little that they could do. However, the VVD and the CDA disagreed about whether or not insurance for the partially disabled should be wholly privatised, as sickness insurance had been, or left with the UWV.49 After months of wrangling, which at times appeared to threaten the fate of the reform, the two parties agreed on a compromise proposal that allowed firms themselves to choose between the public provider and private insurance.

The disability criteria and the re-evaluation issue were largely resolved as part of a complex bargain in November 2004. The autumn had seen a wave of strikes and demonstrations against government proposals on a range of issues, including early retirement, changes to UI, and DI reform, following the collapse of tripartite talks on the replacement of early retirement schemes with “life-span” leave arrangements. Preliminary talks with the unions and employers culminated in formal tripartite negotiations that produced a package agreement which softened a number of the planned reforms. On DI, the government agreed to limit re-assessments to those under 50 (rather than 55). The government stuck to its definition of permanent disability but gave up its much stricter work-capacity criterion in favour of the traditional earning-capacity definition. Finally, the definition of “sufficient remunerative work” was relaxed, so that the partially disabled would receive wage-related benefits on condition of remaining in employment for 50% of their assessed ability percentage, rather than the 100% implied by the original SER proposals.

These adjustments to the DI reform were exchanged for union concessions on other issues, including wage moderation. DI inflows had already begun falling sharply (Figure 13.1), so the cost of these concessions was not as great as it might have appeared just two years earlier. This owed much to the unexpected effects of the Gatekeeper Act: in 2003, the number of DI recipients fell for the first time in seven years, largely due to the decline in long-term sickness absence following the Act’s introduction the year...
before. Ironically, the lower inflows also seem to have reflected the full effects of the PEMBA system, which was to be scrapped.\textsuperscript{50} This improvement also reduced the apparent urgency of the new DI reform proposals, reinforcing opposition to them. By then, however, the government had invested too much in the reform to do anything but press ahead.

**Policy process: implementation**

The years immediately following the reform saw sharp falls in both inflows into DI and the total number of claimants: by 2008, the total stock of benefit recipients had fallen by roughly one-quarter from its 2002 peak.\textsuperscript{51} Initial assessments suggested that this was primarily a result of the reduction in long-term sickness cases following implementation of the Gatekeeper Act, as well as the second year of employer-financed sick pay: the most noticeable changes were not in the awarding of benefits to applicants but in reducing the number of DI applicants via better handling of long sickness spells. However, there were also significant changes attributable to the restructuring of benefits and the reassessment of existing recipients.\textsuperscript{52} The early stages of this process saw around 40\% of those reassessed either losing benefits altogether or being assigned to a lower category of disability; these initial reassessments focused on younger cohorts in the DI pool and on those most likely to be reintegrated.\textsuperscript{53} Such widespread adjustment of benefits led to mounting pressure to exempt older beneficiaries from being reassessed under the new criteria, and the government in due course accepted that the new, stricter classification criteria would apply only to new DI applicants born after 1 July 1959; those born between 1954 and 1959 would, however, remain subject to periodic reassessment under the old criteria.\textsuperscript{54} Thus, while the incidence of DI receipt in the Netherlands remained exceptionally high by international standards in 2008, it was clear that the changes of the early 2000s had contributed to a substantial improvement.\textsuperscript{55}

New challenges also emerged, particularly the reintegration of those long out of the workforce. Some of the least problematic cases were removed from the system relatively easily, but it was feared that many DI benefit recipients had been out of the labour force too long for a rapid return to employment. Active and effective re-integration support was therefore needed. In response to this, the government in 2006 further extended the “no-risk” hiring scheme; later, it prepared new plans for wage subsidies for disabled employees and additional premium discounts for employers hiring older DI benefit recipients. In addition, a new market in reintegration services began to emerge. A study examining the situation of those reassessed in 2005 found that 51\% were in employment 18 months after their reassessment, and a new study analysing the 2006 cohort reported that 62\% were working after 18 months. Former benefit recipients in the second study also reported receiving more active reintegration support from the UWV.\textsuperscript{56} Partial DI benefit recipients who found employment were increasingly working under temporary contracts (to circumvent sickness-related employment-protection legislation) or on contracts concluded under the government’s “no risk” provisions, which meant that the state bore the cost and the responsibility in the event of renewed illness during the first five years of the contract.\textsuperscript{57} While this reduced incentives for employers to discriminate against the partially disabled, the “no risk” policy, combined with the prevalence of temporary contracts, prompted concerns that an increasing number of employees were not covered by the various measures aimed at strengthening employer responsibilities.\textsuperscript{58}
Conclusion

In many respects, the DI system in 2000 was clearly “ripe for reform”. First, while there was much disagreement about specific issues, there was a widely shared recognition that the DI system was failing, and none of the major players wished to defend the existing system. While the trade unions were anxious to preserve generous benefits and resist what they saw as unacceptably harsh activation measures, they nevertheless accepted that the system was bound to change. Public opinion, too, had come to recognise the need for change, so the parties forming the government after the 2002 and 2003 elections could credibly claim a mandate for reform. Secondly, a series of incremental decisions had already begun to erode the existing DI regime and, crucially, to change the incentives facing the agents involved. These adjustments particularly affected employers, who increasingly bore the costs of the status quo. These two factors created a favourable basis for the negotiations among the principal stakeholders in the SER and the Labour Foundation. These negotiations were underpinned by a large body of commonly available and commonly accepted evidence, which also highlighted the lessons to be drawn from earlier attempts at DI reform and which formed the basis for consensus among experts on many of the issues under consideration. The close alignment between fiscal targets and reform further strengthened the pressure for change.

Despite this generally favourable context for reform, a successful reform was not a foregone conclusion. The issue was so sensitive in 2000 that it was difficult even to get it onto the government’s agenda, particularly given sharp differences of view within the purple coalition. However, that government’s decision to appoint the Donner commission, which helped form the basis for a broad cross-party consensus on the major issues, and its implementation of the Gatekeeper Act set the stage for a major reform effort under its successor. In the end, the Balkenende government opted to push beyond the SER recommendations, provoking a bruising conflict with the unions that was finally settled as part of a much larger tripartite agreement – an instance of “bundling ex post”. Yet in some respects, sequencing was as important as bundling, since prior administrative reforms, as well as the adjustments to employer responsibilities in the late 1990s, made it much easier to adopt and implement the restructuring of the WAO system. Finally, it is important to note that the compromises agreed during the three years of bargaining over the reform effectively protected the acquired rights of a large proportion of DI benefit recipients, so that the full impact of the reform was felt only by new DB applicants. This was perhaps the most important concession won by the unions in return for accepting other elements of the reform.

Notes

1. See OECD (2003); van der Ploeg (2003); de Gier et al. (2003).
2. Based on a definition of working-age as 15-64 years of age.
3. See van het Kaar (2002) on the tensions between the VVD and the PvdA over this issue.

4. In this respect, it resembles the US welfare reform: improving labour market conditions in both cases facilitated the implementation of reforms that tried to reactivate significant numbers of workers in part by reducing their benefits.

5. For an overview, see Visser (2007).


7. For the classic account of Dutch “consociationalism”, see Lijphardt (1968). More recently, see van Waarden (2002); Anderson (2007).

8. The president of the central bank and the director of the CPB Netherlands Bureau for Economic Policy Analysis are ex officio two of the eleven crown representatives at any given time.

9. This was obligatory until the mid-1990s.

10. See van der Ploeg (2003) and OECD (2003) for comparisons of the Dutch DI system with those of other European/OECD countries.

11. Overall, Dutch employment-protection legislation was only slightly stricter than the OECD average, but it was among the toughest in respect of protection against individual dismissals.

12. Empirical estimates of the “hidden unemployment component” of DI uptake in the 1980s and 1990s range from 10 to 50%. Even on relatively conservative assumptions, up to 38% of the inflow into DI in some years appears to have been hidden unemployment, though in other years the figure may have been as low as 2-5%. See, inter alia, Aarts and de Jong (1992); van Vuren and van Vuuren (2005); and Koning and van Vuuren (2006a and 2006b).

13. Actual replacement rates fell by a smaller margin, owing to the use of collective agreements to offset some cuts. De Jong (2007) estimates that the average after-tax replacement rate for a fully disabled worker fell by around 5 percentage points in the 1980s.


16. The prime minister of the day, Ruud Lubbers, threatened to resign if the number of DI recipients surpassed 1 million.

17. GAO (2008:51).

18. The inquiry in question was the so-called “Buurmeijer inquiry”, named for the chair of the Commission. Its prominence stemmed in part from the fact that parliamentary commissions were relatively rare in the Netherlands, being used primarily to investigate scandals. In addition, its hearings were televised. See de Gier et al. (2003:33) and GAO (2008:52).

19. See de Gier et al. (2003:26) for details.

21. Berendsen et al. (2007) estimate that around 5% of DI inflows in 2000-02 were re-entries. See also OECD (2008b).


27. See van het Kaar (1999, 2002).

28. When the second purple coalition was formed after the 1998 elections, it was not yet clear that the progress achieved after the reforms of the early 1990s was being reversed.

29. Under the proposal, entitlement was to increase on the basis of years of employment, reaching a maximum after 15 years. Only claimants with 40 years of employment or more would qualify for a higher entitlement.

30. For an overview of the January 1999 proposals and reactions to them across the political spectrum, see van het Kaar (1999).

31. The Council of State is an advisory body to the government established under the constitution, which gives advice on draft legislation to the cabinet before it is submitted to parliament.

32. Mr. Donner was already heading another committee (“Donner I”) to look at the increase in disability resulting from psychological problems.


34. Around one employer in eight with a worker reaching the two-year limit ends up paying sickness benefits for a longer period; OECD (2008b).

35. See Breninkmeijer (2003:6) for the entire schedule of actions defined by the act. Since 2003, employers also have to contract private reintegration firms to help a worker in finding a new job with another employer.

36. Van Vuren and van Vuuren (2005); and de Jong et al. (2006).

37. The maximum daily wage was defined as the median wage from at least 30 jobs representing the three best-paid functions the beneficiary could have performed before becoming disabled. The upper limit was set at EUR 37 000 per annum.

38. When such an individual was not working, the benefit would be equal to 70% of the (maximum) daily wage in the first stage and 70% of the minimum wage times the assessed degree of disability. The 70% subsidy rate is intended to avoid unemployment traps for disabled low-wage earners. A person previously earning the minimum wage who has lost 70% of his earning capacity can achieve an income equal to 79% of the minimum wage by working, as against 49% when not working (increased to 70% if the household has no other income).

39. If he were not working or not doing enough remunerative work, the partially incapacitated individual could claim a follow-on benefit equal to 70% of the statutory minimum wage multiplied by the percentage of incapacity. If he were doing sufficient
remunerative work, the wage supplement would equal 70% of the difference between
the (maximum) daily wage and his work-related income.

40. Definitions based on work capacity are much more common internationally; see
OECD (2004); and van der Ploeg (2003).


42. In terms of numbers of studies, the breakdown was as follows: 32% from commercial
research consultants, 30% from public research institutes and 38% from national
government, advisory and supervisory bodies, client organisations or the
implementation agencies themselves.

43. This is the conclusion reached by de Gier et al. (2003) and Aarts et al. (2002) in two
assessments of the impact of WAO research on policy.

44. In practice, this was achieved by stipulating that new DI recipients could not have
received more than 170% of their previous salary during their two years on sick pay.
This was very unpopular with the unions, who accepted it in the end in return for
other concessions.

46. Such firms now pay a flat 2.4% of the wage bill.

47. The government definition read as follows: “A fully and permanently disabled person
is someone who as a direct and medically objectively observable consequence of
sickness is permanently unable to earn income from work”. This definition was very
close to that recommended by Donner II. See OECD (2004).

49. The liberal VVD opposed privatisation, arguing that the transaction costs and loss of
economies of scale involved in a privatised scheme would outweigh the benefits of
private provision – a view that was reinforced by the first estimates of what the
private insurance premia might be, which were higher than the finance ministry had
expected. There was also concern about financing the transition from a pay-as-you-go
public system to private insurance, which would require the formation of adequate
reserves. The CDA and the Ministry of Economic Affairs nevertheless favoured
privatisation, as offering the best chance to ensure strict administration of the rules.

Koning (2004) finds that the experience rating system brought about a 15% reduction
in the DI inflow after one year. See also OECD (2004, 2008b).

51. OECD (2008b); GAO (2008).

52. Berendsen et al. (2007).


54. OECD (2008a).

55. OECD (2008a).

56. Van Deursen et al. (2007).


59. Ross (2007) refers to these two processes as **de-legitimation** (the policy regime comes to be seen as a contributory cause of the policy problem rather than a solution) and **de-institutionalisation** (incremental erosion of the programme over time as a result of piecemeal early reforms).
Bibliography


Chapter 14

Spain: Labour market reform, 1994 and 1997

Summary and main conclusions

In an effort to address longstanding structural rigidities in the labour market Spain substantially liberalised temporary employment contracts in the middle of the 1980s. However, nothing was done to relax employment protection legislation (EPL) for permanent workers. This led to a dramatic increase in the use of temporary contracts, which soon came to be recognised as a serious problem in its own right. In an effort to reduce this labour-market dualism, Spain adopted two reforms, in 1994 and 1997, that sought to diminish incentives to use temporary contracts and to make indefinite contracts less expensive, mainly by lowering dismissal costs. Both reforms were accompanied by changes to the system of collective bargaining. A number of factors contributed to governments’ ability to press ahead with these reforms:

- **An “employment crisis”**. With unemployment rising to almost 25% on the official figures in 1994, there was a strong consensus on the need for action. Such high levels of joblessness also put enormous strain on public finances. By 1997, unemployment was falling but it was still very high, as was the share of temporary contracts in total employment, so the issue remained a reform priority. It had loomed large in the 1996 election campaign. At the same time, the improving labour-market situation did facilitate reform in 1997, as rising employment made changes in EPL less threatening to workers.

- **EMU convergence pressures.** In 1994, the government argued that labour-market reforms could help rein in inflation and bring down public-sector deficits as Spain tried to qualify for the first wave of Economic and Monetary Union (EMU). In 1997, the link between EMU and labour-market reform was less about qualifying for EMU than about making the economy more flexible in order to cope with the disciplines of the single currency. Given the strength of public support for European integration, the connection between reform and EMU made it harder for the opponents of reform to win public support.

- **Government leadership.** The 1994 reform was adopted over the objections of the unions. The 1997 reform, by contrast, was negotiated by the social partners, with the government acting as facilitator. The government’s commitment to reform was important in both instances: in 1997, the unions’ willingness to bargain was affected by the experience of 1994 and the fear that the government might once again act unilaterally if agreement were not reached. Moreover, while unilateral action in 1994 created some problems at the implementation stage of the reform, it also contributed to a process of institutional learning. The positive results of
some elements of the 1994 reform contributed to a softening of union attitudes, which facilitated the success of the negotiations over the 1997 reform.

- **Shared objectives among the protagonists.** By the mid-1990s, the share of temporary employment had reached levels seen nowhere else in the OECD, creating problems for employers and workers alike, so both employers and trade unions wished to encourage the use of more indefinite contracts. The two sides also shared common ground with respect to collective bargaining reform. This created the basis for a bargaining process that was eventually extended to include more contentious issues, and the wide range of issues under discussion created the possibility for complex bargains involving concessions to different players on different issues.

- **Respect for acquired rights.** Although the reforms involved some reductions in EPL for a very large proportion of new indefinite contracts, the rights of those already employed under indefinite contracts were largely unaffected.

### The context for reform

#### Framework conditions for reform

Both the 1994 and 1997 reforms were adopted under minority administrations: the Socialist (PSOE) government of Felipe González had lost its absolute majority in Congress at the 1993 elections, and its successor, the centre-right Partido Popular (PP) cabinet of José Maria Aznar, had won power in 1996 without securing a congressional majority. Both governments therefore relied on the support of the Catalan nationalist Convergence and Unity (CiU) and other small regionalist and nationalist parties in order to govern. Since these smaller parties were relatively well disposed towards labour-market reform, the governments’ lack of congressional majorities did not turn out to be a key factor. The political context in 1997 differed from that of 1994 in two distinct but closely related respects. First, the PSOE government that pushed through the 1994 package had been in power for 12 years and was suffering from the gradual erosion of its support base. The PP government in 1997, by contrast, was relatively new, having won election the year before on a platform that included labour-market reform. Since the PP had never held power before, other actors were initially somewhat cautious in dealing with it, particularly the unions. Secondly, the two major trade union confederations had, in the period since 1993, changed their strategies for dealing with the government and the employers, adopting a less confrontational stance on a range of issues. Their greater readiness to negotiate in 1997 was partly the product of lessons drawn from earlier defeats.¹

The macroeconomic contexts for the two reforms could hardly have been more different. The 1994 reform was formulated in the midst of Spain’s deepest recession in a generation and adopted in the very first stages of the recovery; though growth had resumed, unemployment was still rising, reaching a peak of just over 24% on the official numbers, just a few months after the reform was adopted.² This was also a time when policy was largely driven by Spain’s determination to qualify for the first wave of European Economic and Monetary Union (EMU), an ambition that was threatened, in part, by the explosive growth of expenditure on unemployment benefits.³ By contrast, when the 1997 reform was adopted, Spain was in its fourth year of relatively strong growth, and unemployment, though still around 20% on the official figures, had been falling since late 1994. With a decision on EMU entry imminent, the government’s
medium-term programme was focused on the structural challenges that would face Spain after EMU entry, above all on the need to make the economy more flexible and to reduce the constraints on non-inflationary growth. Thus, the approach of monetary union strengthened the position of reform advocates and helped legitimate reform in both 1994 and 1997, though the primary drivers of reform in both cases appear to have been domestic.\(^4\)

**Actors and interests**

The PP’s position, as a party with strong support from the business community, was relatively straightforward: it favoured greater labour-market flexibility but did not wish to pursue that objective in a fashion that would provoke potentially costly labour strife. The PSOE’s position was much more complex. There was considerable suspicion of reform on the left of the party, but by 1993 economic policy was largely in the hands of the PSOE’s more economically reformist wing, the so-called “renovadores”, who had long occupied the principal posts in the Ministry of Economy and Finance (MEH).\(^6\) Since the party as a whole was strongly pro-European, reforms presented by the renovadores as necessary to prepare for the Single Market or EMU were hard for the left to oppose. Thus, despite the intense competition between the PSOE and the PP during the 1990s, neither the 1994 nor the 1997 reform was contested on party lines. In 1994, the PP was broadly receptive to the kind of changes proposed by the PSOE government, which were backed by the business community, and the 1997 reforms were the fruit of negotiations between unions and employers, with the PP government acting as a facilitator. The PSOE could hardly attack such a negotiated reform, especially insofar as it built on the Socialists’ own 1994 measures.

In the early 1990s, fewer than 20% of dependent workers in Spain were unionised. However, the trade unions’ influence was far greater than this figure might suggest, owing to exclusive jurisdiction rules, which limited rivalry among unions at firm level, statutory extension of wage agreements\(^7\) and public financing of trade unions.\(^8\) The unions’ principal goals in the reform debates of the mid-1990s were to prevent any erosion of EPL for regular workers, to press for measures that would promote the stability of employment and to secure changes in the system of collective bargaining that would strengthen their position *vis-à-vis* employers. While the unions were chiefly concerned with their core constituency – regular workers on indefinite contracts – they were not indifferent to the needs of those on fixed-term contracts, who by the early 1990s represented almost 35% of all employment.\(^9\) The major national-level trade union bodies, the Trade Confederation of Workers’ Commissions (CC.OO) and the General Union of Workers (UGT), managed, despite some differences of view, to maintain a united position *vis-à-vis* the government and employers.

The Spanish Confederation of Employers’ Organisations (CEOE), founded in 1978, is the principal peak-level association representing business interests, bringing together around 200 employers’ organisations, including the main small business lobby, the Spanish Confederation of Small and Medium Enterprises (CEPYME). Though consistently in favour of greater labour-market flexibility – and, in particular, of lower severance costs – employers attached tremendous importance to labour peace and wage moderation, and they were reluctant to jeopardise these objectives by pushing too hard for reforms and thus risking confrontation with the unions. The CEOE and CEPYME are generally regarded as having great political influence but little capacity to lead or discipline their members.\(^10\) Their influence reflects their involvement as representatives of Spanish business in national-level tripartite negotiations, while their limited authority
over their members is a product of their very loose internal structures: their constituent member organisations are very autonomous. Much the same may be said of the unions, albeit to a lesser extent: the regional, provincial and sectoral affiliates of the major confederations enjoy a high degree of autonomy and often pursue their own agendas. This complicated the reform of collective bargaining during the 1990s, because some agreements reached between the peak-level union confederations and employers’ organisations could not be implemented effectively without the co-operation of their territorial and sectoral affiliates.

**Issue history**

The 1980 Workers’ Statute (LET), which forms the legal basis for labour relations in Spain, provided very high levels of protection against both dismissals and occupational or geographical mobility. It also provided a legislative framework for collective bargaining (Box 14.1). Strong EPL for permanent workers coexisted with relatively limited coverage of unemployment insurance (UI), which was mainly targeted at long-term employees and left many of those on temporary contracts unprotected. This policy mix was consistent with a “breadwinner model”, the main goal of which was to protect permanent workers. The first major revisions to the LET concerned fixed-term contracts. The law initially allowed such contracts only for clearly temporary activities, as initial contracts for young workers and in circumstances where the government chose to introduce them as employment-promotion measures. The government first used this last option in 1982, when it introduced a new “Fixed-Term Employment-Promotion Contract” (CTFE), with a minimum duration of 6 months, renewable in increments of at least 6 months up to a total of 3 years. When it expired, the firm had to convert it to a permanent contract or pay the worker severance pay equivalent to 12 days’ wages per year of seniority. In the latter case, the firm could not hire another worker to do the same job, though this requirement proved relatively easy to evade in practice.

The introduction of the CTFE had little impact until 1984, when the government eliminated most of the restrictions on the use of fixed-term contracts, and a number of other temporary contracts, such as training and practice contracts, and contracts for new activities, were introduced. Employers at the time urged the government to adopt a more extensive reform, but since the government was at that point unwilling to alter conditions for permanent workers, it opted for increasing “flexibility at the margin”, a pattern seen in labour market reforms in many other OECD countries. This was an approach the unions could accept and, in some cases, even support, albeit reluctantly.

During the late 1980s, the Spanish economy grew rapidly and experienced job creation at rates unprecedented in its history, but rapid labour-force growth meant that unemployment rates nevertheless remained relatively high throughout the period. For a time, strong employment growth largely pushed the issue of flexibility off the policy agenda, but the period saw explosive growth in the use of fixed-term contracts. The share of temporary employment in total employment rose from around 8% in 1984 to just under 34% in 1992, almost triple the EU average of 13%. These developments, combined with a spurt in wage growth, prompted some to argue that collective bargaining had acquired an inflationary bias, since unions chiefly represented permanent workers, who perceived themselves to be protected by the “buffer” of temporary workers. Permanent workers appeared to earn around 10% more than temporary employees with similar profiles in similar jobs. In 1991, the government commissioned a study of the reform of collective labour contracts. The resulting proposals were not adopted, but the report focused
attention on the growing segmentation of the labour market and the costs of labour-market dualism, issues that became the focus of increasing public debate.

Box 14.1. Collective bargaining under the Workers’ Statute

The provisions of the Workers’ Statute that regulated collective bargaining were very flexible, leaving the parties to decide which precise structure to follow. Firm-level, provincial, regional and national agreements could be, and were, negotiated. The structure of bargaining tended to be dominated by provincial-level sectoral agreements, which were the determinants of de facto minimum wages in individual sectors that were well above the statutory minimum. Provincial-level sectoral bargaining likewise tended to be most important in regulating working hours and other conditions of employment. Firm-level agreements were also important, mainly in respect of very large firms. Because the law did not establish a hierarchy of agreements or specify matters to be negotiated at each level, there was a great deal of “cascading” negotiation, with the same issues (chiefly wages) being negotiated in respect of the same firms and sectors at different levels. Any negotiation under way had to take as a baseline any applicable collective agreements negotiated at other levels. In general, sector agreements were understood to establish minimum pay and other basic conditions, which could only be improved at firm level. Prior to 1987, economy-wide annual agreements did establish a “wage range” for collective agreements in specific sectors, but no such agreements were concluded between 1986 and 2002, leaving things much less centralised and allowing for more collectively bargained wage dispersion. Critics argued that this left Spain mid-way between centralisation and decentralisation in its collective bargaining, arguably the worst place to be.

2. Since 2002, the employers’ associations and the main union confederations have concluded annual “Interconfederal Agreements for Collective Bargaining”. These fix general criteria for collective bargaining at lower levels on a range of issues. They do not impose specific wage ranges but they do limit wage dispersion somewhat, by linking wages across sectors to the same inflation target and to sectoral productivity growth.

The evidence suggested that the temporary segment of the labour market was becoming more precarious as it grew in size, while the permanent core was becoming more secure as it shrank. The problem concerned employers, as well as trade unions, since such heavy reliance on temporary contracts generated high levels of uncertainty, reduced workers’ motivation and impaired human capital formation within firms. Dismissal costs for permanent employees, which were among the highest in the OECD (Box 14.2), were widely seen as the cornerstone of labour-market dualism. For most employers, the main attraction of fixed-term contracts was that they reduced adjustment costs when labour-shedding was required. A secondary motivation was internal flexibility: the regulatory provisions of the LET and many collective agreements meant that it was often easier to achieve internal flexibility by relying on short contracts than by altering conditions for permanent workers.

Contrary to expectations, the recession that followed the 1980s boom did not reduce the share of fixed-term contracts in total employment: as unemployment surged, two-thirds of employment losses corresponded to permanent employees and only one-third to temporary workers – roughly equal to their shares of the workforce by then. This had an important effect on the politics of labour-market reform, since it meant that, by 1993, the unemployed and temporary workers together outnumbered permanent employees: the “median voter” was now a labour-market “outsider”. Nevertheless, the
first major changes to the system of temporary contracts were prompted by budgetary
considerations: the cost of unemployment insurance (UI) skyrocketed during the
recession of the early 1990s, and the government responded with a substantial reform of
the UI system that, among other things, raised the eligibility threshold for UI benefits
from 6 to 12 months of contributions. This triggered a reform of the CTFE, which was
henceforth concluded for a minimum 12 months, renewable in 12-month increments to a
maximum of 36 months.

Box 14.2. Dismissal costs under the LET

Under the LET, firing costs depended on a worker’s seniority and the reasons for dismissal.
Individual dismissals could be for “objective” reasons (e.g. incompetence), “economic,
organisational or technical reasons” or disciplinary reasons. Individual dismissals for economic
reasons were permitted only in firms with fewer than 50 employees. Collective dismissals were
allowed only for “economic” reasons and required advance notice to the trade unions and
administrative approval. If an individual was being dismissed for non-disciplinary reasons, he
was entitled to 30 days’ notice in writing and 20 days’ pay per year of seniority up to a
maximum of one year’s pay, including bonuses and overtime pay. The worker could appeal his
dismissal to a labour court, which could find the dismissal fair, unfair or void. If the dismissal
was found to be void, the worker was reinstated with back pay; if it was judged to be unfair, the
worker received 45 days’ pay per year of service up to a maximum of 42 months’ wages. The
employer continued to pay the worker during the appeals process (so-called “procedural
wages”) (Bentolila et al., 2008). Since the criteria for dismissal were construed very narrowly
by the courts, workers tended to win 75-80% of appeals (OECD, 1994), and many employers
simply reached agreements with workers to pay fairly generous compensation rather than
fighting cases before the courts.

Policy process: the 1994 reform

Formulation of the reform

With unemployment rising to unprecedented levels, the government in early 1993
presented a “Green Paper” on labour market reform, with a view to discussing it with
the social partners in the Economic and Social Council (CES), which brought together
representatives of trade unions and employers’ organisations, as well as
government-appointed experts and representatives of other social groups. The paper
provided an analysis of labour market issues and a menu of possible reforms to serve as
the basis for discussion. However, before it could be discussed, the government called an
early election and the green paper was shelved. It was clear nevertheless that labour-
market reform would be among the top priorities for whatever government emerged from
the June 1993 poll. In the event, the PSOE won the election but lost its majority in
Congress and found itself reliant for support on the CiU, which was somewhat more
liberal in its view of labour-market reform. At this point, the renovadores took control
of the labour ministry. While the outcome of the reform effort did not depend on the
identity of any one minister, some analyses suggest that this changeover did help to
ensure smooth co-operation between the MEH and the labour ministry.

In August-September 1993, senior government officials held a number of meetings
with representatives of Spanish business organisations, including the employers’
federation CEOE and the Círculo de Empresarios. The business groups made a concerted
effort to convince the government of the necessity of reform. This reflected both
immediate concerns – the recession had made them keener than ever to gain greater
internal flexibility – and a longer-term conviction that they would be hard-pressed to maintain their competitiveness in an integrating European market if labour-market rigidities were not addressed.\textsuperscript{30} The employers were also anxious to regularise the position of the 100 or so temporary work agencies (ETTs) then operating, whose activities remained legally questionable under existing legislation. The PSOE’s economic policy makers were receptive to these concerns, and many of the employers’ proposals were close to the government’s own, which were put forward in late September, replacing the Green Paper.\textsuperscript{31} The renovadores worked to persuade both the party and the public that such reforms would help Spain qualify for EMU: they would facilitate wage restraint and thereby hold down inflation, while reducing the cost of hiring would help reduce unemployment and thus ease the burden on the social welfare system. More generally, they argued that greater labour-market flexibility was required to increase overall competitiveness in the run-up to EMU.

\textit{The government reform proposals}

The reform proposals that the government put forward were wide-ranging and left little of the LET untouched. The major provisions tended to shift the balance of power in the workplace towards employers:

- \textit{Dismissal procedures} were to be restructured in an effort to reduce costs, and the grounds for “justified” individual and collective dismissals were expanded and clarified. Procedures for collective dismissals were to be streamlined somewhat.\textsuperscript{32} While these changes were limited, and the rules governing severance pay remained unchanged, this marked an important watershed, insofar as the reform attempted to alter conditions for regular permanent workers – i.e., for labour-market “insiders.”

- Major changes were also proposed for the regulation of working time, in an effort to allow firms greater internal flexibility. In addition, overtime hours were no longer to be paid at an obligatory 50\% premium; instead, overtime pay would be set by collective bargaining.\textsuperscript{33}

- A new \textit{apprenticeship contract} of 6 to 36 months’ duration was to be introduced for 16-25 year-olds without post-secondary education. The minimum wage for these contracts was to be set at 60, 70 and 80\% of the legal minimum in the first, second and third years, respectively; employers would not pay UI contributions for such workers, nor would the workers be eligible for unemployment benefits. This was politically bold for a minority Socialist government: five years earlier, a general strike had blocked a much more modest proposal to introduce cheaper contracts for people under 25 without work experience.

- In respect of collective bargaining, the government proposed a process of decentralisation and strengthening the role of collective bargaining in the regulation of labour relations. The crucial innovation was an amendment to the LET allowing lower-level collective agreements to include better or worse conditions than those agreed at a higher level; previously, lower-level bargaining could only improve on what had been agreed at higher levels. The old Franco-era “Labour Ordinances” that still regulated many sectors were to be abolished and the social partners were to replace them with collective agreements.\textsuperscript{34}

The government also proposed to legalise the ETTs. There was, however, one significant concession to the unions: the CTFE, which accounted for around one-third of
temporary hires by the early 1990s, was to be effectively eliminated, remaining an option only for employing the long-term unemployed aged 45 or older. This was effectively a quid pro quo for the simplification of dismissal procedures. The unions had long objected to the use of the CTFE to circumvent all other restrictions on the use of temporary contracts. Its abolition meant that fixed-term contracts would again have to be justified by the content or temporary nature of the job in question, the characteristics of the employee (e.g. apprenticeship contracts) or other specific circumstances (e.g. temporary contracts for workers employed in “new activities”). The legislation also lifted the minimum limit on the duration of temporary contracts, allowing employers to roll them over more frequently.

Adoption: a unilateral reform

The government hoped to secure the social partners’ approval of the proposals within the framework of consultations in the CES, but the unions were reluctant to cooperate. Though the UGT was historically affiliated to the PSOE, its leaders had long been opposed to the party leadership, and both the CC.OO and the UGT had clashed with the government repeatedly over the preceding decade, resisting its austerity and deregulation policies. When the September proposals were put forward, the UGT and the CC.OO, who were reportedly unaware of the discussions between government officials and employers’ representatives over the summer, initially refused to consider them. They eventually relented and briefly took part in the CES discussions, but they quickly rejected the major elements of the reform – the new apprenticeship contracts, the changes in dismissal procedures, and the legalisation of ETTs, which the unions claimed would lead to the privatisation of employment services. With unemployment high and rising, the value of EPL to those who had it was arguably greater than ever, so even the very modest reductions in dismissal costs proposed by the government were seen as threatening to permanent workers. The unions also viewed the proposed reform of collective bargaining as a process of “disorganised” decentralisation that was decidedly pro-employer.

The government’s position was straightforward: while it had promised during the spring election campaign to negotiate a social pact concerning measures to tackle unemployment, it held that the unions’ refusal to participate in the CES process in the midst of an economic crisis left it no choice but to act without them. The government therefore approved the reform in December 1993, involving only minimal adjustments to the September proposals. In response, the unions called a general strike for 27 January 1994 – their third in six years. However, turnout was low (below 30% of the workforce), and the unions thereafter offered little resistance to the reforms. The UGT was in any case in the midst of a financial scandal and change of leader, which reduced both its authority and its effectiveness, and the CC.OO was internally divided over strategy. This made things easier for the government, which mounted a strong public relations effort linking the reforms to EMU accession. The European connection put organised labour in an awkward position: some union leaders later acknowledged that they could not win the battle for public opinion once the government had convinced much of the electorate that the reform was needed to help ensure early Spanish entry into EMU. The legislation was confirmed by Congress virtually unchanged: the main concession was to raise the minimum wage for apprenticeship contracts to 70, 80 and 90% of the legal minimum in the first, second and third years of the contract.
**Bargaining in the implementation phase**

Faced with the adoption of a collective bargaining reform that they had opposed, the unions threatened to adopt a very aggressive stance in coming wage rounds. Anxious to avoid escalating conflicts over wages, the employers reached a tacit agreement with them: in return for wage restraint and union moderation, employers would not exploit those elements of the collective bargaining reform that were most troubling to organised labour – specifically, those allowing derogation from multi-employer agreements and greater flexibility in respect of working conditions. In addition, both sides committed to further efforts to reform the collective bargaining system in a fashion that was mutually acceptable. This was a significant gain for the unions. It meant that the reform did not link wages as closely to firm-level productivity as had been envisaged. However, the unions’ new readiness to bargain and the employers’ desire to maintain labour peace and wage moderation bore fruit in other ways, as the two sides concluded the first partial agreements on replacing the old Labour Ordinances and then signed a bipartite agreement on a mechanism for resolving labour-relations conflicts outside the courts. In early 1995, the social partners and the government reached agreement on the broad lines of pension reform in the so-called “Toledo Pact.”

Although employment growth in 1995-97 proved extremely robust as the recovery took hold, the reform had only a marginal impact on the deep duality of the labour market. There was rapid growth in both temporary and permanent employment, and labour-market turnover, measured in terms of hirings and separations, rose. The proportion of the workforce on fixed-term contracts remained more or less constant at around one-third, and the average duration of those contracts fell. The share of workers with less than a year of tenure in their jobs reached almost 35%. The new restrictions on the use of the CTFE simply led employers to resort to other temporary contracts, apparently with the tacit consent of the employment offices in many cases; indeed, the knowledge that these restrictions could be circumvented may well have contributed to employers’ willingness to accept them. The proportion of employees with less than a year of tenure who were still employed one year later, which had fallen dramatically during 1990-94, recovered as growth picked up but had risen only to about 18% by 1997. What was less clear was whether or not those workers who entered the “secondary” labour market were becoming trapped there. Conversion rates into permanent contracts were very low by the mid-1990s – around 5% in 1994, down from 18% in 1987. They then recovered somewhat, but remained in the 10-15% range even during the period of very fast growth in the mid-1990s. However, the average age of temporary workers was more stable than that of permanent workers during 1992-97, implying considerable entry and exit in the former group.

**Policy process: the 1997 reform**

**Formulation and adoption: a negotiated reform**

The PP government that took office after the 1996 elections was in a good position to pursue a negotiated reform, despite the fact that it did not command a congressional majority. Unemployment had been a major issue in the campaign, and the PP had promised to pursue further labour-market reforms if elected. However, it had avoided making specific proposals, stressing instead its intention to negotiate any reform with the social partners, in an effort to allay the fears of many on the left about what the election of a right-of-centre government – the first to win election since Spain’s democratic
transition – might bring. The party was anxious to demonstrate that it could govern effectively and without undermining labour peace, so it adopted a very moderate stance. The unions, for their part, were also ready to bargain. The UGT had undergone a change of leader and both confederations were pursuing less confrontational strategies vis-à-vis the government and employers than in the early 1990s. Moreover, the effects of the 1994 reforms had not been as the unions had feared, and they had begun to accept some of the innovations, such as the operation of ETT’s on the market. Recalling the 1994 reform, the confederations were anxious that the new government should not act unilaterally, as its PSOE predecessor had done.54 While the government strongly expressed its commitment to concertation, there remained the possibility that it would act unilaterally if the talks failed. Such a course of action would have been costly for the government and the unions alike.

Informal negotiations began in June 1996, with more formal talks following in January 1997. Some issues were negotiated more or less exclusively between the social partners, with the government informed of, but not directly involved in, the talks; other issues required tripartite discussions. Initially, the discussions focused on the issues that both unions and employers recognised as problems – the structure of collective bargaining and the extremely high level of temporary employment. As the talks proceeded, however, the unions’ concern with the job insecurity associated with temporary contracts led them to negotiate on compensation for dismissals under permanent contracts, as a quid pro quo for tighter regulation of temporary work. Dismissal costs for permanent workers had previously been a taboo subject, and it would have been difficult for the unions to come to the table in the first place if severance pay had been on the agenda at the start. In the end, the two sides agreed on a trade-off, involving some reduction of dismissal costs for some permanent contracts, combined with measures to raise the costs of temporary contracts. These negotiations were facilitated by the on-going economic recovery, since with employment growing fast, few employees faced any immediate negative impact from reform.55

Agreements and legislation

In April 1997, after ten months of talks, the unions and the employers signed the “Labour Agreement for Employment”, which in fact comprised three separate agreements.

- The centrepiece of the reform was the agreement on employment stability. It provided for a new permanent contract – the so-called “star contract” – targeted at those most exposed to unemployment (younger workers, older men, women in certain sectors and the disabled) and at workers on temporary contracts who converted to indefinite ones within one year of the implementation of the relevant legislation.56 The new contract involved substantially lower severance payments for unjustified dismissals: 33 days per year of service, up to a maximum of 24 months’ pay, as compared with the figures of 45 and 42, respectively, then prevailing. It also sought to clarify when employers were justified in laying off workers on permanent contracts, since justified dismissals entailed even lower severance payments. The new contract was to be used for four years and then reviewed in 2001.

- The agreement on collective bargaining aimed to reduce disconnected and uncoordinated bargaining by defining more clearly the “most appropriate” level of negotiations for specific contract issues, in an effort to reduce the strong
inflationary bias of the existing system. Greater reliance on bargaining at lower levels was to be preferred, since this would permit wage settlements better to reflect specific labour market conditions.

- The third agreement dealt with the regulation of working conditions in those sectors where collective agreements had not yet filled the gaps left by the abolition of the old Labour Ordinances following the 1994 reform.57

The unions’ acceptance of a new indefinite contract involving lower severance payments – even one available only to specific groups – was an important milestone in the reform process. It represented the first time that organised labour in Spain had accepted the idea that a reduction in EPL for indefinite contracts might have to be part of the solution to the problem of labour-market dualism. However, incumbent workers on regular contracts experienced no change in their situation – their acquired rights were respected. This meant that the reform produced no clearly identifiable “losers.”58 In addition, the unions obtained a number of concessions:

- To encourage the use of the new permanent contracts and discourage the use of temporary ones, the government undertook to reduce employers’ social security contributions under the new star contracts. During the first two years of the contract, these contributions were to be cut by 40%, rising to 60% for the over-45s and the disabled. Similar incentives applied to new regular indefinite contracts, albeit only under certain conditions.

- Stricter conditions were imposed on the use of temporary contracts. Temporary contracts for “new activities”, previously the subject of much abuse,59 were eliminated, as were some others, and their maximum length shortened. The government was authorised to subject some types of fixed-term contracts to a higher contribution rate for unemployment insurance.

- Collective bargaining reform was something the unions very much wanted: employers had little interest in modifying the existing collective bargaining system, which, though conflict-prone, put them in a strong position vis-à-vis the unions.60 In particular, the unions secured confirmation of the pre-1994 arrangements, under which national sectoral agreements set sectoral minimum wages and defined overall wage structures; lower-level negotiations could improve on the terms agreed at higher levels but could not make them worse.61

- Employers agreed to begin discussions on the transition to a 35-hour work week.

- The “sunset clause” authorising the new star contract only until an evaluation in 2001 further allayed union concerns about the innovations being adopted.

The adoption of the 1997 reforms was an entirely straightforward affair. Since they had been agreed by the social partners, with the government’s involvement, there was no significant opposition to the passage of the relevant legislation through Congress, although some within the union movement continued to express reservations about the agreements that had been reached.

**Implementation of the 1997 reforms: uneven progress**

The initial impact of the contract reform was quite pronounced: the share of indefinite contracts in all new contracts concluded jumped from 4% in 1996 to 9% in the second
half of 1997. However, there were probably significant deadweight losses involved, since a part of this change consisted of employers registering existing indefinite contracts with the authorities in order to secure lower social security contributions, and the reduced social charges also applied to some new permanent contracts that would have been concluded even in the absence of reform. Nevertheless, ex post analyses found the effects of the reform on the awarding of permanent contracts to be both significant and positive, if not dramatic. They also found that the social security rebates were a very important part of the picture, suggesting that there might have been little change otherwise. Under the structure of incentives created, permanent contracts were cheaper over a two-year horizon given the tax breaks and dismissal costs involved. The years following the 1997 reforms saw further strong, employment-intensive growth, with real GDP rising by 3.5-4.0% per year and employment growing at 2.5-3.0% rates. While the proportion of workers on fixed-term contracts fell back only marginally after the reform, it did not rise as it had during the previous period of intensive job-creation, implying that the economy was creating permanent jobs much faster than it had a decade earlier. By the early 2000s, around 40% of workers on indefinite contracts – or one-quarter of all employees – were on variants of the “star” contract.

Implementation of the collective bargaining reform proved very difficult and made only limited headway, largely because of resistance from within both the trade union movement and the employers’ organisations. What the national-level confederations had agreed was in essence a model bargaining structure that agents actually engaged in collective bargaining could adopt, modify or ignore. In the event, unions and employers in most sectors chose the last option, largely because agents at every level resisted the transfer of bargaining authority to other levels. This was particularly true of provincial-level sectoral unions and employers’ organisations, who had long occupied the key rung in the collective bargaining ladder. It would have been very difficult for the trade union confederations to impose a more structured arrangement from above, and virtually impossible for the CEOE and CEPYME to do so in respect of employers. Employers’ organisations were simply too weakly structured for this, and in some cases, a substantial reorganisation of the employers’ bodies would have been needed to make the new system work. Even so, the implications for wage determination were significant, since the unions accepted much wider use of variable, productivity-related pay in return for clearer regulation under national sectoral agreements.

**Feedback and adjustment**

The 1997 agreement provided for an assessment and revision of the reforms in 2001. Further reform of both labour contracts and collective bargaining thus figured in the PP’s electoral platform at the 2000 elections, and the re-elected PP government sought to extend the 1997 reforms. The star contract, which was due to expire (for new hires), was to be re-authorised, as were the fiscal incentives for using it. Eligibility for the star contract would be extended to prime-age males who had been unemployed for more than six months, rather than one year, as under the previous rules. In an effort to cut dismissal costs further, the government proposed reducing the compensation paid for unfair dismissals under all new permanent contracts from 45 days’ pay per year worked to 33 and the maximum payable from 42 to 24 months’ pay. In addition, employers were to be given greater freedom to set shifts and hours for part-timers, to alter job descriptions and hours of work. Finally, the government planned to adopt a law defining the issues to be negotiated at each level of the collective bargaining system. This amounted to an attempt
to impose a modified version of the structure agreed in April 1997, which lower-level labour and employers’ organisations had resisted in practice.

This time, the unions took a tougher stand: while the CC.OO was more open to negotiate than the UGT, it was unwilling to conclude a deal without the UGT, and both confederations objected to the fact that the government put forward its own proposals before the talks began. The unions maintained that this prejudiced the negotiations. When it became clear that there would be no new pact involving the social partners, the government used its congressional majority to legislate the extension and modification of the star contract and the abolition of procedural wages. However, it dropped the proposed legislation on collective bargaining, which met with hostility from both employers and trade unions. The government proposals would have limited provincial agreements, in particular, to a very minor role. In 2002, the government sought to reform the unemployment benefit system and abolish “procedural wages”. This provoked further union resistance, including a general strike. In the end, the changes to unemployment benefits were substantially softened, and the principal change adopted was the abolition of procedural wages.

The unions’ hostile reaction to the 2001 reform was motivated by its unilateral character, as well as its substance. The unions took the view that any major reform had to be negotiated with the social partners in order to be fully legitimate and claimed that the government had gone back on a promise not to use its parliamentary majority to bypass the unions and the opposition. The unions might nevertheless have agreed to the reform in return for concessions concerning the use of temporary contracts and moves toward a 35-hour work-week, an issue on which the unions had achieved very little progress in the wake of the 1997 agreements. However, the only concession granted was compensation on dismissal of temporary workers of 8 days’ pay for every year worked.

The PSOE government elected in 2004 undertook a further EPL reform in 2006, following several years of negotiation. This latest reform was wide-ranging and included a number of innovations, three of which were directed at reducing labour-market duality. First, workers who have had at least two consecutive contracts in the same post, covering more than 24 months within a period of 30 months, are entitled to a permanent contract. Secondly, the financial incentives in favour of permanent contracts with low redundancy costs were modified. The period during which employers paid reduced social insurance contributions in respect of these contracts was extended from two to four years (or the entire duration of the job for those over 45), but these subsidies were redefined on a flat-rate basis, in order to favour those on lower wages. Such incentives were also extended to cover new groups of workers, specifically unemployed men between 16 and 30 years of age. Thirdly, as a one-off measure, the government offered an annual subsidy of EUR 800 for three years for all temporary contracts converted into permanent contracts before the end of 2006; this included workers aged 30-45 who did not benefit from subsidised contracts. Temporary contract conversions concluded before end-2007 were made eligible for lower redundancy costs (33 days of wages per year of service). No further moves to reduce dismissal costs were adopted, as the unions continued strongly to oppose any additional steps in this direction. The initial impact of this reform appears to have been significant – the temporary employment rate fell by roughly five percentage points in under two years. However, most of this change took place during the transition period and it is too early to assess its longer-term impact, particularly given that by mid-2008, labour-market performance was being affected by the onset of an economic slowdown.
Conclusion

Clearly, labour-market dualism remains a challenge for Spain; the temporary employment rate remains stubbornly high and labour-market reforms have been on the agenda of every government in recent years. Nevertheless, the reforms of the 1990s, as well as subsequent initiatives, do appear to have helped to stimulate the creation of both permanent and temporary employment, and to reduce long-term unemployment, particularly in periods of strong growth. Despite their limitations, the Spanish reforms helped to move the focus of labour-market policy somewhat from “flexibility at the margin”, with its emphasis on liberalising conditions for labour-market outsiders, towards adjusting conditions for permanent workers as well, in the interests of equity and efficiency. This marks a significant contrast not only with the approach to reform adopted in the 1980s but also with the pattern of labour-market reform seen elsewhere in Europe, where the most common approach has been to relax the restrictions on the recourse to temporary, part-time or other non-standard forms of employment without altering the protections afforded to permanent workers.

This outcome owed much to both the severity of Spain’s employment problems and external factors, such as the pressures of European integration, but it also reflected the degree to which employers and trade unions alike accepted the need to reduce labour-market dualism, and the readiness of the latter to compromise on previously taboo issues like severance pay, albeit while respecting the acquired rights of incumbent regular employees, apart from the limited changes in dismissal procedures adopted in 1994. The array of issues on the table in discussions among the social partners and the government created the possibility for complex bargains involving concessions to different players on different issues (collective bargaining and labour contracts, conditions for regular contracts, as against terms for temporary contracts, etc). Moreover, the most important changes adopted in the 1997 reform were framed in such a way as to create no identifiable losers. The evolution of the Spanish labour-market reform process suggests that concertation need not be considered a sine qua non for adopting a reform, although it may make for a smoother process: unilateral reforms have met strong resistance that has sometimes limited their impact. Yet even where reforms were negotiated, government leadership was important, not only in the form of an implicit threat to act unilaterally if social bargaining failed, but also in the government’s readiness to use its fiscal and legislative powers to underpin compromises agreed by the social partners.

Notes

2. Most observers believed that this was above the true figure, and changes in the labour force survey introduced in 2002 brought some major revisions in labour market statistics and prompted the Bank of Spain to re-estimate the figures for earlier years. On the new series, unemployment peaked just below 20%. See OECD (2003) for details.
3. In an effort to limit the fiscal impact of rising unemployment, the government at this
time eliminated the tax deductibility of unemployment benefits.
4. Qualifying for the first wave of EMU was still an issue, but the focus of the
government’s accession efforts was on the budget rather than on structural measures
that would not bear fruit in time to affect the accession decision.
5. On this issue, see Chari (2001); and Molina Romo (2005).
6. Chari (2001) identifies several factors that contributed to the strength of the
renovadores, including the strong growth of 1986-91, which was attributed in part to
the policies they had pursued, and a gradual strengthening of their role in managing
the party’s finances.
7. Around 75% of Spanish workers were covered by collective agreements in the early
1990s. One reason for low union density is that workers have little incentive to join,
since any bargained agreement automatically applies to all workers in a sector.
8. See Bover et al. (2000). These arrangements were negotiated as part of Spain’s
democratic transition in the 1970s in an effort to strengthen a labour movement that
had been weakened by the previous authoritarian regime.
9. See Royo (2004) on the unions’ efforts to extend their support base beyond their
traditional constituencies.
11. UI benefits were not especially low, amounting to 70% of the last wage for the first
six months and 60% thereafter, with a maximum duration of 24 months. However,
these benefits were dependent on previous contributions and a worker required a
contribution history of at least 3 years to qualify for 12 months of benefits. Dolado
and Jimeno (1997) estimate that fewer than 30% of non-agricultural workers were
eligible for unemployment benefits in the early 1980s.
12. See Boeri et al. (2004); Boeri et al. (2007).
16. It should also be noted that Spain was in recession when the CTFE was introduced
but entering an upswing when the restrictions on its use were relaxed.
17. A consequence of both rising female participation rates and the arrival on the market
of the large cohorts born in the 1960s.
18. Jimeno and Toharia (1993); Bentolila and Dolado (1994); Dolado and Jimeno (1997).
22. See Toharia and Malo (2000) for references to the debates at this time. On the costs of
23. Toharia (1996). The contraction of the early 1990s triggered large-scale downsizing
in industrial sectors where permanent contracts predominated. Many firms shed older,
less productive and more expensive workers.

25. See OECD (1994) for full details of the UI reforms, which included adjustments in replacement rates, changes in the tax treatment of UI benefits, tougher eligibility requirements and stronger measures to combat UI fraud.


31. See Chari (2001) on the similarities between the September proposals and the business groups’ submissions to the government in the summer; and Toharia and Malo (2000) on the proposals’ resemblance to many of the ideas put forward by labour-law experts who worked with the employers’ organisations.

32. See OECD (1994) for details of these measures.

33. For a detailed description of the working-time and overtime changes, see Toharia and Malo (2000).


35. However, the remaining temporary contracts involved no severance payments at termination, as the old employment-promotion contracts had.


40. The CC.OO had called an earlier general strike against the government, to protest a 1985 pension reform, but the UGT did not support it.


42. Even the preamble to Law 11/1994 explicitly linked the need for greater labour-market flexibility to European integration.


44. Saint-Paul (1996).


48. Formally known as the “Report of the Committee for the Analysis of the Structural Problems of the Social Security System and of the Main Reforms that Must be Undertaken”, the Toledo Pact was a set of recommendations for the reform of the pension system presented by a parliamentary commission in March 1995. It formed the basis for a further agreement in October 1996 and subsequent legislation.

49. The official data at the time showed employment growing about as fast as GDP; these figures were later corrected downwards, but the adjusted data were still impressive,
showing employment growth of 1.7, 1.2 and 2.6% in 1995, 1996 and 1997, respectively.

50. See OECD (1998); Toharia (1999).
52. Toharia and Malo (2000); Bentolila et al. (2008).
55. Boeri et al. (2007).
56. In practice, this meant until May 1998. Apart from such conversions of temporary contracts, the contract was applicable only to new hires.
57. The agreement covered occupational structures, career structures, pay structures and disciplinary systems.
58. The workers who lost out under the reform were those employed on star contracts who would have managed to obtain permanent positions with secure, traditional contracts even in the absence of reform. Yet these individuals could not be identified either \textit{ex ante} or \textit{ex post}.
59. The contracts were intended to allow firms to innovate or enter new markets without taking irreversible risks, but many employers were accused of stretching the definition of “new activity” in order to justify using temporary contracts where they would not otherwise have been allowed.
61. On this issue, see especially Molina Romo (2005).
63. See Benito and Hernando (2003). OECD (2005) estimates the cost of the reform at roughly 0.4% of GDP in social security contributions forgone.
67. According to official social security records, some 300 000 people were likely to find themselves in this situation at the time of the reform; OECD (2007).
68. Rebates range between EUR 500 and EUR 3 200 per year, depending on the characteristics of the beneficiary.
71. See OECD (2008) for a recent discussion of the challenges.
73. See Ochel (2008).
**Bibliography**


Chapter 15

Sweden: Sickness insurance, 1990-2002

Throughout the 1990s and early 2000s, Swedish governments tried to tackle the problem of rapidly rising expenditure on sickness insurance (SI), as well as the concomitant growth in the number of days lost as a result of sickness. Successive governments introduced numerous adjustments to replacement rates, waiting periods, screening and monitoring, and the allocation of costs, but they made little sustained progress in tackling the problem, at least until about 2003. This case study examines the reforms of the 1990s and the reasons why they proved so difficult to implement and sustain despite broad agreement on the need to reform SI. The difficulties in tackling this issue reflected, \textit{inter alia}:

- \textbf{Haste caused by fiscal pressures.} The first major changes were adopted in response to the acute fiscal crisis of the early 1990s. Although the fiscal crunch did indeed make reform possible, the focus on making substantial short-term savings and the haste with which the first reforms were adopted meant that the measures involved were somewhat clumsy and created some perverse incentives. Consequently, pressure for reversal of the reforms rose sharply as soon as the fiscal squeeze abated.

- \textbf{Failure to communicate the labour-market goals of reform.} The emphasis on fiscal necessity meant that the government did little to communicate the labour market goals of reform. Yet making the case for change on structural grounds would have raised questions about the larger bargain underlying many Swedish reforms, whereby the maintenance of fairly generous social protection was the “price” governments paid for acceptance of trade openness and product-market liberalisation.

- \textbf{Lack of support for reform from the social partners.} Neither business nor labour strongly supported reform in the 1990s. The unions were determined to preserve high replacement rates. Employers incurred little additional cost when workers fell ill and could use SI as a means of adjusting employment levels. Cuts in public SI benefits were often restored via collective agreements between business and labour, which imposed a significant part of the cost of these agreements on the state.

- \textbf{The changing nature of the problem.} The position of Swedish employers was affected by changes in the profile of SI beneficiaries. At the beginning of the 1990s, the problem was largely one of endemic short-term absences, which firms found disruptive. This changed over the course of the decade (in part thanks to policy changes), so that by 2000, the big problem was the sharp rise in very long (over one year) sickness spells. Short-term absences were by then at fairly low levels. Long sickness spells were less of an issue for employers, since they were
less disruptive. The workers in question effectively ceased to be the enterprise’s concern and passed into the care of the state.

- **Implementation problems linked to perverse incentives.** There were repeated efforts to tighten access to SI benefits. However, those administering the system often had incentives *not* to be too rigorous. Examining physicians tended to accede to clients’ desires, for fear of being blamed for subsequent health problems if they were too strict, and administration of the scheme involved elected local politicians who risked unpopularity if they denied benefits and whose budgets were not directly affected by rising SI expenditure. Many screening requirements were thus implemented poorly or not at all.

More recently, successive governments have renewed their efforts to implement a sustainable SI reform. After the 2002 elections, the SAP government returned to the issue with a significant package of measures aimed at halving the number of people on sick leave by 2008, and its centre-right successor put forward additional measures in 2008. These reforms, which have so far contributed to a sharp decline in sickness absence, owe much to the experience of the preceding decade, which saw a gradual shift in the political consensus from a commitment to passive income assistance towards the kind of mutual obligations approach used in respect of the unemployed.

**The context for reform**

*Framework conditions for reform*

Sweden was governed by minority cabinets throughout the 1990s – a centre-right coalition, in power during 1991-94, was both preceded and followed by minority Social Democratic administrations. This meant that important changes required a degree of cross-party agreement. While reducing governments’ freedom of action, this meant that the responsibility for unpopular benefit cuts was shared with some elements of the opposition, and it reduced the risk of an immediate rollback in the event of a change of government.

The first attempt at reform of SI was initiated in response to, and in the midst of, the worst recession in over half a century. The sharp downturn of 1991-93 was followed by strong recovery in 1994-95, after which growth moderated but remained relatively steady into the early 2000s. The general government balance swung into the red in 1991, with the deficit reaching 12.3% of GDP in 1993, as rising unemployment and a contracting economy pushed public expenditure to over 60% of GDP by the middle of the decade. However, the deficit fell rapidly thereafter and was replaced by a surplus from 1998. A 2% surplus target for the general government was introduced in 1997, reflecting an awareness that population ageing would in future bring about a gradual deterioration in public finances. However, while the importance of meeting this target was recognised by successive cabinets, fiscal policy gradually came to focus more on short-term business-cycle management and reversing some of the austerity measures implemented during the budget consolidation phase in the 1990s.

The early-to-mid 1990s saw a number of social insurance reforms in Sweden despite the fact that the cabinets in office lacked parliamentary majorities. In addition to a series of revisions to the SI scheme, Sweden adopted a structural pension reform, as well as changes in unemployment insurance (UI), disability insurance (DI) and healthcare. In 1996, the replacement rates of all social insurance schemes were harmonised at 75%, well below the levels of the preceding decade. Alone among Nordic countries, Sweden
reduced per capita spending on social services and transfers in real terms (though not relative to GDP) during 1990-95. However, the latter years of the decade witnessed a partial reversal of this trend, as the fiscal pressures that had prompted it began to abate. It was, in any case, hard to push too far in reforming social insurances once the crisis of the early 1990s was over, owing to a strong consensus on the need for a relatively generous welfare state to cushion the impact of increasing external openness and deregulation on those adversely affected by such changes. SI reform thus represented a challenge to the implicit bargain that underlay many other Swedish reforms.

**Actors and interests**

All major parties accepted the need for reform of the welfare state after 1991, though they disagreed about the specific measures to be taken and varied in the extent of their commitment to welfare-state retrenchment. The Social Democratic Party (SAP) initiated the first major changes in early 1991, in response to the financial crisis that had erupted in late 1990. However, throughout the years that followed, the SAP leadership was under pressure from parliamentary backbenchers, trade unions and the party rank-and-file to reverse the cuts when circumstances permitted. Nevertheless, the party’s agenda when it returned to power in 1994 included low inflation and fiscal consolidation, as well as traditional employment and distributive goals, and this limited its ability to reverse the cuts. While some within the SAP recognised the case for reform on non-fiscal grounds, much of the party failed to embrace the structural reform agenda. By contrast, the centre-right parties that governed during 1991-94 viewed reform of SI and other programmes as desirable for structural as well as fiscal reasons, since they believed existing social insurance arrangements created perverse incentives, reduced labour supply and were fraught with moral hazard.

The main employers’ organisations in Sweden in the 1990s were the Swedish Employers’ Confederation (SAF) and the Federation of Swedish Industries (IF). Large firms constituted the backbone of these organisations and of the economy: in the late 1980s, over 60% of Swedish workers were employed in firms with over 500 employees, an extraordinarily high level in comparative terms. This was particularly relevant for SI reform, since changes in the SI regime affected large and small employers differently (see below). Unionisation rates in Sweden were exceptionally high, at around 80-82% of dependent employees throughout the 1990s. Union influence was assured by the unions’ dual role as labour representatives and administrators of the government-financed UI funds. They also had close institutional links to the SAP, although relations between the party and trade unions were strained on account of what the latter saw as the party’s embrace of overly liberal policies. The main trade union confederations included the Swedish Trade Union Confederation (LO), with almost 2 million members, and the Central Organisation of Salaried Employees (TCO) for white-collar workers, with around 1.3 million. While the two confederations generally spoke with a single voice in respect of SI-related issues, SI was more important to the LO, since blue-collar workers tended to have more frequent and longer sick spells than their white-collar counterparts.

On the issue of SI reform, trade unions and employers’ organisations often made common cause in the 1990s. The trade unions, not surprisingly, wished to defend high replacement rates from the first day of any sick spell, and employers did not object, so long as the burden on them was limited. Indeed, employers arguably had little incentive to care about reducing sickness absences: although they funded almost all of employees’ SI, they incurred little additional cost when workers fell ill, and employers often used SI as a means of adjusting employment levels in response to fluctuations in demand.
was more attractive to employees than redundancy, since the SI benefit ceiling was higher than the ceiling for UI. Since SI contributions were not hypothecated, employers did not directly bear the costs of using of SI benefits to downsize workforces, nor did initiatives that reduced sickness absence necessarily reduce their contributions (Box 15.1). Both employers and unions resisted government attempts to extend employers’ responsibility for paying sickness benefits. The employers always insisted that the payroll tax relief accompanying reform proposals was insufficient, and the unions feared that greater employer responsibility would reduce employers’ generosity in the context of collective agreements. The unions also feared that employers would try to get rid of employees with health problems and to avoid hiring those they thought at risk of suffering ill health. Both unions and employers’ organisations also favoured transferring control of the SI scheme to the social partners.

The administration of Sweden’s SI system is overseen by the Ministry of Social Affairs, but many of the changes adopted in the early 1990s were driven by the Ministry of Finance in its pursuit of fiscal consolidation. Although SI is organised as a form of social insurance, SI contributions are not hypothecated – they go into the general budget, and the Ministry of Finance has consistently resisted proposals to separate SI from the state budget and develop an independent insurance system.

**Issue history**

Sweden fared badly in international comparisons with respect to both sickness absence among the working population and sickness-motivated withdrawal from the labour force in the 1990s (Table 15.1). Only the Netherlands and Norway exhibited comparable levels of sickness absence. In addition, Sweden’s absence rates also tended to fluctuate more than any others. Such high levels of absence partly reflect the composition of the workforce: sickness absence rates tend to be higher for older workers and for women, and Sweden has historically had exceptionally high levels of employment among both these groups. Older workers are a small part of the workforce, however, and Sweden’s high sickness absence rates were chiefly the result of exceptionally high levels of absenteeism in the 20-49 age group. Econometric analyses found no evidence that demographic or health factors account for this phenomenon: data on health status (rather than benefit payouts) suggested that Sweden’s workforce was exceptionally healthy.

**Table 15.1. Sickness absence of employees in eight countries, by age group**  
Average (%), 1983-2001

<table>
<thead>
<tr>
<th>Country</th>
<th>20-29</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
<th>60-64</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>1.3</td>
<td>1.8</td>
<td>2.5</td>
<td>4.4</td>
<td>5.3</td>
<td>2.3</td>
</tr>
<tr>
<td>France</td>
<td>1.8</td>
<td>2.1</td>
<td>2.5</td>
<td>3.9</td>
<td>4.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Denmark</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
<td>2.3</td>
<td>2.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Germany</td>
<td>0.9</td>
<td>1.1</td>
<td>1.4</td>
<td>2.7</td>
<td>3.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Netherlands</td>
<td>3.2</td>
<td>3.9</td>
<td>4.5</td>
<td>5.9</td>
<td>4.9</td>
<td>4.1</td>
</tr>
<tr>
<td>Norway²</td>
<td>2.4</td>
<td>2.7</td>
<td>3.1</td>
<td>4.3</td>
<td>7.2</td>
<td>3.2</td>
</tr>
<tr>
<td>Sweden²</td>
<td>2.5</td>
<td>3.1</td>
<td>3.9</td>
<td>4.4</td>
<td>9.1</td>
<td>4.2</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1.8</td>
<td>2.1</td>
<td>2.5</td>
<td>3.9</td>
<td>4.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

To a great extent, Sweden’s anomalous sickness absence rates seem to have reflected the operation of the sickness insurance (SI) system. Sweden has long had one of the most generous SI systems in the world (Box 15.1). A large body of empirical work strongly suggests that changes in the generosity of the SI system have substantial effects on labour supply, particularly of women and older workers. Moreover, the short-term sickness rate has tended to be closely correlated with the economic cycle: absenteeism rises as the labour market tightens (Figure 15.1). This may reflect the disciplining effect of higher unemployment or it may be the product of selection when firms are shedding labour. Long-term sickness, by contrast, was less sensitive to the business cycle but began a sustained rise in the late 1990s, which significantly altered the terms of the SI reform debate.

Figure 15.1. Sickness cases and days lost, 1970-2002

1. The series has been adjusted (upwards) after 1992 so that it includes the waiting day and the days paid by employers (so it is comparable with the pre-1992 figures). The adjustment is based on Henrekson and Persson (2004) and NIER (2004).
2. Sum of waiting days, employer-paid days and public sickness insurance.


The welfare-state retrenchment of the early 1990s was preceded by a period of rapid expansion of social insurances: the generosity of family allowances and UI, as well as the SI system, was increased. The economy was growing fairly strongly, and the expansion of social insurances offered the SAP government of the day a way to retain the support organised labour, while implementing financial and product-market reforms that the unions opposed. With respect to SI, the government first fulfilled an election pledge to reduce the waiting period for SI benefits from two days to one. In 1987, it scrapped the remaining waiting day, something the unions had long demanded. This left many workers with effective replacement rates of 100% from the first day of a sickness spell and no need for a medical certificate for the first week. These changes contributed to a rapid rise in the incidence of short sick spells. By 1988, the average sick leave per insured employee exceeded 25 days per year, rising to 40 in some sectors. This trend worried both employers and the government, but the SAP would have paid a high political price for reversing course on SI at that point.
Box 15.1. Sickness insurance in Sweden

Sweden’s sickness insurance (SI) scheme is a state-administered compulsory programme financed by payroll taxes levied on employers. At the beginning of the 1990s, it was administered by 21 regional social insurance offices and characterised by: very high replacement rates, the absence of any reliance on employers to fund sick pay directly, and the lack of any limit on the duration of benefits – a feature of the system that was almost unique in the OECD. SI benefits could be paid for an unlimited period, so long as the person was unable to work because of sickness and had not been transferred to a rehabilitation benefit or a disability pension. While SI benefits were capped in the public scheme at around 115% of the average production worker’s wage, benefits in most cases were “topped up” by additional insurances provided under collective agreements between unions and employers. This meant that effective replacement rates were often 100% for extended periods.

Eligibility requirements were comparatively lax: the right to a benefit had little to do with one’s employment or contribution history, and a medical certificate was required only after the first week. This was usually provided by the individual’s own GP, so only after two weeks did the social insurance office become involved in assessing entitlement to further benefits. Social insurance physicians generally based their opinions on GPs’ assessments and did not usually re-examine patients. Individual patients therefore had a great deal of influence over sick-listing decisions.

The employer was required to produce a rehabilitation plan after 1-2 months, sometimes in conjunction with the insurance office. If it seemed likely that the reduced work capacity would last longer than a year, the worker would be granted a temporary or permanent disability benefit. In principle, there was a re-evaluation after one year, but these were often performed late, and the doctors at the social insurance office did not, in any case, perform independent medical assessments. After the first year, a re-assessment was required every two years, but implementation of this requirement tended to be patchy.

Since SI premiums were not hypothecated, SI claims in excess of the premiums paid were financed by the budget, which also captured any savings in the event that contributions exceeded pay-outs. This arrangement arguably gave rise to moral hazard, since the social partners did not directly bear the additional costs to the public SI system that arose as a result of collective agreements concerning SI “top-up” benefits.

1. Since 2005, it has been administered by a single central body, the Social Insurance Agency.
2. Ireland has no limit on the duration of benefits, but its scheme is far less generous in other respects.
3. Since 2003, the disability pension has been renamed “sickness compensation” and integrated with the SI system – previously, it was part of the pension system – but this lies outside the scope of the case study.
4. Esping-Andersen (1990) compares SI systems in 18 OECD economies with respect to the link between one’s right to SI benefit and one’s work/contribution record. He finds Sweden and Denmark have the weakest link between benefit entitlement and recent experience in the labour market.
5. Two studies of the issue found that patients were rarely denied a sick listing if they requested it and that GPs rarely offered one unless patients asked for it. See OECD (2003b, 2005) for details.

Policy process: formulation

The search for solutions

The surge in short sickness spells after 1987 prompted considerable debate, but there was strong resistance in many quarters – not least among the unions – to acknowledging the moral hazard involved. Yet the evidence leaves little doubt that the social norms that had long limited abuse of the system were being eroded over time, a process that was probably accelerated by the macroeconomic shocks of the early 1990s. Defenders of the
SI system insisted that rising rates of sickness absence reflected a need to develop healthier working conditions. Some officials involved in administering SI argued that the increase in short sickness spells was actually desirable: workers’ greater propensity to take off when sick would reduce the likelihood of longer illnesses and of their infecting their colleagues. Local social insurance offices in some areas even experimented with scrapping the requirement for a medical certificate after the first week.31

Those concerned with the perverse incentives generated by the SI system focused their attention on two issues: screening and the structure of benefits. Tighter screening was seen as essential to addressing the moral hazard, given that there was a fairly strong consensus in favour of high replacement rates and fairly generous terms for the availability of SI benefits. When it came to restructuring SI benefits themselves, a number of trade-offs were identified:

- The overall level of replacement rates needed to balance workers’ welfare against the awareness that individuals’ work incentives would be reduced if they bore none of the financial costs of ill health.
- Benefits over the duration of a sick spell needed to be structured so as to avoid incentives to take longer or more frequent sick spells. Offering lower benefits (or none at all) at the start of a sickness spell might tend to reduce the frequency of such spells but increase their duration. Failure to take short spells might well lead to more long spells, as untreated minor complaints grew worse, and a worker unsure of whether to return to work would remain on sick leave rather than risk returning to work and falling ill again, thereby incurring additional losses. A decline in replacement rates over time, in contrast, might encourage workers to return to work but at the risk of penalising the long-term sick.
- If employers paid little or nothing directly, they would have less incentive to care about working conditions or employee health, but if they paid too much, they would face incentives to avoid employing workers who seemed “at risk” in health terms. SMEs, in particular, might suffer in the latter situation, since the burden of supporting a sick worker would weigh all the more heavily on them. Employers as a group were also concerned that the introduction of any employer-covered period be offset by a reduction in payroll taxes sufficient to keep the overall cost of the scheme to employers from rising.

In addition to these considerations, there was growing awareness over the course of the 1990s of the interaction between various social insurance schemes: because eligibility, replacement rates and maximum benefits varied, changes in one often had unanticipated consequences for others: tighter access to one would lead to the absorption of workers by other, more accessible insurances. In 1993, a government-appointed commission on Sweden’s economic future (the so-called “Lindbeck Commission”) recommended harmonising the replacement rates of social insurance schemes, but this proved very difficult and did not in any case eliminate unexpected interactions between them, because eligibility criteria, benefit ceilings and other system parameters could not be harmonised entirely.

While SI policy was much debated during the 1990s, and a great deal of empirical work was done on the subject, it was never the focus of the kind of comprehensive, cross-party reform effort that led to the overhaul of the general pension system. The issue was examined by a number of government commissions, but these did not lead to much, owing largely to the social partners’ failure to embrace reform. The big cuts of 1991-93
were adopted rapidly, in the midst of a severe recession, and subsequent adjustments to replacement rates and changes in employer responsibility were taken in piecemeal fashion, often in the context of the budget process, rather than as part of a thoroughgoing reform. The result was a multitude of adjustments to the system (Table 15.2), some of which had unanticipated consequences.

Table 15.2. Replacement-rate history, 1987-2003

<table>
<thead>
<tr>
<th>Sick leave compensation as a percentage of own income (and the share paid by employers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>First day</td>
</tr>
<tr>
<td>Days 2-3</td>
</tr>
<tr>
<td>Days 4-14</td>
</tr>
<tr>
<td>Days 15-28</td>
</tr>
<tr>
<td>Until 90 days</td>
</tr>
<tr>
<td>Until 1 year</td>
</tr>
<tr>
<td>Over 1 year</td>
</tr>
<tr>
<td>First day</td>
</tr>
<tr>
<td>Days 2-3</td>
</tr>
<tr>
<td>Days 4-14</td>
</tr>
<tr>
<td>Days 15-28</td>
</tr>
<tr>
<td>Until 90 days</td>
</tr>
<tr>
<td>Until 1 year</td>
</tr>
<tr>
<td>Over 1 year</td>
</tr>
</tbody>
</table>

Note: Figures in bold represent changes relative to the previous scheme. Figures in brackets show the share paid by employers, including mandatory employer-financed sick pay for the early part of a sickness spell and top-up payments under collective agreements. Throughout the period shown in the table, coverage of the public system was capped, but fewer than 10% of wage earners were consistently above this cap, and collective agreements usually covered income above the cap.

* The waiting day can be partly avoided after 1993 by reporting sick part way through the day.


Policy process: adoption and implementation

Response to the crisis in 1991

The first major attempt to curb the generosity of SI benefits was made in early 1991, as part of the government’s response to the severe foreign exchange crisis that erupted in the autumn of 1990. Anxious to bring public expenditure under control, the SAP government had little choice but to look for savings from social insurance schemes, since they constituted such a large part of total government expenditure – SI benefits alone amounted to 2.5% of GDP in 1990. The SI replacement rate was cut in March 1991 from 90% of earnings throughout the period of sickness to 65% for the first three days...
(excluding top-up benefits), 80% for days 4-90 and 90% thereafter. Financial pressure ensured that the changes were adopted without the usual long consultative process that is typical of Swedish policymaking. There was thus little attempt to explain the reform in terms of labour supply and growth objectives; on the contrary, the SAP made clear that the tax increases and cuts in welfare benefits being adopted as part of the austerity drive were meant to be temporary.

The radical reduction in the replacement rate for the first days of a sickness spell was necessitated largely by two factors. First, the SAP was determined not to violate its 1988 election pledge by re-introducing a waiting day. With elections due later in the year and the government’s relations with the LO and other major unions already under strain on account of its austerity policies, this would have been very costly in political terms. Secondly, the government was anxious to retain a 90% replacement rate for an indefinite period, in order to avoid a regressive impact on income distribution – low-paid workers were more likely than the better off to suffer long sickness spells. In order to ensure that workers on long-term sick leave nevertheless had an incentive to return to work, it was decided that no additional compensation from collective insurances would be allowed past the 90th day of illness. The rejection of both a waiting day and a fall in the long-term replacement rates of the public SI scheme meant that significant savings could be realised only by cutting benefits sharply in the early stages of illness.

These changes had a dramatic effect: average sick leave per worker fell from just under 25 days in 1990 to around 15 in 1992, making for an increase in hours worked of about 5%. One subsequent analysis concluded that, given typical estimates of the elasticity of labour supply in Sweden at the time, this was equivalent to the labour-supply effect of a 40-50% increase in after-tax wages. The effect on short-term absences was particularly pronounced, but sickness spells grew longer on average. Since the benefit rate was cut sharply for the first three days of a sickness spell, those returning to work risked further losses if they fell ill again soon and had to begin a new spell.

Changes under the centre-right

The centre-right coalition that took office in October 1991 moved rapidly to follow up on the previous government’s changes. As from January 1992, employers were required to provide for sickness payments during the first 14 days of a sickness spell. Effective replacement rates for the first two weeks were left unchanged, but the government planned to reintroduce a two-day waiting period from January 1993. These changes were intended to give employers stronger incentives to monitor sick leave, while simultaneously discouraging unnecessary sickness absence by imposing earnings losses on workers at the start of new spells. At the same time, the government adopted a series of measures designed to shift responsibility for monitoring and rehabilitation/reintegration onto employers, arguing that it was in the workplace that sickness absence, and thus the need for rehabilitation, were most likely to be observed at an early stage. A greater direct role for employers was thought to be a natural concomitant of increased financial responsibility. Both unions and employers disliked these innovations – the employers’ organisations took the view that the payroll tax relief provided in conjunction with this new obligation was insufficient to cover the costs involved, and the unions remained strongly opposed to the re-introduction of the waiting period. Many employers also disliked their enhanced monitoring responsibilities, since they had limited resources and little capacity to sanction employees who did not actively cooperate with efforts to rehabilitate and reintegrate them. However, the social partners’ objections carried only limited weight: the fiscal situation remained acute, and the unions
in any case carried far less weight with a centre-right government than they had with the SAP.

Concessions were nevertheless forthcoming. In the event, only one waiting day was instituted, and even this could be partly avoided by reporting sick during the course of the day. In addition, the government set aside an SEK 10 billion “Working Life Fund” to finance measures intended to improve health conditions in the workplace over a three-year period. This was meant to address the concerns of those who argued that working conditions were responsible for high levels of sickness absence and also to make it easier for employers to finance changes that would reduce their sick-pay bills. Overall, private employers proved keener to access and use these funds than did public-sector entities, perhaps reflecting the greater incentives the former had to reduce sickness-related spending. With sickness absence rates and SI expenditure continuing to fall sharply – a reflection of labour-market weakness as well as the 1991 reforms – the government could afford such concessions.

A final innovation introduced by the centre-right coalition proved far more unpopular than the changes relating to the first fortnight of illness: replacement rates for long sickness spells, which the SAP had left unchanged, were reduced. In January 1992, the replacement rates for periods from 91 days on dropped from 90 to 80% and from July 1993, the replacement rate for periods of more than one year was cut further, to 70%. Since the government did not alter the previous cabinet’s decision to limit top-up benefits to the first 90 days of a sickness spell, this meant that those on sick leave for over one year were palpably less well off. A gradual decline in SI replacement rates was intended – as in many other social insurance systems – to increase the incentives for the long-term sick to re activate if possible. It might have offset to some extent the impact of the waiting day in encouraging longer spells – having climbed steadily for over a decade, the number of sickness spells of more than a year began to decline in 1993. Nevertheless, the 70% replacement rate was widely criticised by the unions and the parties of the left for imposing additional hardship on those with the most serious health problems.

Growing generosity and conflict with employers

Although the SAP had opposed the reduction in long-term replacement rates, it did not move to reverse them when it returned to power at the end of 1994. Despite pressure from the trade unions and its own left wing, the new government remained committed to fiscal consolidation and, in 1996, it moved to harmonise replacement rates for a number of social insurance schemes at 75%. This meant a reduction in replacement rates in the SI scheme for most of the first year but a rise for the long-term sick, a trade-off the new government argued would be more equitable than the arrangements it had inherited. However, it became harder to resist pressure for at least a partial reversal of the retrenchment of the early 1990s as both the fiscal position and the situation with respect to sickness absenteeism continued to improve. By 1995, SI expenditure was below 1.5% of GDP, down from 2.5% four years earlier. In 1996, the government yielded to this pressure, raising the replacement rates for both UI and SI to 80% from January of the following year. The 1997 SAP congress went further, committing the party to more generous SI benefits, as well as increases in other transfers. While this hardly marked a return to the policies of the 1980s, it was a sharp break with the austerity of the immediate past. The 1996 Employment Bill aimed to mitigate the effects of higher replacement rates by tightening eligibility for SI and introducing a three-year limit on benefits, but it was repealed before it entered into force. The legislation that replaced it scrapped the three-year limit and retained the higher replacement rates. It did increase the
qualifying period for SI eligibility, but only from five months’ contributions to six, rather than from five to nine, as originally proposed.39

Far more controversial than the rise in replacement rates was the government’s attempt to shift much more of the cost of short spells onto employers. From January 1997, employers were given responsibility for the first four weeks of sick pay. This was still rather short by international standards, but it provoked serious opposition from employers.40 It also had wholly unintended consequences for labour relations. Under collective agreements, employers had long paid 10% of wages to workers from the 15th day of absence. When the Sick Pay Act obliged employers to pay 75% during the third and fourth weeks of absence, they included this 10% payment in the 75% total. Unions challenged this before the Arbitration Board and the Labour Court. Both found for the employers. The government had believed that the employers would pay 75% plus whatever was due under collective agreements, as had happened in the past when employer responsibilities were extended. However, the wording of the new legislation was found to correspond more closely to the employers’ interpretation, resulting in a fall in effective replacement rates that the government had not expected.

The government’s response was to propose adjusting the legislation so as to achieve the outcome it had envisaged all along. The SAF rejected this claim, insisting that employers could not afford to pay 75% of wages for 28 days. The trade unions backed the SAF. The unions had never favoured long employer-financed sick-pay periods; they feared discrimination in hiring and argued that one’s access to rehabilitation and retraining should not depend on one’s employer. The unions feared that imposing such a burden on employers would complicate collective bargaining on other issues. Under pressure from both unions and employers’ organisations, the government dropped its proposed amendment and reversed both the payroll tax cut and extension of employers’ payments. This effectively re-established the pre-1997 position, apart from the fact that replacement rates had been raised to 80% in the meantime.

The SI scheme became even more generous in January 1998, when the government changed policy to allow additional compensation under collective agreements after the 90th day of the sickness absence, something that had not been permitted since the beginning of 1991. The government took the view that the previous policy had discriminated against collectively agreed insurances as opposed to private ones. Following the change, both blue-collar workers in the private sector and white-collar municipal workers negotiated top-up benefits equal to 10% of pre-sickness wages for days 91-360.41 One recent assessment suggests that this change increased the average length of sickness spells of over 90 days by 4.7 days.42

While adjusting replacement rates and the duration of benefits was relatively straightforward, and implementation of such measures was more or less automatic, it was also extremely unpopular. There was a strong consensus that replacement rates should be generous and that the moral hazard problem could be addressed by tougher screening, which would limit access to benefits to those in genuine need, and by more effective rehabilitation and reactivation measures. In 1995, the government adopted a much tougher set of rules governing doctors’ certificates. However, subsequent reviews found that implementation was very patchy.

This was partly a reflection of time pressures on physicians and officials, but other factors were also at work. First, many doctors found the rules too strict and inflexible, and therefore largely ignored them. Many, if not most, GPs were in any case reluctant to deny sick notes to their patients, partly because they feared being held responsible if an ailment
turned out to be unexpectedly serious and partly because they had every incentive to please patients who might have been with them for years. Secondly, elected local politicians were directly involved in administering the system via local social insurance offices, and they were reluctant to court unpopularity by denying benefits. The somewhat perverse incentives involved here reflected the fact that the SI system was centrally financed but locally administered, and local administrations had few incentives to worry about the financial implications of their decisions.43

A further government review also found that the 1992 reforms increasing employer responsibility for monitoring and rehabilitation were not being implemented. Employers in general were unaware of their responsibilities and generally did not investigate employees’ rehabilitation needs. The social insurance offices, the committee found, were failing to remind employers of their duties.44 In short, efforts to use better screening and activation measures were found to achieve little.

Reversal of progress

One reason the gradual relaxation of earlier SI reforms proved so hard to resist in the late 1990s was that reform had been presented primarily as a temporary fiscal measure. Although many economists argued the case for reform in terms of labour-market performance and work incentives, little effort was made to sell reform to the electorate in these terms, and the unions continued to resist this line of argument. Moreover, the sickness absence problem appeared for a time to have lost much of its former urgency. When absentee rates began to pick up in the latter part of the decade, it was initially viewed as no more than a response to the tightening labour market and perhaps some partial reversion to trend. Concern really began to grow only in the early 2000s, when it became apparent that absenteeism was rising much more than the authorities had expected. Moreover, the increase was driven chiefly by a marked rise in the number of long-term sick. By 2002, around 6.8% of the labour force was receiving sick pay at any given time, not counting those on rehabilitation pay or work injury insurance pay, and over half of those absent through sickness had been off work for more than a year. The number of people absent for 30 days or more owing to injury or illness had doubled in five years, while SI benefit payments increased by 170% in real terms.45

These developments prompted much debate. Initially, the government argued that working conditions were primarily responsible for the increase in long-term sick spells. However, this view was difficult to square with indicators of employee health status, and a large part of the growth was driven by stress, depression and other difficult-to-diagnose ailments.46 A large body of evidence pointed to the importance of policy changes in and around the SI system. Both the rise in replacement rates and the reduction in employer responsibility measurably increased the propensity to take sick leave, a trend that seems to have been reinforced by more generous agreement-based security in cases of long sickness spells.47 Interactions with other social insurances also seem to have played a role:

- The new pension system reduced the payout rate actuarially if retirement was taken early, but this did not affect those leaving the labour force on SI or disability pensions; the incentive to take this path to exit the labour force was strengthened further by the decision to add old-age pension contributions to all benefits from 1999 onwards.48
- SI was preferable to UI for many workers. The ceiling for the former was 115% of the average production worker’s wage, as against 90% for the latter,49 and
agreement-based top-ups ensured that many workers had better replacement rates under SI than UI. Unlike SI, UI was still subject to a waiting period (five days) and was time limited. SI claims thus exhibited a tendency to increase as the expiration of unemployment benefit approached.

- Until 1997, workers over 60 could be granted disability insurance (DI) benefits for labour-market reasons in conjunction with health factors. The medical requirements and requirements to change residence or occupation were not very stringent, and there was no requirement to join rehabilitation or retraining programmes. From 1997, this path to DI was ended, resulting in a significant diversion of older workers into UI or SI, particularly the latter. Transitions from employment into SI were rising for all age groups, but the inflow of older workers accelerated markedly at the time of the DI reform.

By the beginning of the 2000s, SI reform was back on the government’s agenda, but the social partners continued to resist fundamental change. A government committee in August 2000 proposed extending the period of employer responsibility from 14 to 60 days and requiring co-payments from employers throughout the sickness episode in an effort to strengthen their incentives to improve employee health. In return, the commission proposed a reduction in employer contributions to the SI system and limiting the period covered by SI to one year. These proposals met strong opposition from employers and organised labour alike and most were blocked. Both unions and employers retained their traditional objections to extending employer-financed sick pay, and the unions were strongly against the one-year limit. In addition, critics noted that the growth in long-term sickness cases was concentrated in the municipalities and some other parts of the public sector, where increased employer responsibility for sick pay would not have the same incentive effects as in private industry.

In any case, the politics of SI had changed over the decade. First, public sympathy appeared to be far greater for the long-term sick; there was widespread acceptance of the fact that a lot of short absences in the past had been unwarranted, but there was little support for cutting benefits to those with long-term illnesses. Secondly, the long-term sick posed different problems for employers: while frequent short-term absences were disruptive, longer absences tended to be less so, especially for large firms, as those on long-term sickness absences largely became the concern of social insurance offices, almost as though they had been made redundant. After the first two weeks of a sickness spell, the cost to the employer fell sharply, from 106% of gross pay to no more than about 13.3% – lower in the case of employees whose top-up benefits did not extend over the long term.

**Conclusion**

There are a number of factors that explain why reform progress after 1991 was not steadier or more sustainable. First, although many specialists were aware of the perverse incentives created by the SI system, the early reforms were justified chiefly in terms of their fiscal benefits. The labour-market rationale for reform was never accepted across the political spectrum. Thus, as the fiscal situation improved, it was hard to resist pressures to restore some past cuts. Yet making the case for change on structural grounds would have raised questions about the larger bargain underlying many Swedish reforms, whereby the maintenance of fairly generous social protection was the “price” governments paid for acceptance of trade openness and product-market liberalisation. Secondly, there were no
obvious winners from SI reform, because the benefits of increased labour-supply were poorly understood by the public. Workers activated as a result of changes in SI might be financially better off in the end, but they did not necessarily welcome increased pressure to return to work; the unions were understandably committed to generous benefits and relative easy eligibility; and employers found generous social insurances a useful way to increase their own internal flexibility. Moreover, the generosity of the system, in conjunction with the system of top-up benefits and the fact that the system was not separated from the state budget, enabled both social partners to shift onto the state a part of the cost of their own collective agreements. Thirdly, as the nature of the problem changed towards long-term sickness spells, addressing the moral hazard problem without penalising those who were seriously ill grew more difficult – doctors and officials proved particularly reluctant to apply the kind of tough screening that was meant to accompany the commitment to very high replacement rates. Finally, the failure to co-ordinate the reform of various social insurances meant that changes to other programmes had unintended consequences for SI.

Nevertheless, the sickness absence problem remained too big to ignore, and the renewed reform debates of the early 2000s did not prove to be inconclusive in the end. After the 2002 elections, the SAP government returned to the issue with a significant package of measures aimed at halving the number of people on sick leave by 2008. The evidence to date suggests that these reforms have had a remarkable effect: the incidence of sickness absence fell by roughly 40% between 2002 and 2007, despite an improving labour market. This seems to have resulted largely from changes in the SI system. Moreover, the new centre-right coalition put forward additional measures of its own in 2008. The recent reforms owe much to the experiences of the preceding decade: despite the zigzag nature of SI policy in the 1990s, there was a gradual shift in the political consensus from a commitment to passive income assistance towards the kind of mutual obligations approach used in respect of the unemployed. Moreover, recent policy making has been informed by the rich vein of research done on the evolution of SI in the 1990s.

Notes

2. See OECD (2005) for details.
4. The pension reform adopted in 1998 replaced the existing defined-benefit system with a defined-contribution scheme. In the process, Sweden became one of the pioneers of notional defined contribution systems.
5. Eligibility requirements were tightened, replacement rates cut and a waiting period reinstated.
7. Charges for some health services were introduced.
9. Kuhnle (2000). Even then, however, its expenditure on these items’ share was highest, at 35.8% of GDP.
12. Specifically, the Moderate Party, the Liberal People’s Party, the Centre Party and the Christian Democrats.
13. They merged in 2001 to form the Confederation of Swedish Enterprise (SN).
14. The European Community average at the time was 30.4%; in most European economies, almost half of all employment was in firms with under 100 employees. See Olsen (1996).
17. Bergendorff (2003). Norway and the Netherlands were the other two economies with fairly high volatility of sickness absences.
18. Only the United Kingdom, with an employment rate of 47% for males aged 60-64, came close to Sweden (49%) and Norway (58%).
19. See Bergendorff (2003); Rae (2005).
21. On the OECD’s composite index of compensation generosity, Sweden in 2003 scored about 130% of the OECD average, level with Norway and well above any other member. While the index is not available for earlier years, Sweden has long stood out for the generosity of its SI scheme; see OECD (1994:95-96).
24. Employees with poorer health are more likely to be without employment in bad times, and older workers are more likely to be let go in a recession, especially if a path to early retirement exists for the latter. In a comparison of income-security policies in Sweden and Finland, Hytti (2006) finds that the recession and recovery of the 1990s had little influence on the uptake of sickness and disability benefits in Finland. In Sweden, short-term sickness spells fell sharply during the downturn but the incidence of SI payments in excess of 12 months (and disability pensions) rose markedly.
Child allowances were increased for a third child and above, as well as for successive births that were close together in time. Eligibility requirements for UI were eased, the waiting period was abolished and replacement rates were raised.

Including the liberalisation of the financial sector and the opening up of some network sectors to competition. See Olsen (1996) and Kuhnle (2000) on tensions between the SAP and the LO.

The 1987 legislation also changed the formula for calculating benefits in such a way as to help workers with irregular hours, who had been disadvantaged under the old arrangements.

For subsequent econometric estimates of the impact of the 1987 changes, see Henrekson and Persson (2004), who estimate that the abolition of the waiting day led to a 10% increase in the average number of sick days; and Andrén (2004), who observes that the number of 1- and 2-day absences, in particular, jumped after the waiting day was abolished. Absence rates for men rose significantly on or around birthdays or major sporting events, a development that highlights the moral hazard problem; see Skogman Thoursie (2004, 2005).


Skogman Thoursie (2004, 2005); Lindbeck et al. (2008). On endogenous habits and norms, see Lindbeck (1995), who demonstrates how a macroeconomic shock may accelerate the erosion of norms that tend to constrain the influence of welfare-state disincentives on individual behaviour.

The argument underlying such experiments was that physicians were likely to be overly cautious and to sign workers off for longer than the workers would take if left to themselves. In fact, sickness absences were even longer for those not subject to medical screening. See Hesselius et al. (2005).


i.e. after taking account of top-up benefits.


Depending on the time of day, one might lose as little as ¼ of a day’s pay at the start of a sickness absence.

This reduction was effective, inasmuch as the policies in place prevented the social partners from making up the losses under collective agreements.

See OECD (1999).

Germany, for example, required up to 6 weeks, Austria up to 12 weeks, and Switzerland up to 6 months. In the Netherlands, employers bore full costs for 28 weeks, and in the United Kingdom one year. However, the experience of the Netherlands and the United Kingdom with very long periods suggests that re-insurance can blunt the incentives on firms to invest in prevention.

No such top-ups were negotiated for central government employees or white-collar employees in the private sector.
43. This situation was corrected when the Social Insurance Agency was created in 2005.
44. EIRO (1998).
46. See, especially, Rae (2005). There was, however, a modest demographic effect; OECD (2002) estimates that the ageing of the workforce could increase the number of SI beneficiaries by as much as 12% by 2008.
48. See Palme and Svensson (2003) and OECD (2005:82-83) for details.
49. See Hytti (2006), who shows that the recession of the 1990s led to a much sharper rise in unemployment in Finland, where unemployment benefits were relatively more generous, than in Sweden, where the rise in unemployment was smaller but the increase in SI benefits much greater.
52. Until 1991, labour-market conditions alone were sufficient, even in the absence of any evidence of ill health or disability.
54. The 60-day proposal was reduced to 28 days and eventually adopted in the 2003/04 budget at 21 days, but the time limit on the duration of SI was dropped. From January 2005, employers were required to make a contribution equal to 12% of wages beginning at the fourth week and continuing throughout the duration of the sickness spell.
55. An 80% replacement rate plus 0.33 x 80% in social security contributions for the sick employee. See Hytti (2006).
56. 10% in top-up benefits plus 0.33 x 10% in social security contributions.
57. The package included strengthening administration of the scheme, reducing replacement rates and adjusting benefits for unemployed persons reporting sick, an additional week of employer liability, subsidies for those employing people on long-term sick leave. See OECD (2003a, 2005) for details.
Bibliography


Chapter 16

United States: The 1996 welfare reform

The 1996 Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) brought about an overhaul of America’s major social assistance programmes, in an effort to reduce work disincentives for benefit recipients and to improve access to training, childcare and health insurance for those making the transition to work. The existing federal entitlement to assistance was ended and the federal government retained responsibility only for setting broad principles concerning welfare provision. Responsibility for applying these principles and implementing programmes passed to the states. The principles enshrined in the bill included the following:

- no one may receive cash welfare benefits for more than a total of five years during his/her lifetime (although in exceptional cases money may be provided from non-welfare sources);
- no one may receive welfare benefits for more than two years continuously (although other benefits, such as subsidised employment, are allowed after the two years have elapsed).

PRWORA left the states considerable discretion in granting exceptions to those disqualified on the basis of the two- or five-year rules, but it also contained penalties for states that failed to move enough welfare recipients into “work activities”. The legislation increased funding for childcare by USD 4 billion over seven years. While there was enormous controversy over various specific provisions, the overall approach taken by PRWORA reflected an emerging consensus on the need for a welfare reform that strengthened work incentives while providing more assistance to those making the transition back into work. The existence of this consensus was an important factor making reform possible.

Welfare reform was a very difficult issue to tackle, especially in an election year, and the legislation passed only when the Republican Congress and the Democratic President reached agreement in the wake of two previous presidential vetoes. Among the factors facilitating its passage were the following:

- Widespread dissatisfaction with the status quo. Surging welfare rolls, growing evidence of the emergence of a significant class of long-term welfare dependents and a broad awareness that America’s basic welfare programmes were ill designed either to fight poverty or foster labour-market re-entry meant that there was broad agreement on the need for reform. Virtually no one wished to defend the status quo.
- Broad consensus on some elements of the reform. While some elements of the reform provoked fierce controversy, there was broad agreement on certain basic principles. It was generally accepted that an adult judged able to work should not
be able to receive assistance indefinitely without working and that helping recipients back into the workforce would require greater emphasis on education, training, childcare and other forms of assistance to facilitate the transition from welfare to work. The hard bargaining concerned the specifics of how to put these principles into practice.

- **Electoral calculations.** Welfare reform was broadly popular with the electorate, and President Bill Clinton wanted to make good on his 1992 campaign pledge to “end welfare as we know it” before the 1996 general election. Some senior Republicans argued against compromising on welfare reform in order to deny him victory on such a key issue and to use it against him in the election campaign. However, the Republican congressional majority elected in 1994 was also under pressure to deliver on its own campaign promises, which included welfare reform. Both sides thus aimed to present the reform to voters as their own achievement.

- **State governments’ activism.** State governments were already experimenting with welfare reform and were eager to support a federal-level reform that would offer them both greater flexibility in running welfare programmes and substantial “windfall” federal funds. Governors of both parties thus tended to favour reform, and they played a key role in finding the compromises needed to bring Congress and the White House together. Their involvement made it easier for both the administration and congressional Democrats to accept a bill that largely reflected the views of the Republican majority.

- **Compromise with interests capable of stalling reform.** While the Food Stamps programme was cut, the attempt to end it as a federal entitlement and pass it to the states as a block grant was thwarted by the agrarian lobby, while proposals to roll Medicaid into the reform foundered on opposition from state governments, the healthcare industry and the “grey lobby” – although Medicaid was created to provide healthcare to those on low incomes, much Medicaid spending is for long-term care for the elderly.

### The context for reform

#### Framework conditions for reform

Few observers in early 1996 expected a major welfare reform that year. Even an administration whose party controls Congress can find it very difficult to win adoption of its policies – as, indeed, had been the case in 1993-94, when the Democrats controlled both Congress and the White House. By 1996, the situation was still more complicated, because the Republicans had taken control of both houses of Congress the year before and had begun to push an ambitious and wide-ranging legislative agenda that included welfare reform. While there was some overlap between the Democratic administration’s approach to the issue and that of congressional Republicans, the latter took a tougher line on a number of issues, and the approach of the 1996 elections gave both parties incentives to highlight partisan differences and avoid compromises that might upset their supporters. Many in both parties wished to keep the issue unresolved in order to use it in the election campaign.

PRWORA was adopted and implemented against the backdrop of an extended period of growth. Both unemployment and poverty were already declining (Table 16.1), as were
welfare rolls: during 1995-96, the number of welfare recipients fell by 11.1% from its 1994 peak. This decline reflected falling unemployment, state-level experiments with welfare reform, conducted under federal waivers, the impact of the Earned Income Tax Credit (EITC) and rising public expenditure on childcare.\(^2\) Overall, income growth for the lower quintiles of families with children during 1992-96 was relatively robust, with the incomes of single mothers in the lowest three quintiles rising by around 20%, as their employment rates rose.\(^3\) However, these favourable trends – which in due course made welfare reform much easier to implement – did little to reduce the sense of urgency about the issue. The reform debate was heavily influenced by a surge in the welfare rolls in 1989-94. By early 1996, it was clear that things were improving but not that the improvement marked a change in the trend.

Table 16.1. Economic and social indicators, 1989-96

<table>
<thead>
<tr>
<th>Year</th>
<th>Real GDP growth (%)</th>
<th>Unemployment rate</th>
<th>AFDC recipients (million persons)</th>
<th>Fiscal balance (% of GDP)</th>
<th>General government net debt (% of GDP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>3.4</td>
<td>5.3</td>
<td>10.9</td>
<td>-2.8</td>
<td>44.1</td>
</tr>
<tr>
<td>1990</td>
<td>1.2</td>
<td>5.6</td>
<td>11.5</td>
<td>-3.9</td>
<td>45.2</td>
</tr>
<tr>
<td>1991</td>
<td>-0.9</td>
<td>6.8</td>
<td>12.6</td>
<td>-4.5</td>
<td>48.9</td>
</tr>
<tr>
<td>1992</td>
<td>2.7</td>
<td>7.5</td>
<td>13.6</td>
<td>-4.7</td>
<td>52.4</td>
</tr>
<tr>
<td>1993</td>
<td>2.3</td>
<td>6.9</td>
<td>14.1</td>
<td>-3.9</td>
<td>54.9</td>
</tr>
<tr>
<td>1994</td>
<td>3.5</td>
<td>6.1</td>
<td>14.1</td>
<td>-2.9</td>
<td>54.5</td>
</tr>
<tr>
<td>1995</td>
<td>2.0</td>
<td>5.6</td>
<td>14.5</td>
<td>-2.2</td>
<td>53.7</td>
</tr>
<tr>
<td>1996</td>
<td>2.8</td>
<td>5.1</td>
<td>13.8</td>
<td>-1.4</td>
<td>52.9</td>
</tr>
</tbody>
</table>

Source: OECD Main Indicators Database, ASPE (1998); US Bureau of the Census.

The federal deficit had been shrinking steadily since 1992, owing to tax increases, structural spending reductions and the growth of the tax base, and fiscal outcomes were consistently better than forecast by the Clinton administration or outside observers. Nevertheless, deficit reduction remained a priority for both Congress and the White House. The Republicans’ determination to achieve substantial savings via welfare reform, in order to meet their deficit-reduction targets while still pressing for the tax cuts that they had promised in the 1994 mid-term elections, meant that fiscal considerations strongly influenced their approach to reform.

**Actors and interests**

During the first two years of the Clinton presidency, welfare policy was largely in the hands of the Department of Health and Human Services (HHS), which was charged with preparing the reform legislation that the administration hoped to pass through the then Democratic Congress. After the Republicans took control of Congress in January 1995, HHS was largely sidelined and the process centred on the interaction between the Democratic White House and the Republican congressional majority. Indeed, HHS in some cases tried unsuccessfully to dissuade the White House from making some of the compromises it reached with congressional Republicans.\(^4\) The administration’s room for manoeuvre was affected by long-term changes in the Democratic congressional caucus brought about by population shifts, in conjunction with congressional redistricting.\(^5\) Welfare recipients tended to be heavily concentrated in large urban centres, and by the mid-1990s, the proportion of congressmen representing predominantly urban districts had been falling for decades – down from roughly one-quarter in 1982 to around 17% in the 1990s. The number of predominantly suburban districts had risen from 29 to 48% over the same period.\(^6\) On balance, this trend favoured Republicans and centrist Democrats, who were less anxious to defend welfare programmes. Many Democrats were ready to embrace welfare reform, arguing both that the *status quo* was failing and that reform
would help the Democrats shed their image as a “tax-and-spend” party that was “soft” on welfare.

The other key political actors were state governors. The programmes at issue were all administered by the states and most were substantially state-financed as well. In general, the states wanted greater flexibility in running social assistance programmes, and by early 1996 governors began to coalesce around a package of proposals that would give them that.

Those likely to be most directly affected by the reform – transfer recipients – were not organised and were less likely to vote than the average citizen, so their interests tended to be represented by welfare-rights advocacy groups and politically engaged think-tanks. While several of these commanded considerable respect in both parties, their resources were limited, as was their ability to mobilise grassroots support or to engage more powerful interests as allies in lobbying Congress, and their general orientation towards the Democrats greatly reduced their influence after the Republicans took control of Congress in 1994. The Republicans tended to look to more conservative institutes for input and support. Business lobbies, too, were supportive of welfare reform, seeing it as a means of increasing labour supply in a tightening labour market. However, there was little active lobbying by business groups, apart from the mobilisation of certain sectoral lobbies against elements of the reform they found threatening.

**Issue history**

In the early 1990s, America’s social safety net comprised a range of means-tested programmes offering limited support for very low-income families, particularly those with children (Box 16.1). Joint federal-state financing and administration meant that these programmes typically involved a complex mix of federal requirements and state options. By 1994, expenditure on the most important welfare programmes stood at roughly 3.3% of GDP, with Medicaid alone accounting for around 60% of the total. The best known and perhaps least popular major welfare programme was Aid to Families with Dependent Children (AFDC), created in 1935 as a small programme to support widows with minor children. AFDC rolls began growing rapidly in the mid-1960s, rising from around 4 million in 1964 to 11 million by 1973, largely as a result of benefit increases granted by state governments, court decisions on eligibility and the growing willingness of eligible individuals to claim benefits. In response, most states stopped raising benefits at the rate of inflation; increases in Food Stamps and other in-kind benefits only partially offset the decline in the real value of cash assistance, so that by the early 1990s, AFDC and Food Stamps in most states were insufficient to lift a family to even 80% of the poverty line. Between 1989 and 1994, welfare rolls surged again (Table 16.1), driven by rising rates of unwed pregnancy and, latterly, a recession. This increase was a major factor pushing welfare reform onto the policy agenda in the early 1990s.
Box 16.1. Federal social assistance programmes before PRWORA

1. From its birth in 1935, AFDC had been a federal entitlement – it guaranteed welfare support to those who met the eligibility requirements. However, the extent of this entitlement was somewhat ambiguous, since it was the states that set benefit levels and eligibility rules, within federal guidelines. Under AFDC, federal matching funds were provided to the states without limit and depended on the state’s chosen benefit level, the number of recipients in the state and the state’s “percentage match”. Thus, the state’s entitlement to federal funds was clear, but the individual recipient’s entitlement to benefit depended on where he/she resided. AFDC rolls peaked at 14.2 million in 1994, when total federal and state AFDC expenditure reached USD 26.1 billion.

2. Supplemental Security Income (SSI) is a Federal income supplement program funded by general tax revenues and designed to help aged, blind, and disabled people with little or no income, to meet basic needs for food, clothing, and shelter. In 1994, SSI expenditure reached USD 31.8 billion, including both administrative costs and benefits paid to some 6.3 million persons.

3. Administered by the US Department of Agriculture, the Food Stamp programme is a federal entitlement that aims to increase the food purchasing-power of low-income families. Food stamp benefits make up the difference between a household’s expected contribution to food costs (around 30% of counted cash income) and an amount judged to be sufficient to buy an adequate low-cost diet. In 1994, 28.8 million people received food stamps, and total food stamp expenditure reached USD 27.3 billion.

4. Medicaid is an entitlement program providing medical and health-related services to people with limited income. It is jointly funded by the states and federal government, and managed by the states. While the vast majority of Medicaid beneficiaries have always been children and working-age adults, by the early 1990s the great bulk of Medicaid expenditure was on the disabled and long-term care for the elderly. Combined federal and state spending on Medicaid totalled USD 122.9 billion in 1994.

5. The Earned Income Tax Credit (EITC) is available to households with earnings, with the size of the credit rising as earnings increase; it can both offset taxes and provide a wage supplement. First introduced on a very small scale in 1975, the EITC was expanded substantially in 1986 and 1990 and doubled in size in 1993. By 1996, some 18 million taxpayers were receiving the credit, claiming a total of roughly USD 25 billion, around 86% of which was refunded as direct payments to families (Ways and Means, 1996). A single mother with two children working half-time at the minimum wage could, with the benefit of the EITC, earn more than her welfare benefits would have been in most states (Ellwood, 1990).

* States did not receive dollar for dollar what they spent on welfare: federal funds matched a percentage of state expenditure, which varied with the per capita income of the state. States with high per capita income, like New York, received USD 0.50 for every dollar spent, while poorer states received up to USD 0.80.

Welfare reform debates

AFDC was subjected to increasing criticism from the late 1960s onwards, largely on account of concern about fraud and the perception that many adults who were able to work were choosing not to. Early welfare reform efforts thus addressed perceived “abuses”, focusing on purging ineligible individuals from the rolls and reducing overpayments. Over time, the profile of AFDC recipients was changing in ways that
profoundly altered the terms of the debate over welfare reform. Families in which the father was deceased became eligible for (usually more generous) Social Security benefits and largely dropped off the AFDC rolls, so by the 1980s, the vast majority of adults on AFDC were single mothers. Almost none were employed, because benefit withdrawal rates had been increased sharply in the early 1980s, effectively removing most working recipients from the AFDC caseload. At the same time, attitudes towards female labour-force participation were also changing. When the programme was created, labour-market attachment was not among its aims – widowed mothers were expected to stay home and AFDC was intended to help them do that; a half-century later, however, a growing majority of non-AFDC mothers were working, while 90% of AFDC mothers were not. This made the lack of any emphasis on getting recipients into work increasingly difficult to defend. Over time, criticisms of welfare went further, as it was blamed for family breakdown, rising rates of teen pregnancy and unwed births, and the emergence of pockets of concentrated poverty in the cities. Given that members of ethnic minorities dominated the AFDC rolls, such arguments could not but tap into racial sensitivities, especially since much of the public discussion was based on anecdotal and impressionistic evidence. This contributed to a polarisation of the welfare reform debate.

Three lines of research in the 1980s, however, began to change the debate substantially. First, Murray (1984) advanced the case that existing welfare programmes were bad for the poor; he argued that they created a system of perverse incentives, encouraging out-of-wedlock births, discouraging labour-market entry and lowering the value of learning. Secondly, Bane and Ellwood (1986) found that almost half of the AFDC caseload at any given time had been drawing AFDC for over two years and that around 65% of recipients on the AFDC rolls at any given time would spend 8 years or more on AFDC. Their results, and others like them, were important in shaping the debate on welfare dependency and the need to time-limit assistance. Finally, a research programme undertaken by the Manpower Demonstration Research Corporation (MDRC) showed that tightly administered welfare-to-work programmes that emphasised job search consistently produced significant increases in employment and earnings, while reducing welfare rolls. In addition to these, there was a growing body of evidence that public welfare did reduce work. The empirical evidence concerning the claim that welfare contributed to illegitimacy and related social ills was far less clear, as even some advocates of reform acknowledged, but the argument made sense to many voters and was widely accepted by politicians in both parties; it thus became an important “political fact” in the reform debate.

A final issue in the welfare reform debates of the 1990s – and one that had a big impact on the reform that was eventually adopted – concerned evidence of rising rates of enrolment among non-naturalised immigrants. Overall, immigrants were about one-third more likely than natives to access welfare programmes, and non-naturalised immigrants were far more likely than naturalised US citizens. Immigrants accounted for 13.3% of adult AFDC recipients and their share was rising. Moreover, although only 1.5% of SSI recipients were aliens, non-citizens accounted for 28.2% of elderly recipients of SSI, largely because many elderly immigrants had not worked enough quarters in covered US occupations to qualify for Social Security benefits. For many immigrants, too, SSI was a bridge to Medicaid, which was crucial given their ineligibility for Medicare.

Welfare reform prior to PRWORA

Apart from the cuts implemented in the Reagan administration’s budget package at the beginning of the 1980s, the only major federal-level reform of AFDC prior to
PRWORA was the Family Support Act (FSA) of 1988. Heavily influenced by research like that of Bane and Ellwood and the MDRC, the FSA represented a compromise between liberals and conservatives on the basis of a broader safety net combined with stronger work requirements and more rigorous enforcement of child support in respect of non-custodial parents. The FSA provided transitional childcare and Medicaid for recipients leaving welfare for work, and it imposed some modest work activity requirements via the Job Opportunity and Basic Skills (JOBS) Training Programme. The states were required to have 10-15% of AFDC families in work-related activities by 1993. The net new cost of the FSA was estimated at USD 3.3 billion over four years, mainly for training, childcare and Medicaid. For the working poor, the most significant change in policy was the steady expansion of the EITC, which grew almost 7-fold in real terms from the mid-1970s to the mid-1990s. The empirical evidence overwhelmingly suggests that the growth of the EITC increased the labour supply of single parents and raised the incomes of the working poor substantially.21

Efforts to reform welfare at federal level coincided with a rapidly growing trend towards the conduct of welfare reform “experiments” at state level. In 1962, states gained the power to waive, with the permission of the federal government, certain provisions of national welfare law in order to experiment with reforms to AFDC. Such waivers were uncommon until the 1980s, but during 1982-95, almost all states applied for at least one, and the average state government applied for three. Moreover, the willingness of the federal authorities to grant waivers also grew over time. During 1993-96 alone, 43 states conducted 78 welfare reform “experiments” under waivers granted by the administration. The Clinton administration was generally reluctant to deny waiver requests, with the White House occasionally overruling HHS to grant them.22 Many of the most prominent elements of PRWORA were first implemented on a provisional basis under state waivers,23 suggesting that the 1996 reform was the culmination of a gradual relaxation of the federal government’s grip on welfare.24

The evolution of public opinion

Polling data across six decades demonstrate a certain ambivalence in the electorate’s attitude to welfare: strong support for helping people in genuine need has always been in tension with a belief in self-reliance and individualism that left voters wary of supporting those who could support themselves. During the early 1990s, public support for welfare fell sharply; a large and growing majority of the public – and, indeed, of welfare recipients themselves – believed that welfare was not working, and voters increasingly associated welfare with fraud, social decay and dependence.25 This erosion of support appears to have reflected in part increased media attention to the issue and the tenor of the political debate.26 However, it may also have been a natural response to the Democrats’ recapture of the White House. Survey data over the long run reveal that support for spending on the poor has tended to rise when Republicans controlled the White House and to fall under Democratic administrations.27 Uncertain of the appropriate level of welfare spending, voters appear to believe that whichever party is in power will go too far – the Democrats will be too generous and the Republicans not generous enough. Polls in the mid-1990s, moreover, did not show a clear, consistent advantage for one party or the other when respondents were asked whom they trusted more on welfare reform.28 This suggests that the electorate saw the issue as one that needed a bipartisan solution.

Public opinion varied considerably with respect to forms of assistance and categories of recipient. Cash assistance to able-bodied, working-age adults has historically been least popular, and by 1995, support for it had been declining steadily for a generation. By
contrast, aid to children, the elderly and the disabled was consistently popular, as was support for nutrition programmes and for training and employment programmes designed to get recipients off welfare. During 1994-96, polls showed very high (80-90%) levels of support for job-training and the creation of public sector jobs, as well as strong support for childcare and transport subsidies. Work requirements after two years were popular as well. Voters were willing to exempt mothers with small children from such a requirement, but substantial minorities did favour limiting benefit increases to mothers who bore children while on welfare and denying benefits to unwed teenage mothers who did not live with an adult.  

**Policy process: formulation**

In 1992, the successful Clinton campaign had made welfare reform a prominent part of its electoral platform. The general approach promised during the campaign was to combine stronger pressure for labour-market re-entry with more support for those returning to work, including employment subsidies, if necessary. Initially, the newly elected administration focused on healthcare reform and deficit reduction, but a Working Group on Welfare Reform, Family Support and Independence was appointed in 1993. Its mandate was to craft a legislative proposal based on four principles: (1) making work pay by creating incentives that ensure that benefit recipients who went to work would benefit; (2) strengthening enforcement of child-support payments by non-custodial parents; (3) providing education, training, childcare and other supports to help transfer recipients return to work; and (4) creating a time-limited transitional support system followed by work. These principles commanded such wide agreement among both Democrats and Republicans that they may be seen as having constituted a basic consensus. It was generally accepted that an adult judged able to work should not receive assistance indefinitely without ever working, and it was recognised that significant resources should be devoted to helping recipients back into the workforce. However, there was no consensus concerning how these principles might be realised in practice – what the balance between support and sanctions should be, how firm the time limits should be and to whom they should apply, or how extensive the support for the transition to work would need to be.

The administration finally put forward its welfare reform bill in mid-1994. It proposed a two-year lifetime limit on receipt of cash assistance by those who could work, to be followed by mandatory public service if other employment could not be found. Exemptions were allowed for the seriously ill and for those caring for a child with a serious illness or disability. States would be allowed to extend eligibility for those in education or training programmes and for those facing particular obstacles to labour-force entry, such as learning disabilities. However, such extensions could not exceed 10% of a state’s caseload. In addition, the plan excluded legal immigrants from access to benefits during their first five years of residence and limited their access thereafter. The other key element of the reform plan was substantially increased public expenditure on training, childcare, transportation and other services to support those making the transition from welfare to work. Paid work assignments were to be provided to those who could not find private-sector employment. While the plan stopped short of ending AFDC or Food Stamps as an entitlement or devolving responsibility for welfare to the states, it did envisage giving them greater flexibility in designing and administering their programmes.
While the president’s working group was preparing its bill, congressional Republicans were preparing their own, more radical alternative, the major elements of which were included in the “Contract with America”, which formed the basis for the party’s mid-term congressional campaign. The Republicans aimed to cap federal funding for a range of social assistance programmes and to roll these funds into a single “block grant”, which would then be awarded to the states to use as they saw fit, within certain basic federal guidelines. This would end welfare as a federal entitlement: even if an individual or family met all of a state’s eligibility requirements, the state would not be legally obliged to provide benefits if state funds combined with the federal block grant proved insufficient to cover all eligible beneficiaries.

While the Republicans had to negotiate a large number of issues among themselves in order to reach a common position, two issues stood out. The first was illegitimacy: social conservatives within the party regarded reducing out-of-wedlock births as a policy goal at least as important as labour-force activation, if not more so, and they won inclusion in the “Contract” of provisions that would severely limit unwed young mothers’ access to assistance. The second problem was the formula for the block grant, an issue which could have provoked considerable conflict among the states. The Republican proposal simply set the block grant in the first year at 95% of the aggregate funding of the individual programmes involved. The grant would thereafter be adjusted for inflation and changes in the population in poverty. This approach would have preserved the existing highly unequal distribution of federal welfare funds among the states – at the time, there was a roughly 12-fold difference in federal AFDC spending per poor child among the states – but by ensuring that no states emerged as conspicuous winners or losers, it avoided political conflict.

Policy process: adoption

The reform impasse

Even before the 1994 mid-terms, the administration’s bill had little chance of moving forward. Few congressional Democrats wished to support the plan, finding the work requirements and time limits too harsh, while Republicans insisted that it did not go far enough, arguing that it would perpetuate a system that was more generous to benefit recipients than to the working poor. Republicans also objected to the fact that the short-run impact of the administration’s plan would be to increase spending, owing to the need to provide meaningful assistance to mothers on benefit who were returning to work. Nor was the administration proposal popular with the states; they disliked its highly prescriptive character, which they felt would leave them too little flexibility and which they found to be at odds with the administration’s relaxed approach to the granting of waivers. It is hardly surprising, therefore, that the Democratic leadership in Congress did not take up the White House bill. In any case, since 1994 was an election year, the administration understood that action was unlikely on such a sensitive issue, and it planned to return to welfare reform in 1995. The mid-term elections, however, gave the Republicans control of both houses of Congress and a net gain of eleven governorships. As a result, the administration effectively lost control of the welfare reform process; the working group’s bill was not even re-submitted in the 104th Congress.

The new Republican majority put forward a bill based on its “Contract with America” proposals in late 1995, as part of the 1996 Budget Reconciliation Act. It provided for a five-year life-time limit on cash-assistance, requiring work after two years of any individual spell on welfare, and denying benefits to unwed teen-age mothers and children.
born to welfare recipients. It would also have restructured federal funding for AFDC and other cash assistance programmes into a new programme called Temporary Assistance to Needy Families (TANF), which would be given as block grants to the states. Altogether, the welfare provisions of the bill (including EITC cuts and changes to other non-TANF programmes) would have saved an estimated USD 81.5 billion over seven years. States receiving the TANF block grant would be subject to a series of targets for participation rates of adult TANF recipients in work-related programmes, reaching 1.5 million by 2001. The bill passed both houses of Congress on party lines but was vetoed by the president. The Republicans tried again in January 1996 with a separate Welfare Overhaul Bill, which essentially repeated the main provisions of the earlier proposal and would also have capped spending on Food Stamps, limiting maximum benefit increases to 2% per year and offering the states the option of taking Food Stamp funds as a block grant. This would have ended Food Stamps as a federal entitlement. The Republicans also proposed capping the growth of Medicaid, which would, in practice, have amounted to turning that programme, too, into a block grant. The administration argued that the bill did too little to help benefit recipients move into work, and the White House Office of Management and Budget concluded that it would push over a million children into poverty. Backed by congressional Democrats and child advocacy groups, the president vetoed it.

The PRWORA compromise

At this point, the nation’s governors stepped in to try to broker a compromise. In February 1996, the National Governors Association unanimously endorsed a set of proposals for modifying the Republican bill in order to turn it into something the states could support. Their demands included, inter alia, a big infusion of childcare funding outside TANF, reductions in the proposed savings from curbing the EITC, greater flexibility in the states’ use of TANF dollars, a contingency fund states could access in a severe economic downturn, less stringent targets for the states regarding recipients’ participation in work activities, and provisions for including families that left welfare for work in the calculation of those participation rates. The states also insisted that they be able to retain any waivers that were at odds with the new legislation until those waivers expired. The governors had been discussing a common approach to welfare throughout 1995, but the president’s veto message helped tip the balance. Although the administration did not favour block-granting welfare programmes, the veto message made no reference to the issue. This allowed Democratic governors to come out in favour of block grants for some programmes. The idea also appealed to many Democrats in Congress, since it would pass the issue of welfare to the states once and for all, and Democratic candidates in federal elections would no longer have to carry this “political albatross” around their necks.

The bill that became PRWORA thus emerged as a compromise between the administration and the Republican congressional majority that drew on many of the governors’ ideas. There were no real negotiations between Congress and the White House, apart from some informal contacts, and there was no administration counter-proposal on the table. Instead, the Republican leadership simply removed from the bill a number of so-called “poison pills” – elements that would effectively guarantee a veto, like the caps on Food Stamps and Medicaid. Many of the provisions were in any case opposed by the governors and antagonised some Republican interest groups, as well as the president. The really difficult issue for the Republicans was whether they wished to pass a welfare bill in 1996 at all. Many were reluctant to allow the president to campaign for re-election having just signed a major welfare-reform into law. However,
congressional Republicans were in any case pessimistic about re-capturing the White House and increasingly concerned about their own failure to enact more of their 1994 “Contract with America” agenda. Adoption of PRWORA would enable them to show that they were moving forward on their promises to reform welfare, cut the deficit and devolve more power to the states, all at the same time. In the end, therefore, they crafted a more moderate bill that they thought the White House might accept.

PRWORA replaced AFDC with TANF and authorised six years of TANF block grants to the states, fixed at USD 16.5 billion per year. States receiving TANF funds would have to maintain 75-80% of their own pre-reform AFDC spending and to meet strict, escalating targets for participation in work activities; 50% of their caseloads were to be enrolled in such activities by 2002. A state’s TANF block grant allocation was to be based on the highest of: its FY 1994 spending, its FY 1995 spending, or the average of its spending for FY 1992 through FY 1994. Thus, no state would be worse off under TANF than it would have been under the old rules. Indeed, since AFDC rolls had begun falling, the states would, in the short term at least, profit from the block grant. In addition to ending unlimited matching funds for AFDC (and thus any notion of a federal entitlement), the bill reduced federal spending on Food Stamps by 13% by tightening the eligibility requirements, and it restricted legal immigrants’ access to benefits under TANF, Food Stamps and other programmes.

As a concession to the administration, the third bill included much more generous provisions regarding support for people moving into work than the Welfare Overhaul Bill had done: it retained guaranteed Medicaid for recipients of cash assistance and increased childcare funding. The toughest questions concerning illegitimacy and benefits for teenage mothers were largely left to the states, and proposals to reduce SSI expenditures on disabled children by tightening the definition of “disability” were softened. PRWORA also included a provision requiring the states to maintain their own spending on childcare at 1994 or 1995 levels (whichever was higher) in order to access the additional federal funds; the previous bills had contained no such “maintenance of effort” requirements, prompting fears that the states would just take the federal monies and cut their own spending. The bill was projected to save USD 55 billion over six years, almost all as a result of cuts in non-AFDC programmes, especially Food Stamps (USD 28 billion) and SSI. Restrictions on legally resident aliens’ eligibility for these two programmes were to generate an estimated 40% of the total savings from PRWORA.

Democrats were unsure of how to respond to the new bill. Leading HHS officials strongly opposed it, arguing that up to 2.6 million people, including 1.1 million children, would be pushed into poverty, but a minority of congressional Democrats supported the revised bill and were even prepared to help the Republicans override a veto if need be. Another minority, on the left, was firmly against it. In between, there were many prepared to vote for it if the president signalled its acceptability, thereby providing them the political cover they needed. Despite misgivings about the bill and, in particular, about what he regarded as overly harsh treatment of immigrants and the insufficient commitment to support poor mothers making the transition into work, the president signalled his readiness to sign PRWORA into law, thereby allowing large numbers of Democrats to vote for the final version; a veto override was in any case a real possibility and would have meant that the Democrats could not share any of the credit for a popular reform in an election year. However, the administration promised to pursue changes in the law’s harshest provisions, including the restoration of some Food Stamp spending and the relaxation of restrictions on immigrants’ access to benefits. In the 1998 budget, it did
indeed win USD 3.3 billion for a new welfare-to-work initiative for long-term AFDC dependents and the softening of provisions affecting immigrants. Some Food Stamp cuts were also reversed.52

**The fate of Medicaid and Food Stamps**

PRWORA’s adoption was made easier by the fact that those most directly affected by the overhaul of AFDC were poorly organised and represented. The provisions concerning Food Stamps and the Republicans’ decision to leave Medicaid out of the reform illustrate how much more difficult it was to bring about change when powerful vested interests felt threatened. Unlike earlier Republican bills, PRWORA did not dismantle Food Stamps or Medicaid and turn them into block grants.53 Nor did it include an expenditure cap on Food Stamps, which thus remained a federal entitlement for those who met the revised eligibility criteria. The preservation of these two programmes owed much to the power of specific lobbies that supported them:

- There was never much chance that Medicaid would be block-granted or capped. In early 1996, a rumour that the president was considering this Republican proposal was sufficient to trigger the brief but effective mobilisation of lobbies representing the elderly, including the powerful American Association of Retired Persons (AARP) and the healthcare industry. The idea was soon dropped.54 The AARP and its allies saw Medicaid as a crucial programme for seniors in need of long-term care, while the healthcare lobby was also deeply attached to it.55 Moreover, most governors did not wish to block-grant Medicaid, fearing that the states would then be saddled with the burden of spiralling healthcare costs.56

- The governors were divided over the desirability of a Food Stamps block grant, and farm-state Republicans, both in Congress and in state governments, were strongly against the idea. A US Department of Agriculture analysis had concluded that a Food Stamp block grant would lower total nutrition spending in FY 1996 by more than USD 5.1 billion (12.7%), reducing farm incomes.57 In any case, nutritional assistance was more popular than cash handouts, and many congressional Republicans had long supported the Food Stamp programme. The compromise reached, therefore, was to preserve Food Stamps as an entitlement but to tighten the eligibility rules. Since the savings counted towards the deficit-reduction targets assigned to the House and Senate agriculture committees by the chambers’ budget committees, other forms of farm support suffered fewer cuts, an outcome the farming lobby welcomed.58

**Policy process: implementation**

PRWORA’s emphasis on devolution meant that it was implemented overwhelmingly at state level. Contrary to the expectations of many critics, however, it did not lead to a “race to the bottom” in state welfare policies.59 While there was a great deal of adjustment and experimentation at state level, there was no evidence of convergence on a single model.60 This was particularly true in respect of the treatment of immigrants – altogether, some 28 states substantially softened the provisions affecting non-natives in one way or another.61 TANF also provided an immediate windfall to the states, since welfare rolls shrank while the block grants did not.62 By 2000, AFDC-TANF rolls were down 56% from their 1994 peak. This allowed many states to cut taxes or increase spending on other things. By the end of FY 1999, the states had accumulated roughly USD 7 billion in unspent TANF funds.63
The sustained strong growth of the 1990s facilitated the implementation of PRWORA by providing employment opportunities for those under pressure to leave welfare. Employment rates among single mothers rose from 63.5% in 1996 to 72.8% in 2000, and the total income of the bottom quintile of families with children rose by 24.5% over the period, thanks to a 57.8% increase in their earnings. Many former welfare recipients went on to find and hold jobs that provided them with greater financial and family stability. The overall poverty rate fell 2.4 percentage points during 1996-2000, while the child poverty rate fell by 4.3 percentage points. There was – and is – much debate about the extent to which cyclical factors were responsible for these positive outcomes, but a significant body of research suggests that PRWORA, including the sanctions on those who did not comply with the legislation’s work requirements, and the expansion of the EITC both played a significant role. There were, nevertheless, losers under PRWORA – former AFDC recipients who failed to obtain jobs ended up worse off than under AFDC. This was a significant group: the proportion of single mothers with neither employment nor welfare rose from 16.4% in 1996 to almost one-third in the early 2000s. Empirical research suggests that such “disconnected” families were disproportionately likely to face exceptional barriers to labour-market entry, such as addictions, health problems or learning disabilities, and that children in such families were at significantly greater risk of nutritional or health problems. Also among the losers were those who found themselves in unstable, low-wage employment. The income growth of the lowest decile of families with children was slower in the first years after the reform than it had been in 1992-96.

The new system’s performance during the economic slowdown of the early 2000s triggered much debate. Unexpectedly, the TANF caseload continued to fall in 2000-02, even as unemployment and poverty began to rise, leading critics to insist that the system was not fulfilling its safety-net functions. However, this seems to have reflected, at least in part, the fact that the 2001 recession did not much affect the consumer sectors where low-skilled women were most likely to be employed. Overall, therefore, TANF was generally regarded by both parties as a success, even if there was continuing debate about the steps needed to address the needs of those especially disadvantaged families who had lost out under PRWORA. It was therefore expected that reauthorisation of PRWORA in 2002 would be fairly automatic, but the legislation was not in fact reauthorised until 2006, owing to conflicts over new funding for childcare, marriage promotion, the level of increase in the TANF block grants and changes in work-participation requirements. In the end, the block grants were left untouched through 2010 – they were not even indexed for inflation. Childcare funding was increased by USD 200 million and a further USD 150 million/year was allocated for marriage promotion/fatherhood programmes. However, the most significant change concerned the work participation requirements: here, the legislation became much more prescriptive, tightening definitions and adding new rules in an effort to achieve greater uniformity across the country. This greatly reduced the states’ freedom of action, which had been one of the governors’ key priorities in 1996.

Conclusion

By the mid-1990s, virtually no one in US political life wished to defend the status quo with respect to welfare. The elements of a compromise were already there, inasmuch as both parties were in broad agreement on the need to shift the balance between the safety-net and labour-market goals of the major welfare programmes. There was broad
agreement that assistance to adults able to work should not be extended indefinitely without any requirement to work and that helping welfare beneficiaries return to employment would require resources for education, training, childcare and other forms of assistance to facilitate the transition from welfare to work. This emerging consensus resulted not only from mounting evidence that federal welfare programmes were performing poorly but also from a growing body of empirical research that identified some of the causes of that failure, as well as potential solutions.

While there were sharp disagreements about how to apply these principles in practice, the existence of a consensus on the basics made welfare reform much more tractable than issues like Social Security (pensions) and healthcare, where no such area of common agreement existed. Moreover, both the Democratic White House and the Republican congressional majority believed that they needed to deliver a welfare reform ahead of the 1996 elections. Both could claim a share of the credit for the generally popular reform of a very unpopular programme. For Democrats, PRWORA, though based more on Republican proposals than those of the administration, had the added attraction of passing welfare to the states, thereby relieving Democratic candidates in federal politics of the need to defend it. Welfare reform also offered significant administrative and fiscal benefits to state governments, a fact that prompted the governors’ decision to intervene in the federal welfare reform debate in order to secure a compromise between the White House and Congress. The compromise that emerged ensured that no state would be worse off financially as a result of the reform. While this preserved significant inequalities in the allocation of federal funds, it ensured that PRWORA was not blocked by disputes over the grant formula, as has often been the case with block-grant programmes.

The final shape of the reform owed much to the Republicans’ fiscal goals: whereas Democratic advocates of reform expected it to cost more than the status quo in the first instance, the Republicans were determined to make significant savings up-front. This underlay many of the legislation’s provisions, particularly those regarding immigrants, as well as what critics believed were the inadequate resources devoted to supporting those in transition from welfare to work. In addition, the block granting of AFDC, which gave the states an initial windfall, led to significant federal savings later on, as a result of non-indexation of the grant combined with a tightening of federal regulations. The intervention of some powerful organised interests also affected the bargains that were done as PRWORA was negotiated. The agrarian lobby played a key role in maintaining Food Stamps as a federal entitlement, while the healthcare industry and the “grey lobby” defended Medicaid. Finally, there is little doubt that the robust economic conditions of the 1990s facilitated the implementation of PRWORA. Such a reform would have imposed much higher costs in a recessionary environment.
Notes

1. Changes adopted in 1997 did give the states some flexibility on these and other points, but much of this flexibility was lost when PRWORA was re-authorised in 2006. See below for details.

2. See OECD (1997); CEA (1997); Murray and Primus (2005).

3. For all households with children, the lowest quintile experienced faster income growth during this period than did any but the top 20%.


5. The impact of the flight to the suburbs and exurbs on congressional districts was compounded by the practice in many places of creating compact, minority-dominated districts, often in urban centres, in order to facilitate the election of minority candidates to Congress. While this helped increase the number of representatives of ethnic minorities in the House of Representatives, it reduced the number of districts representing urban centres.

6. The number of rural districts fell sharply over the period and the number of mixed districts was roughly stable. See Caraley (2002:546-48).


8. The US Chamber of Commerce identified welfare reform as one of its top priorities in 1996, and other business groups expressed support for the Republican reform agenda. See Klarner et al. (2007).


12. Since 1967, benefit recipients could earn USD 30 per month before losing any benefits and then lost USD 2 in benefits for every USD 3 earned. After 1981, benefit withdrawal was dollar for dollar once an individual had been receiving benefits for four months. Paradoxically, then, the drive to intensify work requirements came at the expense of concern with the lack of work incentives and work opportunities.


14. The data in ASPE (1998) show that the share of African-Americans among AFDC recipients was about triple their share of the US population as a whole.

15. On their impact, see Weaver (2000); Haskins (2006).

17. See Danziger (1999); Haskins (2006:–8). By the mid-1990s, around 18% of the 1.4 million adult recipients of AFDC were single mothers under the age of 20; see ACF (1994).


19. ACF (1995). Many of their children had been born in the United States and were therefore citizens.


21. See Greenstein and Shapiro (1998); Eissa and Leibman (1996); Eissa and Hoyes (2005); Holt (2006); OECD (2007). Studies have found some negative effects on the labour-supply of secondary earners but these are small and more difficult to discern.

22. HHS (1996). The surge in waivers actually began in 1992, but it is not clear why. One hypothesis is that the Bush administration was trying to get out in front on an issue that Democratic candidate Bill Clinton was pushing hard; alternatively, the process may have been demand-driven, as reform-minded Republican governors sought to secure waivers before the Democrats could take power. In any case, the Clinton administration’s policy towards waivers was extremely liberal.

23. Including the family cap and time limits on cash benefits.


25. See the data presented in Weaver (2000), chapter 7; and in Schneider and Jacoby (2005).

26. Schneider and Jacoby (2005) argue strongly for the impact of these factors.

27. See Weaver et al. (1995); and AEI (2003).


29. For detailed discussion of this topic, see Weaver et al. (1995); Danziger (1999); Weaver (2000), chapter 7; and AEI (2003).

30. An immigrant’s sponsor was to be financially responsible for him/her for five years; after that, the immigrant would be eligible only if the sponsor’s income were at or below the national median family income or the immigrant became a citizen.


32. The plan allowed, *inter alia*, for states to limit benefit increases when beneficiaries had more children.

33. Including AFDC, SSI, public housing benefits and ten nutrition programmes, including Food Stamps and the school lunch programme.

34. The general block grant proposed in the Contract with America was rejected by Republicans on the House Ways and Means Committee and dropped soon after the election.

35. Unmarried mothers under 18 would be ineligible and states would have the option of declaring those aged 18-20 ineligible. Mothers would be required to establish paternity to receive benefits, except in cases of incest or rape.

36. The draft text of the welfare bill presented as part of the Contract with America can be accessed at www.house.gov/house/Contract/persrespb.txt.
37. These differences reflected the interaction of state policies and variations in the states’ “percentage matches” under AFDC. Food Stamp spending varied much less, as it was a wholly federal programme and differences in expenditure largely reflected differences in price levels. Altogether, in 1995, there was a roughly 4-fold difference in AFDC and Food Stamp spending per poor child between the least and most generous states, excluding Hawaii and Alaska, the outliers on the upper end of the spectrum. See Douglas and Flores (1998).

38. There was never even a committee vote on it.


40. Since the states already had enormous leeway in administering Medicaid, there was little difference in practice between a cap and a block grant.


42. The campaign of Republican presidential nominee Robert Dole strongly pressured the Republican congressional leadership to keep Medicaid reform bundled with AFDC reform, in order to ensure that the president could not sign it; see Brookings (2006:20).

43. See Weaver (2000) and Haskins (2006).

44. PRWORA divided legal immigrants into two groups: those who were in the country as of 22 August 1996, when the bill was enacted, and those who came afterwards. Most immigrants arriving after PRWORA were barred from federal means-tested benefits, including SSI, Food Stamps, TANF, non-emergency Medicaid and the state Child Health Insurance Program (CHIP) for the first five years. States were free to bar earlier arrivals from TANF and non-emergency Medicaid programs. Since 1996, however, eligibility for SSI and Food Stamps has been restored to many immigrant children, elderly, and disabled persons who arrived prior to PRWORA. See Zimmerman and Tumlin (1999).

45. Specifically, children and adults eligible for Medicaid under existing AFDC rules would remain eligible for Medicaid even if they no longer qualified for cash assistance under TANF.

46. The vetoed welfare bill would have denied SSI benefits to a large number of children who qualified as disabled under existing guidelines and would have reduced benefits by 25% to most disabled children who remained on the programme. In general, assistance for children was relatively popular, but SSI was a special case. In 1990, the US Supreme Court had held that the test for determining whether children were “disabled” was unlawfully restrictive. The number of children in the programme subsequently grew dramatically, prompting claims that children were being “coached” to misbehave in order to win disabled status. The GAO concluded that many children on SSI were not disabled on any reasonable definition; hence the Republicans’ desire to curtail this programme.

47. Super et al. (1996).


49. This claim was based on an Urban Institute study of the earlier Republican bills; HHS officials claimed that nothing in the new version would affect these findings.
50. Weaver (2000) shows how many congressional Democrats moved with the president, trying to avoid positions to his left (which could be seen as “soft on welfare”) or too far to his right (which could look too “Republican”).

51. In the Senate, only one Democrat facing re-election in 1996 voted against the bill.

52. See Greenstein and Primus (1997).

53. The vetoed Welfare Overhaul Bill gave states the option of block granting Food Stamp benefits.


55. Critics argued that the industry feared that further cuts in the provision of basic healthcare services to the needy would in the long run increase the pressure for comprehensive healthcare reform.

56. In March 1996, an attempt by some Republicans to claim bipartisan support for a Medicaid block grant among the governors backfired when Democratic governors denounced it.


58. King (1999). If the committees of each house do not write and pass the legislation needed to meet the targets specified by the budget committees, the budget committees have the power to write the legislation themselves; Haskins (2006:242-43).

59. See Notes (1996) for warnings that PRWORA would lead to a “race to the bottom”.


64. This continued a trend that had begun in 1993, when the employment rate of this group was just 57.3%. The employment rate for mothers with very small (under 3 years of age) children also rose, by roughly 12 percentage points, to reach 57% at the end of the 1990s.

65. Unless indicated otherwise, all data in this paragraph are from CRS (2007).


67. A number of these studies are surveyed in Greenberg (2007).


70. See Brookings (2006:46); the slowdown hit manufacturing and traded goods sectors hardest.

71. One might see this as a protection of “acquired rights” in the course of a reform, but they were the acquired rights of the states, not of benefit recipients.
Bibliography


CBO (1993), Forecasting AFDC Caseloads, with and Emphasis on Economic Factors, Congressional Budget Office, Washington, DC.


Greenberg, M. (1993), “The Devil is in the Details: Key Questions in the Effort to ‘End Welfare as We Know It’”, Center for Law and Social Policy, Washington, DC, July.


Poverty Research News 4:4, July-August.

Law, Brookings Institution Press, Washington, DC.

Reconciliation Act of 1996”, Department of Health and Human Services, Washington, 

Brief, The Brookings Institution Metropolitan Policy Program, Washington, DC, 
February, www.brookings.edu/~/media/Files/rc/reports/2006/02childrenfamilies_ 
holt/20060209_Holt.pdf

the Food Stamp Program”, Political Science Quarterly, 114:3, Fall.

Temporary Assistance to Needy Families”, Social Science Quarterly 88:1, March.

Ku, L. and M. Broadus (2003), “Why Are States’ Medicaid Expenditures Rising?”, 
Centre on Budget and Policy Priorities, Washington, DC, www.cbpp.org/1-13-
03health.pdf.

State Welfare Innovation under Devolution”, Political Research Quarterly 53:2, June.

Analysis and Management 15:4, Autumn.

New York.

Murray, K. and W. Primus (2005), “Recent Data Trends Show Welfare Reform to Be a 
Mixed Success: Significant Policy Changes Should Accompany Reauthorization,” 


Summer.

Case of Welfare Spending”, Political Research Quarterly 58:3, September.

WECNF813.HTM.


Weaver, K. (2000), Ending Welfare as We Know It, Brookings Institution Press, Washington, DC.

Part IV

Case studies in product-market reform
Chapter 17

Australia: Power-sector reform, 1990-2004

Summary and main conclusions

Australia was among the first countries to undertake substantial market reform in the electricity sector. Reform involved breaking up the vertically integrated monopoly of production, transmission and distribution, introducing competition into those activities where it was feasible, such as generation and retail supply, and reforming the regulation of transmission and distribution, which were natural monopolies, so as to encourage greater efficiency. Limited wholesale market trading between New South Wales (NSW) and Victoria began as early as 1994, and a National Electricity Market (NEM) commenced operation in late 1998 as the wholesale market for the supply of electricity to retailers and end-users in southern and eastern Australia. An open access regime governs use of the transmission and distribution networks. Retail market opening, or full retail contestability (FRC), as it is called in Australia, was introduced in New South Wales (NSW), Victoria, South Australia (SA) and the Australian Capital Territory (ACT) in 2002-03, with Tasmania and Queensland following somewhat later. FRC gives all electricity customers the right to choose their retail suppliers, but customers can still opt not to enter a contestable market and to remain under regulated electricity tariffs. Most assessments credit the reform with increasing efficiency and reducing electricity prices.

A number of factors contributed to Australia’s ability to make substantial progress in the field of power-sector reform during the 1990s:

- **The momentum generated by a broad structural reform agenda.** Economic policy in the early-to-mid 1990s was driven by recognition of the need for microeconomic reform. The opening of the tradables sector in the 1980s had revealed the need for reform of infrastructure and non-tradable sectors, in particular, as it was understood that the inefficiency of many of these sectors was a competitive handicap to the exposed tradable sectors.

- **The creation of new federal co-ordinating bodies.** The first electricity reforms were undertaken by individual states. These first movers then sought, together with the Commonwealth government, to promote the creation of a national market. This process, in turn, was facilitated by the creation in the early 1990s of new institutions for co-ordinated federal-state policy making, most notably the Council of Australian Governments (CoAG). Power-sector reform became a core element of the “National Competition Policy” (NCP) reforms launched with the conclusion of a number of CoAG agreements in 1995. Under the NCP, the Commonwealth made payments to the states in return for fulfilment of their NCP reform commitments. The incentive effects of these payments proved significant, despite the relatively small sums involved. Commonwealth leadership proved
critical to this effort: although policy towards the power sector was mainly in the hands of state governments, national-level reform has tended to move forward when the Commonwealth government was actively engaged.

- **Fiscal pressures on the states.** Victoria, the state that moved first and furthest on electricity reform, did so largely on account of a severe fiscal crisis, and fiscal pressures helped drive reform in other states as well. In some cases, however, excessive concern with fiscal outcomes led to flaws in reform design.

- **Assimilation of new technologies and new ideas.** While successive studies highlighting the inefficiencies of Australian power utilities provided strong evidence of the need for reform, the market-oriented direction of the reforms reflected developments elsewhere. Changes in technology meant that economies of scale in power generation were falling, opening up the prospect of decentralised production, involving real competition, across a common delivery network, and there was growing attention internationally to the possibility of replacing vertically integrated monopolies with market-based models in electricity. Institutions like the Productivity Commission were effective in both assessing the need for power-sector reform and educating opinion-forming elites about new ways of managing the electricity-supply industry.

- **Scope for adaptation and adjustment.** There was no “big bang” in electricity: reform was rolled out over a number of years, at different speeds in different states. The benefits of reform in the states that moved first therefore tended to strengthen the case for reform elsewhere. At the same time, the gradual nature of the process meant that rules and institutions could be, and were, adjusted in response to problems that arose as the reform unfolded.

### The context for reform

**Framework conditions for reform**

Electricity reform in Australia posed a number of challenges. First, it was a lengthy and technically complex process: power-sector restructuring was not a “big bang” reform but a sustained effort involving the gradual implementation of a number of steps over nearly a decade. Moreover, the peculiar properties of electricity as a commodity (Box 17.1) meant that the design of a workable market would have to address a host of complex regulatory and institutional issues. Secondly, the powers of the Commonwealth in respect of electricity reform were indirect: it was essentially a state matter. Taken together, these two factors meant that the reform required co-operation between the Commonwealth and the states over an extended period, involving governments of changing political outlooks at both federal and state levels. Maintaining a commitment to reform in such circumstances was all the more difficult, given the risks involved: since Australia was near the international “frontier” of power-sector reform, there was only limited international experience on which to draw and much uncertainty about how things would unfold.

Two developments in the early 1990s helped create conditions that were perhaps uniquely favourable for a sustained commitment to reform among national and sub-national governments of both left and right over a long period. First, economic policy in the early-to-mid 1990s was largely concerned with questions of microeconomic reform. A combination of financial liberalisation and the opening of tradables sectors in the 1980s had revealed the need to reform infrastructure and non-tradable sectors, in
particular, since the inefficiency of many of these sectors was seen as a competitive handicap to the exposed tradable sectors that relied on them for services and other inputs. These concerns were underscored by assessments showing that the productivity of labour and capital in Australia’s utilities sectors was about half the OECD average. Secondly, the Commonwealth and the states were engaged in a broad effort to create closer, more efficient cooperation between different levels of government, not least in order to overcome the impediments to mobility, portability and uniformity that had, in the government’s view, rendered the Australian economy more fragmented than Europe after the advent of the single market. This process led, inter alia, to the creation of the Council of Australian Governments (CoAG), which has served since 1992 as the peak intergovernmental forum in Australia. While the CoAG process has not always proceeded smoothly, electricity reform proved one of the areas in which successive CoAG meetings proved most fruitful.

Box 17.1. Economic implications of the physical properties of electricity

Many of the technical challenges for policymakers trying to create electricity markets derive from three peculiar physical properties of electricity: it cannot be stored economically, apart from water stored in reservoirs available to hydroelectric generators; demand tends to be inelastic in the very short term, as most consumers lack the information or the means to react rapidly to changes in price; and supply and demand must balance at all times. Electrical imbalances at any point within an interconnected transmission network can have immediate and severe repercussions throughout the entire network. In addition, policy makers must consider the binding nature of generator and network capacity constraints when they emerge. The challenge of balancing supply and demand in real time across the whole network is magnified by the dynamic nature of flows on interconnected transmission systems, which follow the path of least resistance determined by the constantly changing interaction between generation and load, rather than contract paths. As a result of these factors, electricity markets can be quite volatile and may be particularly susceptible to the exercise of market power (IEA, 2001b; Hunt, 2002). For example, a generator who knows he must be dispatched to keep the system in balance is well placed to abuse his position. Electricity’s attributes also give rise to the need for centralised or coordinated system operation and dispatch to manage reliability. It is these considerations that underlie the need to create a number of special institutions to regulate and operate electricity markets, as well as to prepare fairly complex rules governing electricity trading.*

* For further discussion on the unique characteristics of electricity and their implications for market performance see Kellow (1996); OECD (2003); and Joskow (2003).

These two policy trends – the emphasis on microeconomic reform and the effort to improve inter-governmental co-operation – came together with the launch of the “National Competition Policy” (NCP) reform process launched by CoAG in 1995, of which power-sector reform became a core element. A combination of binding intergovernmental agreements and incentive payments offered by the Commonwealth to the states for implementation of NCP commitments drove competition-oriented reform in many domains, one of the most important of which was electricity (Box 17.2). The payments reflected the expectation that the structural changes wrought by competition reforms would lead to a shift in tax revenues from the states to the Commonwealth. NCP payments were conceived as a way for the federal government to share the fiscal benefits of reform with sub-national governments. Some AUD 5.7 billion in Commonwealth funds was earmarked for this purpose over the period to 2005-06.*
Box 17.2. The National Competition Policy (NCP)

In April 1995, CoAG, acting on the recommendations of a 1993 National Competition Policy Review Committee report (the so-called “Hilmer report”), agreed to implement a package of measures designed to extend pro-competitive policies to sectors of the economy that had previously been exempt from them. The competitive conduct rules of the Trade Practices Act were extended to these enterprises and sectors, including state and territory government business enterprises (GBEs), and a new national regime for access to infrastructure facilities was created. The Price Surveillance Act was amended to encompass oversight of state- and territory-owned GBEs. The reform package also included three intergovernmental agreements on aspects of competition policy. Under the third of these, the Agreement to Implement the National Competition Policy and Related Reforms, the Commonwealth agreed to provide payments to states and territories meeting their obligations under these agreements plus reform commitments in electricity, gas, water and road transport. Two new institutions were established to oversee and advise on the implementation of the NCP: the Australian Competition and Consumer Commission and the National Competition Council (NCC). The latter was created to coordinate reform efforts, to assess whether jurisdictions had met their reform obligations under the agreements and to make recommendations concerning the award of NCP payments to the states.


The major external pressures for reform stemmed from globalisation and technical change. The former fed growing concern about Australian competitiveness, while the latter raised questions about traditional ways of managing electricity supply. The spread of small, low-cost gas turbines in the 1980s and the easing of restrictions on the use of gas for electricity production made it more likely that real competition in generation could be realised: the economies of scale in power generation were falling to a fraction of what they had been, raising the prospect of decentralised production across a common delivery network. The economies of scope that had long justified vertical integration of electricity utilities were thereby called into question, opening up the possibility of replacing vertically integrated monopolies with market-based models in electricity.

While economic growth was fairly robust during the period when power-sector reform was formulated and implemented, a short but sharp recession at the beginning of the 1990s played a key role in getting the process started. First, the slowdown reinforced the pressure for efficiency-enhancing reforms, as firms looked for ways to reduce input costs and remain competitive. Secondly, it pushed mounting fiscal pressures in the state of Victoria to crisis point, triggering a wave of privatisations and reforms, among the most ambitious of which concerned power-sector restructuring. Fiscal pressures played an important role in the reform processes of other states, as well, though with varying results.

Actors and interests

Although it loomed large in a number of state elections, power-sector reform was not really a party-political issue. While the degree of their engagement varied, Commonwealth governments of both left and right remained fairly consistent in their overall commitment to market-oriented reform during 1990-2007. The Commonwealth’s position was reinforced by the work of national-level bodies like the Industry Commission and its successor, the Productivity Commission, on the potential benefits of reform. The states’ enthusiasm for reform was far more variable, but there were
supporters and opponents of reform in both the Australian Labor Party (ALP) and the
centre-right Liberal and National parties. Opposition within the ALP was motivated
largely by the concerns of organised labour about potential job losses, while among the
coalition parties the strongest opposition to reform stemmed from the fear that reform
would lead to much higher electricity prices for small businesses and rural consumers,
particularly in Queensland and NSW. This was an especially sensitive issue for the
primarily rural National Party (NP), but elements of the Labor and Liberal parties also
remained attached to the use of state-controlled utilities as instruments of regional policy.

Australian business was largely supportive of reform. Since the pre-reform
arrangements involved a good deal of cross-subsidy in favour of households, business
consumers had reason to believe they would face lower prices in a competitive market. Cross-subsidies to households and rural customers meant that the average tariffs charged
to non-farm commercial consumers in Victoria were 28% higher than otherwise. Business consumers were thus the most active and best organised supporters of reform, seeing it as part of a broad menu of options for reducing the costs of doing business in an increasingly competitive international environment. The Business Council of Australia (BCA), the country’s principal employers’ association, lobbied actively through its Electricity Task Force. In Victoria, the state that moved first and furthest on electricity reform, there was also business sponsorship for “Project Victoria”, a wide-ranging programme of market reforms promoted by business groups and economically liberal think-tanks in the early 1990s. However, business support for reform was not uniform. The very largest consumers, including some aluminium and chemicals producers, resisted the idea that they should buy electricity in the market pool rather than striking their own bilateral deals. Prior to the 1990s, such consumers had often played the states off against one another when making location decisions and had thereby negotiated very attractive long-term electricity-supply contracts that they were determined to keep. The real burden of cross-subsidies thus fell not on the biggest players but on “second-tier” industrial and commercial consumers. It was they who had the most to gain from reform.

Households and small businesses made up a large but rather diffuse constituency of
high-fixed-cost, low-volume customers, who were poorly organised and for whom electricity accounted for a relatively small share of total expenditure. Households, in particular, benefited from cross-subsidies. Nevertheless, while having little power in the market, small consumers exercised it at the ballot box. Consequently, the retail market proved the most politically sensitive segment of the market and the one where states proved least willing to deregulate or even to agree on a common regulatory framework. Voters were generally suspicious of privatisation, which proved unpopular in all the states that attempted it, but they did not generally react much to other reforms, in part because the states tended to protect households from any adverse effects of reform via price caps and other measures. Such protections notwithstanding, consumer groups argued consistently that reform would help big business at the expense of households and small firms.

Labour unions in the power sector tended to oppose both vertical separation and privatisation, chiefly owing to fear of job losses, a fear reinforced by the arguments of some privatisation advocates concerning the link between public ownership and union-backed working practices that undermined competitiveness. This was largely a question of protecting existing jobs in existing power plants: in general, the unions were not opposed to private participation in the construction of new power stations. Unions in Australia were growing steadily weaker in the 1990s, as a result of structural change in the economy and industrial relations reforms, and the electricity sector was no exception.
Power-sector employment had already been cut drastically in the last years prior to the reform. In general, therefore, the unions needed allies to prevail: they thus had more success in opposing privatisation, where the public and some power-sector executives supported the unions’ position, than in their resistance to vertical separation.

Finally, the reform was shaped in its later stages by new actors, who were created or entered the market as a result of the reform process, including independent power producers and energy markers, as well as a number of new regulatory bodies.

**Issue history**

Australia’s states initially developed separate electricity industries, consisting largely of vertically integrated public monopolies owned by state governments. These utilities were run in a highly politicised fashion, pursuing regional development and social goals, often with little regard for efficiency. The pursuit of such objectives reflected the electoral appeal of electricity investment as an instrument of regional development, as well as the interests of the utilities themselves, which pursued aggressive system expansion based on very optimistic demand forecasting, while using arguments about the specialised nature of the industry to rebuff criticism and avoid public accountability and scrutiny. Moreover, the states continued to develop their power sectors largely in isolation from one another, even after technology advanced to a stage that would make integration both feasible and efficient. This “electricity autarky” reflected both the path-dependence of state policies – each system had its own technical standards, voltage systems, governance philosophies, etc. – and the desire to give priority to state policy objectives, such as local employment creation or the use of state resources. The result of such policies was the emergence of structures characterised by excess generating capacity, over-staffing, inflexible pricing and lack of accountability. In due course, the existence of substantial overcapacity in a number of states actually made reform easier, because it greatly reduced the risk that the transition to competitive markets would lead to rapid price increases.

During the 1980s, rising concern about the inefficiencies involved in these arrangements led to a wave of managerial reforms in state utilities, often adopted by state governments in the face of opposition from incumbent management. While there was variation among the states, the common trend was towards greater commercial efficiency and greater accountability. These reforms resulted in appreciable efficiency gains: falling staff numbers and operating costs, reductions in reserve plant margins, increased sales and lower prices. Nevertheless, reviews of industry performance by the forerunners of the Productivity Commission – the Industry Assistance Commission (1989) and its successor, the Industry Commission (1991) – found that Australia’s power sector still performed relatively poorly in international comparisons and pointed to need for further reform. The latter report carefully documented the inefficiencies and distortions that characterised the status quo and explicitly advocated vertical separation, competition and the minimisation of political interference in operational matters. The reports’ methods and findings were both hotly contested, not least by the industry itself, but their estimates of the potential benefits of reform aroused the interest of policy makers, as well as the BCA and other business groups: the commission estimated that reform could increase national output by 0.7% annually, raising annual disposable income by around AUD 300 per household.
CHAPTER 17. AUSTRALIA: POWER-SECTOR REFORM, 1990-2004 – 381


Policy process: formulation and adoption

The search for solutions

The shift from the “managerial” reforms pursued by the states in the 1980s to the more market-oriented structural reforms of the 1990s reflected the coming together of a number of policy agendas and currents of thought:

• The states – particularly Victoria – were increasingly aware of the costs of their autarkic policies in the electricity sector. It was largely financial concerns that prompted the Special Premiers’ Conference of October 1990 (a forerunner of CoAG) to appoint a working group to explore the development of interstate electricity trade. That group’s report in July 1991 led to the creation of the National Grid Management Council to take the idea forward.

• There was growing interest among academics and policy makers in new approaches to the reform of infrastructure sectors being developed in other countries, particularly the United Kingdom, New Zealand and the United States. Thus, when electricity reform emerged on the national policy agenda, there was already a set of ideas which had currency in the Treasury, the Industry Commission and other Commonwealth institutions and which could be used as the basis for reform design.

• These ideas, in turn, meshed well with the government’s emerging microeconomic reform agenda, which reflected the Hilmer report’s emphasis on competitive neutrality, structural reform of public monopolies, removal of legislative restrictions on competition and the establishment of non-discriminatory access regimes for public infrastructure. The Hilmer report helped provide a basis for approaching electricity reform as part of a much broader microeconomic reform agenda.

This thinking was initially concentrated within the public administration, and the main impulse for reform came in the first instance from officials and politicians concerned with Australian competitiveness. While business groups came to support reform, this was largely a response to the work of bodies like the Industry Commission (later renamed the Productivity Commission) to expose and, where possible, quantify the costs of inefficient regulatory and other policies, particularly in sectors dominated by GBEs, and to publicise the potential benefits of change.

The CoAG reform principles

Electricity reform was a multi-faceted process, involving an array of decisions and innovations at state and national level, as well as a number of intergovernmental agreements. However, the closest thing to a single overall blueprint for reform was the set of principles for the NEM confirmed at the CoAG’s 1994 Darwin meeting. The most important principles included:

• structural separation of generation, transmission, distribution and retail supply activities;

• customer choice of supplier, including generators, retailers and traders;

• non-discriminatory access to an interconnected transmission and distribution network;
• the elimination of discriminatory legislation or other barriers to entry for new participants in generation or supply; and

• the removal of discriminatory legislative or regulatory barriers to inter- or intra-state trade.30

The governments also agreed on basic principles relating to the recovery of the fixed-cost component of network pricing, which was critical to making a reality of non-discriminatory access commitments.31 Each jurisdiction was to introduce competition into its own electricity industry, including structural separation, before the launch of a National Electricity Market (NEM) comprising Queensland, NSW, the ACT, Victoria, SA and Tasmania. Western Australia (WA) and the Northern Territory were not to join the NEM, owing to the long distances between their load centres and the interconnected network in southern and eastern Australia. Government-owned electricity utilities were to be corporatised, but privatisation was not among the agreed principles, as only Victoria was prepared to privatise at that point. The aim of the reform was to create fully competitive electricity markets by mid-2001. The CoAG package pre-dated the NCP, but it reflected the principles of the Hilmer report, which were to form the basis for the NCP, and it thus proved relatively easy to integrate power-sector reform into the NCP framework, making NCP payments available to reforming states.

Policy process: implementation

Reform in Victoria

Although the issue had been under discussion for several years, it was a fiscal crisis in Victoria that triggered the first wave of really radical reform. In 1992, the incoming Liberal-National government inherited an AUD 2.2 billion deficit on a state budget of around AUD 13 billion, a net public sector debt of 33 billion.32 The crisis was in fact partly a result of policies towards the power sector: in the late 1980s, the previous government had opted to proceed with construction of the AUD 1.7 billion Loy Yang B power station rather than accept an offer from NSW to complete and commission the much cheaper Mount Piper power station – which NSW was in the process of moth-balling – and to sell electricity to Victoria at prices lower than Loy Yang B was expected to require. The State Electricity Commission of Victoria (SECV) later recognised that it could not afford to finish Loy Yang B, but by then enormous debts had been incurred on account of the project.33 In order to close the deficit and reduce the debt burden, the government undertook a wave of reforms, including the restructuring and privatisation of the power sector in the face of fierce opposition. The reforms encompassed the vertical separation of the SECV in 1993, the creation of a wholesale market (Vicpool) in 1994, and the privatisation of the state’s generators in 1996-97. This was followed by the gradual opening of the retail market to competition.

Critics accused the government of exaggerating the depth of the fiscal crisis in order to justify its pursuit of the radical reforms outlined in “Project Victoria.”34 In fact, while the state’s loss of its AAA credit rating left little doubt about the gravity of the fiscal situation, the government’s opponents were correct to claim that Victoria’s Liberal leadership was committed to market reforms anyway. The significance of the crisis was that it enabled the government to overcome opposition to unpopular measures like electricity privatisation.35 The Victorian Labor Party was in disarray following its defeat and was in any case poorly positioned to oppose the reform, since it had lost the 1992 elections on account of the fiscal crisis, which was attributed in part to its management of...
the power sector. Moreover, it was Labor that had undertaken the first, albeit partial, privatisation of a power station at the end of its term of office. At the same time, a labour relations reform then being introduced, combined with the shift from trade- to industry-based unions, was putting many trade unions under financial stress and sharpening the competition among them for members and collective bargaining rights. The labour unions were therefore unable to unite in opposition to the reform. Finally, the government replaced senior SECV officials who opposed vertical separation and privatisation with managers committed to the reform.

Although it moved quickly to exploit the window of opportunity created by the fiscal crisis, the Victorian government did not rush implementation of the reform. A “paper market” was set up to simulate the wholesale market and test its rules for a year prior to the launch of Vicpool, while the reformers studied the experiences of the United Kingdom, the United States and New Zealand closely. The decision to split the generating and retail assets of the former SECV into a number of different companies, for example, was prompted by the belief that UK governments had paid too much attention to sale prices and too little to market structure in some major privatisations. A major public relations effort was also mounted, in an attempt to win public support for the reform. While the polls suggested that the voters were never convinced of the benefits of privatisation, the government’s re-election in 1996, when the reform was being hotly debated, did allow it to claim a mandate to carry on.

Concessions to key stakeholders and interest groups also helped defuse opposition. Retail tariffs were frozen until 1996, and existing long-term supply contracts with very large industrial consumers were upheld. The NP threatened to leave the coalition if the government did not preserve uniform pricing of electricity for urban and rural consumers, despite the substantially higher cost of supplying the latter; in the end, the government rejected such “postage-stamp” tariffs in return for significant increases in spending on rural infrastructure, and the coalition held. This decision ultimately led to a doubling of power prices in some rural areas and contributed to the government’s electoral defeat in 1999, as did job losses in the small communities of the Latrobe Valley, where Victoria’s generation was concentrated: employment in the sector fell by almost two-thirds during the 1990s.

The diffusion of reform

After 1992, the creation of a national market was largely driven by a cross-party alliance involving the governments of Victoria and NSW and the ALP federal government. In 1995-96, NSW followed Victoria’s lead down the path of vertical separation, corporatisation and the introduction of a wholesale market, though it did not privatise. In May 1997, the two states formed a single wholesale market, to which Queensland adhered six months later. This alliance between the Commonwealth and the two largest states facilitated the adoption of a number of important CoAG decisions on electricity issues, including: the formation of a National Grid Management Council to manage the transmission network (transmission assets remained in the hands of the states); the launch of an NEM “paper market”; agreement on a National Electricity Code, laying out rules for NEM participants, based largely on the Victorian and NSW codes; creation of a National Electricity Code Administrator to enforce those rules; and the formation in 1996 of the National Electricity Market Management Company (NEMMCO) to serve as both market operator and system operator for the NEM. The NEM itself began operating in December 1998.
This broadly successful drive to create the framework for a single wholesale market stood in marked contrast to the much more variable experiences of the states in implementing reforms within their own jurisdictions. Vertical separation proceeded relatively smoothly in NEM-participating jurisdictions, but WA resisted it until 2005, when it adopted legislation on the restructuring of Western Power, partly in order to release suspended NCP payments. However, the degree of vertical separation in the NEM states varied, since vesting contracts and other instruments were used to sustain the links between vertically separated but still publicly owned generators and retailers. Critics spoke of “pseudo-vertical integration”, as state-owned retailers and generators in most states tended to contract overwhelmingly with each other.

Willingness to invest in interstate interconnectors varied significantly from state to state and over time. Even in Victoria, critics claimed that the government’s desire to maximise the sale price of its power-sector assets led it to neglect the potential for developing the inter-state grid, despite the fact that the fiscal crisis in Victoria had resulted partly from its predecessors’ reluctance to buy cheap electricity from NSW. Anxious to maximise privatisation revenues, SA was reluctant to strengthen interstate connections prior to disposing of its electricity assets, and in Tasmania, which relied overwhelmingly on hydroelectricity, proposals to connect that state to the mainland were resisted not only by environmentalists but also by those who feared that an integrated market would force Tasmanians to pay higher “continental” prices. Opponents argued that construction of the Basslink interconnector would lead to the privatisation of Hydro Tasmania. There was also resistance to strengthening the interconnectors between Queensland and NSW. In each case, opposition was motivated by the fear that market opening would lead to a loss of rents.

Given the sensitivity of the issue, each state tended to adopt its own approach to retail market opening. All jurisdictions retained some sort of price caps in the retail sector, and some, such as Queensland, were extremely resistant to full retail contestability (FRC) and introduced it relatively late. Queensland retained “postage stamp pricing”, the elimination of which would have been very unpopular with its rural electorate. NSW opened up the retail sector but created a so-called “Electricity Tariff Equalisation Fund” (ETEF): retailers supplying customers at regulated rates had to contribute to the fund when spot market prices fell below the strike price set by the NSW Treasury and were compensated by it when spot prices rose above it. For generators, the fund operated in reverse. It was, in essence, a hedging mechanism operated by the state treasury for the benefit of state-owned incumbents, and the NCC criticised it as a potential barrier to entry.

NCP payments were small relative to state budgets (less than 0.1% of GDP throughout the period), and they were rarely suspended or withheld: in the end, well over 95% of the funds available for NCP payments were paid to state governments during 1997-2006. However, they played an important role in reform implementation for several reasons. First, the publicity associated with the NCC’s annual assessments was a source of pressure on state officials to keep to their NCP commitments. Secondly, the sums involved, though small as a share of total state spending, were substantially larger as a share of states’ discretionary spending. Thirdly, the lack of penalties in the early years meant that the states came to rely on the funds when budgeting, which made deductions or suspensions more disruptive. Finally, the NCP payments empowered state treasuries and state departments of premier and cabinet, which were usually driving the reforms, vis-à-vis other departments. This was because delays and deductions were often
taken from the budgets of the “offending” state departments, something no state minister wished to accept.\(^{50}\)

Privatisation, however, stood outside the NCP process, as the states were unwilling to include it in the CoAG agreement, and, in the end, only Victoria opted for full-scale privatisation. In SA, only the retailer was privatised outright; the state’s generation, transmission and distribution companies were put under management by private companies with long-term leases. SA’s experience with privatisation was unfortunate and tended to reinforce opposition in other states (Box 17.3). Several attempts at privatisation under successive Labor governments in NSW were blocked, owing to strong opposition from the unions and Labor backbenchers.\(^{51}\) The NP also remained wary of privatisation in NSW, fearing that it would lead to higher prices for rural consumers. In 1998, a minority Liberal government in Tasmania lost an election over its plans to privatise Hydro Tasmania.\(^{52}\) Queensland privatised its retailers only.

**Box 17.3. Electricity privatisation in South Australia**

The experience of SA in the late 1990s highlights the pitfalls facing state governments that pursued inconsistent policies towards the sector. A Liberal government won re-election in 1997, having opposed privatisation of the Electricity Trust of South Australia (ETSA), only to announce plans to break ETSA up and sell its generation, transmission, distribution and retail businesses to distinct investors, though the state was to retain freehold rights to all but the retail assets. In justifying this *volte-face*, the government cited a combination of serious fiscal pressures and new information about the development of the NEM. In what critics argued was an attempt to maximise the sale prices, the government then withdrew its support from plans to build a new interconnector with NSW, citing “potential reliability concerns”. It thereby limited the competition that SA generators would face. The sales went ahead, to be followed during the unusually hot summer of 2000-01 by periodic blackouts. While the government argued that the problem was not privatisation-related, its opponents emphasised its earlier rejection of the Riverlink interconnector, which the government, reversing course yet again, embraced. It lost the 2002 state elections.

*Source:* See Parkin (2001); Quiggin (2001); Sharma (2003); and Cahill and Beder (2005).

The states’ failure to privatise more has been criticised in numerous assessments of the reform, in part because of doubts about whether or not state-owned generators and retailers are operating on commercial principles, but chiefly because state ownership is regarded a deterrent to new entrants. In a situation where the competitive neutrality regime largely boils down to an appeal to state governments, private investors have been reluctant to enter states where their up- and down-stream competitors are owned by those same governments. The tendency of some state governments to grant derogations to national rules to publicly owned incumbents reinforces such concerns.\(^{53}\) Initially, the presence of substantial over-capacity in many states meant that these concerns were not seen as urgent, and the Hilmer principles, which underlay the NCP, focused on market structure and competition rather than ownership. This was probably crucial to the acceptance of the reform, given how deeply entrenched opposition to privatisation was. However, by the early 2000s, when there was no longer so much over-capacity in the system and cash-strapped states faced the need either to attract private investment or to finance the renewal of their ageing base-load thermal power stations, privatisation moved back onto the agenda. It was largely the belief that state ownership was a deterrent to new
entrants that prompted unsuccessful attempts by two NSW governments to privatise that state’s electricity assets.

**Feedback and adjustment**

The launch of the NEM was followed by something of a hiatus in the reform process, while agents adapted to the new conditions and saw how the market worked. In any case, the coalition that had driven the reform in the early 1990s was gone by this time: power in Victoria and NSW passed to governments that were somewhat ambivalent about further reform, and the new federal coalition shifted its focus away from policy making within the CoAG framework. By the early 2000s, however, there was growing discussion of the unfinished business of power-sector reform. Successive reviews of the progress of reform concluded that it had yielded significant benefits but also identified and attempted to address a number of deficiencies, particularly with respect to the clarity of governance arrangements, the number of regulators (often with overlapping responsibilities), and continued perceived or actual conflicts of interest where governments remained asset owners. The Commonwealth Treasury estimated at one point that there were no fewer than 17 regulators involved in various segments of the market. This partly reflected the fact that regulation of transmission and the wholesale market was assigned to the Commonwealth, while distribution and retail supply were left to state regulators. Moreover, the states were increasingly dissatisfied, and at times in conflict, with national-level regulatory bodies over which they believed they had too little control. There was also concern about the poor incentives for transmission investment, inflexible price signals for retail customers and market features that impeded the use of long-term contracts.

The most important of these review exercises was the 2002 CoAG Energy Market Review, which concluded that these problems contributed to a de facto regionalisation of the NEM that prevented full realisation of many of its potential benefits. The review panel therefore proposed, and CoAG approved, a major overhaul of regulatory governance in the sector. A single Australian Energy Regulator (AER) was established under the aegis of the Australian Competition and Consumer Commission to replace some 13 different but overlapping regulators, with responsibility for economic regulation of wholesale markets and transmission networks. At the same time, in an effort to create a more structured and transparent process for rule-making and market development – and to separate rule-making from rule-administration – CoAG opted to create a new Australian Energy Market Commission (AEMC), replacing three previously existing bodies. Members of both new bodies were to be chosen largely by the participating states and territories. The Ministerial Council on Energy, itself a CoAG body, was left as the sole ministerial-level decision-making body. These arrangements helped satisfy the states’ desire to gain greater influence over national-level regulators in return for accepting the transfer of more regulatory authority from state to national level.

These measures helped to improve the overall framework for the NEM, but they stopped short of addressing retail markets, which remained politically sensitive. FRC was being introduced in the early 2000s, and the states, all of which had retained some sort of price caps on retail tariffs while FRC was being rolled out, refused to surrender control over retail markets to the new AER. Critics insisted that price caps would limit competition and impose a disproportionate share of the costs of rising fuel prices on generators, but CoAG agreed only that the AEMC would “review” retail competition in every jurisdiction before reporting and making (non-binding) recommendations.
concerning price caps. This left the states and territories connected to the NEM with one wholesale market but six retail markets. However, the agreement was not without impact: an AEMC review did lead to the lifting of price caps in Victoria from the beginning of 2009, and the Commonwealth and the states began working to establish a National Energy Customer Framework, which would eventually transfer state regulatory functions in respect of retailers to the AER, with the exception of price regulation.

Yet not all subsequent policy change occurred as a result of early difficulties with the reform: its early successes also contributed to changing policies. Thus, as confidence in the NEM grew, opposition to interconnectors within the states largely abated. Indeed, the blackouts in SA (Box 17.3) in 2000-01 prompted a policy reversal there, and the construction of the Basslink interconnector joining Tasmania to the mainland has proved critical to that state, as the current prolonged drought has prompted that hydro-dependent state to import a great deal of power from its northern neighbour. This change in attitudes made it easier for the states to agree on a new, cost/benefit-based regulatory test for network investment in 2003. This is an important step, since the test not only considers emerging transmission constraints but also considers whether or not local generation or demand-side solutions might be more efficient than new power lines. The transparency of this process is important, given the scope for lobbying and conflict over such issues, since the value of generating assets can be hugely affected by the grid investments that are (or are not) undertaken. Similarly, while resistance to privatisation remains strong in some quarters, it appears less entrenched than it was, largely because several of the state governments see increasing need for new investment in power generation and do not wish to finance it themselves. Even the unions increasingly accept that new stations are likely to be privately built, but attracting private investment may yet compel some states to privatise at least partially, in order to allay investor concerns about competitive neutrality.

**Conclusion**

Most assessments suggest that there is more still to do in order to overcome the geographical segmentation of the market, which becomes highly regionalised when physical transmission constraints intervene, leading to divergence of pool prices among regions. The planning and management of transmission networks remains regionalised. Nevertheless, the benefits of reform have been substantial. Market opening was followed by falling electricity prices for commercial consumers and households in all NEM-participating states except SA, and TFP growth in the power sector accelerated after the onset of structural reform. The Australian Bureau of Agricultural and Resource Economics (ABARE) estimated that the benefits of electricity reform in the first three years following the launch of the NEM amounted to AUD 1.5 billion at 2001 prices, equivalent to around 0.25% of 2001 GDP.

As a case study, the Australian power sector is interesting in a number of respects. First, it was a politically and technically complex process that required both high-level political commitment and the creation of new institutions and policy processes to sustain it. Indeed, many of the public institutions and private actors which came into being as a result of the reform subsequently helped to drive it forward. Secondly, it was the larger states, rather than the Commonwealth, that really took the lead in the beginning – particularly Victoria but also, to some extent, NSW. This created a situation in which the first movers demonstrated some of the benefits of reform, allowing other states to opt in later on. However, as the reform unfolded, the Commonwealth increasingly assumed
the leading role, which was probably necessary if it was to be coordinated across the
states. Thirdly, the decision to roll power-sector reform into the NCP process was critical:
the NCP payments, though modest, helped empower reformers in state governments,
while the breadth of the NCP reforms meant that power-sector restructuring was merely
one part of a much larger policy shift. Since the electricity reform was getting under way
even before the NCP was launched, it might be viewed as a sort of ex post bundling of
reforms.

The fact that the NCP was concerned with competition and market structure rather
than ownership enabled the authorities in many states to finesse the privatisation issue.
While this omission was not without its costs, it was important in allowing the reform to
proceed – if privatisation had been a non-negotiable element of the reform, it might never
have been undertaken. Moreover, the operation of the market since 1998 has itself
generated increasing pressure on the states to create a more investor-friendly
environment, in part by privatising some of their power-sector assets. Recent assessments
also suggest that prices have risen less and productivity has grown faster in south-eastern
Australia, where the sector is largely in private hands, than in NSW, where the state
retains ownership of all the major assets. It is also in New South Wales that the signs that
electric companies are wielding market power are most pronounced.65

For all the difficulties involved, this was a reform for which many of the
preconditions were in place: employment, prices and excess capacity in the sector were
all falling before 1990, and management culture was changing, while technical
developments and policy innovation abroad were changing the range of options available
to policy makers when it came to managing electricity. To a great extent, the 1980s may
be seen as a period in which the old regime grew ripe for more fundamental reform. At
the same time, there was increasing pressure for efficiency-enhancing reforms in a range
of infrastructure sectors, a pressure reinforced by the Industry Commission’s work in
assessing the costs of the status quo and the potential benefits of reform. Of course,
conjunctural factors also played a role: even Victorian policy makers later doubted
whether or not they could have privatised the sector in the absence of a fiscal crisis and an
opposition that was itself still in disarray following its defeat. Finally, it is important to
recognise that the reform was gradual enough to allow flaws in the process to be rectified:
mistakes were made, some of them quite costly. In a number of instances, they were
corrected later on; in others, they remain part of the unfinished business of the reform.
The critical point is that they were not fatal to the reform, which, for all its imperfections
and incomplete elements, still delivered many of the benefits the reformers sought.

Notes

1. Specifically, Queensland, New South Wales (NSW), the Australian Capital Territory
(ACT), Victoria, South Australia (SA) and, since May 2005, Tasmania.

2. In Queensland, retail contestability (RC) was granted from 1 July 2004 to connection
points consuming more than 100 MWh per year, with FRC following on 1 July 2007.
Tasmania, too, has begun phasing in limited RC.

4. OECD (1994); on the impact of such findings on the Australian debate, see Cox and Warner (2007).


6. The highest-level CoAG meetings bring together the prime minister, the state premiers, the chief ministers of the territories and the president of the Australian Local Government Association. Ministerial committees bring together the relevant ministers and officials for more specialised discussions.

7. Once restructured into corporations, state-owned GBEs, which paid dividends to state budgets, would pay taxes to the Commonwealth.

8. This funding level was developed with reference to the Industry Commission’s initial estimates of the benefits of implementing the reform program.


10. For contemporary reportage concerning the major political parties’ positions, see The Age, 9 January 1993, 29 August 1993, 4 September 1994, and 10 January 1995.

11. White (1996) shows that reform in the United States proceeded fastest in regions where the price gap was greatest.


13. In this, Australia was typical of most economies undergoing electricity reform; see Joskow (1996). White (1996) shows that reform in the United States proceeded fastest in regions where the gap between regulated prices of incumbent utilities and those expected to prevail in a deregulated market equilibrium was greatest.


15. Project Victoria was an initiative of the Melbourne-based Institute of Public Affairs and the Tasman Institute, together with 13 employer groups, including the Australian Chamber of Manufacturers, the BCA, the Victorian Chamber of Commerce and Industry, the Victorian Employers Federation and the Victorian Farmers Federation. See Cahill and Beder (2005).


17. See the surveys cited in Cahill and Beder (2005); Parkin (2001); Sunday Herald Sun, 5 February 1995.


20. Quiggin (2001) observes that reliance on vertically integrated monopolies and de facto electricity autarky were probably economically rational, given the geography of Australia and the technology of the time.

21. Sharma (2003). Kellow (1996) argues that this institutional configuration promoted reverse adaptation – the ends of policy were adapted to fit the means available rather than the other way around.

22. Sharma (2003). For decades, the only major exception to this rule was the Snowy Mountains hydroelectricity and irrigation complex, which involved NSW, the ACT and Victoria and was jointly owned by the Commonwealth and the governments of Victoria and NSW.
23. Victoria, for example, was keen to maintain a market for its brown coal.

24. These reforms included: independent inquiries to review performance, the appointment of independent directors to utility boards, closer scrutiny of investment plans, more rigorous accounting practices and the commercialisation of operations. In some places, the first steps towards vertical separation were also taken.


27. See the overviews of the industry position provided in Sharma (2002, 2003); also ESAA (1996).


29. In Victoria, this understanding was largely a result of the catastrophically expensive fiasco surrounding construction of the Loy Yang B power station; see Barton and van Onselen (2004).


31. Including common asset valuation methodologies, rates of return and pricing methodologies.

32. The government invoked debt figures as high as AUD 70 billion, citing large unfunded liabilities bequeathed to its by its predecessors; see, for examples, the conflicting claims made in The Age, 13 February 1993; and 18 April 1993; The Herald Sun, 30 January 1993.


34. See, e.g., Barton and van Onselen (2004); Cahill and Beder (2005); The Age, 18 April 1993. Much of the dispute centred on the treatment of unfunded contingent liabilities referred to above.

35. The Age, 10 January 1995; and 25 February 1995. While some in the Liberal leadership had been associated with Project Victoria, the party’s election campaign had been cautious, focused on the incumbent Labor government’s failings rather than the Liberals’ plans. See The Age, 9 January 1993.

36. As noted above, Labor had rejected an offer of surplus electricity from NSW in favour of building the 1000MW Loy Yang B power station, which in the event proved prohibitively expensive. In 1992, it then sold a 49% stake in the still unfinished station, as this appeared the only way for the heavily indebted SECV to finance completion of the plant. See Barton and van Onselen (2004).

37. On the labour relations issue, see Barton and van Onselen (2004); and Gahan (2005). The government’s labour relations reform was ultimately thwarted, but not before the upheavals it triggered had provoked considerable conflict among the unions.

38. Conflicts between some of the unions and the leadership of the Labor Party in Victoria also made it harder for the opponents of reform to unite. See Economou (1999).


41. The Age, 4 September 1994.
42. The two roles are closely related. The market operator is responsible for the operation of the wholesale electricity spot market. Generators and buyers send bids to the market operator, who matches them. If the resulting basic daily schedule is not feasible due to transmission constraints, the market operator incorporates offers for congestion relief to establish the definitive feasible daily schedule. The system operator is responsible for generation dispatch, the provision of ancillary services, and operation and control to ensure safety, power quality, stability, reliability and security of the transmission system.

43. NCC (2005).

44. See Box 17.3 below for details.


47. Queensland did so only after the NCC recommended a suspension of 25% of its 2004 NCP payments.


49. The proportion paid for 1997-2004 was 98%, but there was somewhat more withholding in the last two years; see NCC (2005:xx-xxi) and Productivity Commission (2005:30-32).

50. Productivity Commission (2005:31) also draws attention to the extent and significance of non-compliance on the part of the Commonwealth itself, which did not have the incentive of NCP payments to encourage ministers to meet their reform commitments.


54. In NSW, the Carr and Iemma governments both tried to privatise power-sector assets but were thwarted by opposition from the unions and from within the NSW Labor Party.

55. Davis (1998); Stewart (1999); Galligan and Wright (2002); and Hollander and Patapan (2007).


57. IEA (2005).

58. The AER’s remit does not include Western Australia or the Northern Territory.

59. Specifically, the National Electricity Code Administrator (NECA), the National Gas Pipelines Advisory Committee and the Code Registrar.

60. Subnational governments appoint two of the three members of both the AEMC and the AER, including the chairman of the former. The chairman of the AER must be agreed by the Commonwealth and five of the states and territories.
61. This outcome is somewhat ironic. When Basslink was first proposed, it was expected that Tasmania would import cheap Victorian base-load power during slack periods, thereby conserving water in its reservoirs, and then export hydro back to the mainland in periods of high prices. Things have not turned out this way, but Basslink has nevertheless been crucial to meeting Tasmania’s energy needs, a fact that underscores the flexibility that more integrated markets provide in a changing environment.


63. OECD (2004); IEA (2005); Abbott (2006); ERIG (2007).


65. ERIG (2007).
Bibliography


Chapter 18

Australia: Water reform, 1994-2004

Summary and main conclusions

Australia undertook a major reform of water resource policies with the 1994 Water Reform Framework, which aimed to establish a market-based system by 2005. This was to be achieved by allowing prices to reflect the full cost of resources, eliminating cross-subsidies and making other subsidies transparent. Regulatory functions were to be separated from water provision, and trading in water rights was to be introduced to increase allocative efficiency. This ambitious agenda put Australia at the leading edge of water management internationally. However, progress in implementing water reform over the ensuing decade was slow and uneven, both across water-use categories and between states, and fell far short of the goals of the original reform programme. As a result, a new water reform initiative was launched in 2004, in an effort to reinvigorate the reform process. Water reform proved difficult to advance for several reasons:

- **A weak constituency for reform.** Because the status quo with respect to water consumption was reckoned to be unsustainable over the long term, water reform was largely about allocating losses, particularly among rural users, who tended to be cohesive and have political influence. While the economy as a whole, as well as the environment, would benefit in the long run from more sustainable consumption and more efficient allocation of water, most agents would be worse off under the reform than they had been previously, particularly those who were intensive water users and were thus more engaged in water-policy debates.

- **Difficulties in co-ordinating policies among the states.** Water reform highlighted the sharp conflicts of interest between up- and down-stream states in the Murray-Darling Basin (MDB), which accounts for around 40% of national income derived from agriculture and almost three-quarters of irrigated agriculture in Australia. As a result, the scope for using inter-governmental institutions to co-ordinate state policies was more limited than it was in respect of electricity reform, which offered positive benefits to the states.

- **High social costs of reform.** A shift away from water- and labour-intensive irrigated agriculture would not only reduce employment in the farm sector substantially, it would have serious knock-on effects for rural towns, particularly in the MDB. Many rural communities exist chiefly as a result of past irrigation policies that were adopted with a view to encouraging rural settlement. The agricultural lobby and other rural interests used their considerable political muscle to resist changes that threatened the future of irrigated farming.

- **Scientific uncertainty.** There was considerable uncertainty about what levels of extraction really were sustainable, particularly given uncertainties surrounding the impact of climate change, which was hotly debated. This made it harder to
produce authoritative estimates of the benefits of reform (or the costs of non-reform). Moreover, since judgements about technical and hydrological matters had potentially enormous distributional consequences, those for whom reform would be very costly had every incentive to challenge the scientific case for reform, making it hard to establish a consensus even on the most basic issues

- Technical complexity of implementation. Practical difficulties in measuring consumption and in assessing the impact of water diversion on other users made the implementation of reform principles exceptionally complex. The potential externalities involved in trading water, for example, are very hard to address, since the impact of a transaction on third parties may depend not only on the volumes traded but on where they are traded from and to, and on any change in the use of the water as it changes hands.

The context for reform

Framework conditions for reform

Although Australia is the driest inhabited continent in the world, its relatively small population means that water is abundant on a per capita basis, and it has the fourth-highest rate of water abstraction per capita in the OECD. Abstraction since 1980 has risen faster than in almost any other OECD country. The challenges Australia faces in trying to ensure that water supply is both efficient and sustainable stem chiefly from three factors. First, it has the highest year-to-year variability of rainfall in the OECD: many rivers have highly variable flows and protracted droughts are common. There is also enormous spatial and seasonal variation in rainfall. Secondly, data from the Australian Bureau of Meteorology show a long-term trend towards declining rainfall in eastern Australia and rising rainfall in the North-west. Changing rainfall patterns are particularly important in view of the third factor: the concentration of both population and irrigated agriculture in the east of the country. Roughly 70% of water use is for irrigated agriculture, which accounted for up to 95% of consumption growth in the mid-1990s. Almost three-quarters of Australia’s irrigated land depends on one river system – the Murray-Darling Basin (MDB). Covering over 1m square kilometres of eastern Australia, with a population of about 2 million, the MDB accounts for 50% of water consumption and 40% of national income derived from agriculture. Average rainfall in the MDB has fallen in recent decades, even as average temperatures have risen, increasing both evaporation and the water needs of plants. Diversions are thus strongly and inversely related to rainfall.

The framework for addressing these challenges is largely defined by Australia’s federal structure. Section 100 of the Constitution has been held to prevent the Commonwealth from implementing irrigation policy. Water is state property; entitlements to use it, usually tied to land ownership, were granted as statutory rights under state legislation. Thus, although there are national imperatives for water management, responsibility for water issues is vested in the eight state and territory governments. There has long been a consensus in Australia on the need for integrated catchment management (ICM), involving a holistic assessment of water quality and quantity issues, land, water and air interactions, and externalities between up- and down-stream users. Since catchments do not correspond to state boundaries, agreement between states is required to permit such an approach to water management, particularly in the MDB, which occupies parts of four states.
Economists in Australia first began discussing the question of water pricing in the 1960s and 1970s, but the issue did not make it onto the policy agenda until the 1980s. Little progress was made until the 1990s, when reform initiatives were undertaken within the context of a much broader set of microeconomic reforms, many of which concerned infrastructure sectors and utilities like water and power, within the framework of the National Competition Policy (NCP). The reform of such sectors was regarded as critical to maintaining the competitiveness of Australia’s tradable sectors, which depended on them for inputs and services. The potential gains from infrastructure reform were highlighted by international comparative assessments suggesting that labour and capital productivity in Australian utility sectors was only about half the OECD average. The reforms undertaken in many of these sectors shared a common approach, based on the introduction of rational pricing for scarce resources and the development of competitive markets in sectors previously organised as public monopolies. The prevalence of this approach meant that policy makers approached water reform in the early 1990s with a template in mind that was being applied to other sectors.

The other critical factor shaping water reform debates was, of course, climate change. Advocates of reform had drawn attention to both the efficiency of urban and rural water supply and the environmental impact of some water-use practices long before the 1990s, but these concerns were reinforced by evidence that climate change might well trigger long-term changes in rainfall patterns that would render many existing policies and practices even harder to sustain. The onset of a severe and protracted drought in the early 2000s gave substance to these fears. The climate change debate and the drought also raised public awareness of water policy issues.

**Actors and interests**

The range of actors involved in water policy in the 1990s was relatively limited. In the early 1990s, the complexity of the issues involved and the lack of immediate signs of crisis meant that the public was not much engaged in the reform debate, which was largely confined to academic experts, representatives of the affected interest groups – chiefly the irrigated agriculture sector – and Commonwealth, state and local officials involved in water policy. In respect of urban water supply, in particular, the major players were state governments and water utilities: there was little public support for water-pricing reform, but water was too small an item in most household budgets for it to be a major political issue for urban Australia. Politically, water has been much more sensitive in rural areas. The largely urban base of the Australian Labor Party (ALP) meant that water reform was less sensitive for it, and it was generally more closely identified with many of the ideas associated with ICM. By contrast, the predominantly rural-based National Party (NP) tended to be wary of reform and keen to protect farmers in the irrigated sector. Its traditional coalition partner, the Liberal Party (LP), also relied on substantial support from rural towns where the effects of water reform were potentially severe. The surge in support for the populist and largely rural “One Nation” party in the late 1990s, which was partly a reaction against the effects of the liberal economic reforms embraced by the major parties, reinforced the incentives for the NP and LP, who controlled the Commonwealth government at the time, to be sensitive to rural concerns.

The farming lobby’s principal concern was with the security and predictability of water entitlements, but there was also a strong commitment to keeping prices down. Since many activities would probably be rendered unsustainable if water prices were determined by the market, there was a fear that water reform could leave the sector with a potentially huge “stranded asset” problem. This problem was not limited to irrigation...
assets: if changes in water supply prompted a shift towards less intensive dry-land agriculture in a region and a corresponding decline in the farming population, then many rural communities would experience second-round effects, with schools, hospitals, banks and other services closing. The influence of the rural sector was being steadily eroded, as both the rural population share and the relative weight of agriculture in GDP declined, but the irrigation sector continued to lobby MPs effectively. Farmers also had considerable influence over the administration of water policy: the states typically devolved a great deal of authority over water to local water districts and catchment management authorities, which tended to be dominated by irrigators.

Prior to the 1990s, the states were the main drivers of water reform, as it was their treasuries which bore the costs of subsidising irrigated agriculture and urban consumers; many early experiments with pricing and water trading were first initiated by state governments. However, national-level reform posed more complex problems, since the costs and benefits of reform often had to be allocated among the states rather than within them. Differences among the states had long been among the major obstacles to agreement on the collaborative management of the MDB. To a great extent, this reflected the different ways in which the states planned irrigation development: Victoria and South Australia (SA) favoured horticulture, viticulture and dairy production, activities that could take years to recover from a water-supply disruption. They therefore managed allocation policies fairly conservatively, in order to assure a high degree of water security. By contrast, New South Wales (NSW) focused on annual crops like rice, production of which could resume following a disruption as soon as rainfall and inflows to storages improved. NSW therefore attached much less importance to security and adopted a less restrained approach to the expansion of allocations. This led to conflict between Victoria and NSW in the 1980s over the management of water held in joint storages.

During the 1990s, these differences in irrigation planning underlay differences in the nature of the tradable water rights created by the states. There were also conflicts among the states of a purely upstream-downstream nature: South Australia, which suffered most from the MDB’s increasingly severe salinity problems, lobbied hard for action on this issue and, when frustrated, fought court battles in NSW to block irrigation projects there. Victoria, too, was increasingly irritated with rising levels of extraction further upstream. Yet even the downstream states, while anxious to press for changes upstream on environmental grounds, often proved reluctant to scrap practices of their own that critics argued were environmentally damaging or unsustainable.

The dearth of clear potential “winners” from the above list of actors is significant. Water reform was largely about the allocation of losses rather than the realisation of gains. The case for reform rested on the need to ensure long-term sustainability of water use, to reduce environmental damage and to ensure that a scarce resource was allocated efficiently. While both environmentalists and advocates of market reform favoured change on these grounds, most water users stood to lose, to the extent that they were being asked to pay meaningful prices for a resource they had previously received for free or at a nominal price. There was little understanding of the long-term economic benefits of reform or the costs of inaction, partly because there were no authoritative studies really quantifying these costs and benefits.
CHAPTER 18. AUSTRALIA: WATER REFORM, 1994-2004


Issue history

Before the 1980s, policy was dominated by social and strategic objectives with little regard for the opportunity cost of water use. There was much emphasis on engineering solutions (dams, drought-proofing, etc.) and on increasing supply to address shortages. Little attention was paid to economic analysis, efficiency or cost recovery. The development of irrigation was linked to policies explicitly intended to encourage rural settlement, and institutional arrangements often encouraged exploitative management. The result was over-allocation of water supplied below cost and a lack of incentives to conserve. Water charges were based primarily on land values rather than consumption, and different valuation systems were used for residential and non-residential urban sectors and for the rural sector. In many places, virtually every consumer faced a different average cost of water, while the cost of the marginal unit consumed was usually zero. Non-farm businesses typically paid 2-3 times as much per litre as did households, and agricultural users were charged a small fraction of what households paid. Only a small part of this gap reflected differences in the cost of supply and in the quality of water supplied. Charges in many irrigation networks did not even cover operation and maintenance costs. There was thus little incentive to ensure that water migrated to its highest-value uses, and its marginal productivity varied widely.

This approach came in for increasing criticism as Australia became a “mature” water economy, characterised by increasing scarcity, rising marginal supply costs, intensified competition for water resources and increased interdependencies among water users. Policy debates from the 1970s onwards reflected growing awareness of the inelastic supply of “new” water, the need for expensive rehabilitation of ageing infrastructure and the agro-environmental challenges posed by decades of irrigated agriculture (salinity, rising soil acidity, loss of biodiversity). Fiscal considerations also played a role, as state budgets came under pressure. By 1990, most states were committed to recovery of most, if not all, the costs of supply and some had begun experimenting with pricing reform and, on a small scale, with tradable water entitlements. Efforts were made to improve the management of municipal water utilities and irrigation networks alike.

Progress was variable but significant: Victoria cut the deficit arising from the operation of its irrigation systems by 80% between 1984 and 1994, while NSW effectively privatised much of its irrigation network to the irrigators themselves. The 1980s also witnessed increasing efforts to coordinate state policies towards the MDB, with SA, Victoria, NSW and the Commonwealth signing the Murray-Darling Basin Agreement in 1987, leading to the creation of three new inter-governmental bodies, including the Murray-Darling Basin Commission. A common Salinity and Drainage Strategy was agreed the following year, which brought about some improvement in salinity levels. However, sharp differences between SA and NSW over the detail of the strategy also highlighted the distributional conflicts among the states that integrated MDB management would face, particularly given the continued requirement for unanimity among the states. National-level actors in water policy were largely confined to advisory roles.

Urban water was less controversial, partly because inter-state conflicts did not arise but also because quality and reliability were satisfactory. All major metropolitan water authorities generated sufficient revenue to cover operation and maintenance costs, as well as depreciation, and to service their debts, while paying a return on capital. The main problem was the disconnect between consumption and price arising from water rates based on property values. The large gap between households and non-farm businesses...
was a clear case of cross-subsidy, since the former were more expensive to supply. These arrangements were first revised by the Hunter District Water Board in NSW, which introduced user fees in the early 1980s, in response to financial stress, rising demand and the need to invest in a major new dam. The result was an improvement in the board’s financial position, as well as a fall in average consumption per household that enabled it to postpone construction of the dam. The Hunter example underpinned calls for pricing reforms in urban water by the Industry Commission.

Policy process: formulation and adoption

By the early 1990s, it was increasingly clear that the MDB was in bad shape environmentally and that this was having a particular environmental impact on the downstream states. An audit of water use in the MDB, commissioned at SA’s request in 1993, found that median annual flows at the mouth of the Murray were less than 30% of what they would have been under natural conditions and that the percentage of years in which the lower Murray was experiencing drought conditions had risen from 5 to 60%. Diversions had tripled over the preceding 50 years and were still growing. In most years, more water was approved for diversion than was available, and man-made changes in the seasonal pattern of flow had significantly affected water quality and biodiversity in the Basin. These concerns were rising up the domestic political agenda at a time when environmental issues were the subject of increased political attention at the highest levels. As a signatory of Agenda 21, the UN Conference on Environment and Development action plan adopted in 1992, the Commonwealth was committed to applying the principles of environmentally sustainable development set out in Agenda 21. These thus became embedded in water policy. A National Strategy for Ecologically Sustainable Development was ratified in 1992 by the Commonwealth and the states. Though not legally binding, it added to the pressure to protect the MDB.

The environmental case for reform was reinforced by a 1992 Industry Commission study advocating a range of water reforms on economic grounds. Yet while the Commission report was both wide-ranging and deep, considering both urban and rural sectors, it stopped short of trying to quantify the costs of the status quo or the potential gains from the reforms it was advocating. Moreover, the evidence submitted to the Commission and cited in the report suggested that many of the interest groups and actors involved were focused on distributional issues, seeing reform largely in static zero-sum terms – an early indication of the political challenges the reform would face. Thereafter, the Commission did little work on water for more than a decade. In part, this probably reflected the resource constraints facing the Commission and cited in the report suggested that many of the interest groups and actors involved were focused on distributional issues, seeing reform largely in static zero-sum terms – an early indication of the political challenges the reform would face. Thereafter, the Commission did little work on water for more than a decade. In part, this probably reflected the resource constraints facing the Commission, but it also seems to have reflected national policy priorities: when the drought of the early 2000s pushed water reform up the Commonwealth government’s agenda, the Commission produced a series of major reports on water issues.

In response to these developments, the Commonwealth and the states, acting through the Council of Australian Governments (CoAG), established a Working Group on Water Resource Policy to develop a framework for water industry reform. In February 1994, its recommendations were approved by CoAG almost unchanged. The basic template for the reform largely reflected the principles set out in the 1993 National Competition Policy Review Committee report (the “Hilmer report”), which formed the basis for the emerging National Competition Policy (NCP) agenda. Indeed, one recent assessment of the National Water Framework concludes that economic efficiency and competitiveness, rather than resource management experience or environmental concerns, were the
principal drivers for the biggest policy changes adopted in 1994-95, although both sets of concerns mattered. When the NCP was finally launched in 1995, the Water Reform Framework was incorporated in it. Under the NCP, the National Competition Council (NCC), an independent statutory body, was given a mandate to monitor and report on the states’ fulfilment of their NCP reform commitments, as well as to recommend the granting of NCP payments by the Commonwealth to the states in return for making good on those commitments. If the NCC found that states were not fulfilling their NCP obligations, the Treasury could delay or withhold such payments.

The CoAG Water Reform Framework

The water framework included both pricing and institutional reforms and changes to future policy making in respect of water:

- Charges were to be consumption-based. Where possible, urban water utilities were to rely on two-part tariffs, comprising a fixed charge for access to service and a variable user charge based on consumption. By 2001, urban and rural water charges were to reflect full cost recovery with transparent subsidies and the elimination of cross-subsidies.

- Many of the institutional changes approved were needed to create efficient water markets, including the creation of volumetric and tradable water entitlements that were separate from land titles, and the institutional separation of regulation, water service delivery and resource management functions. The states were to create comprehensive systems of water allocations or entitlements, specifying who could abstract water, where, when and on what terms. These were to include allocations for the environment.

- Future water policy was to provide for the use of water for environmental and other public benefit objectives. This entailed more rigorous appraisal of new irrigation projects, based on both economic and environmental sustainability criteria; the allocation of water for environmental uses; and measures to address water quality.

The most immediate change, however, was the imposition of a temporary cap on diversions from the MDB at 1993-94 levels. The cap was conceived as an interim measure to prevent unsustainable demand growth and establish a stable division of water between the environment and production, while a more market-oriented regime was put in place with the aim of reducing extraction to environmentally sustainable levels. Indeed, the cap provided a stimulus to the development of water markets, since administering it required the development of a comprehensive system of water allocations and rights without which water trading would not be possible. It was hoped that trading under the cap would ensure that scarce water resources migrated to higher-value uses.

To assist the states, CoAG created a Water Reform Task Force, consisting of senior federal and state water agency officials, to monitor progress and to promote a uniform approach to the implementation of the reform, particularly by providing guidance on the most controversial questions, such as the best methods for taking account of externalities in establishing the full cost of water consumption, and the development of guidelines for pricing. The complexity of these tasks reflected the exceptional difficulties associated with the creation of markets in water, which, though not insurmountable, were formidable (Box 18.1).
Box 18.1. Economic properties of water and the creation of water markets

If water from a given source is allocated efficiently, the marginal social benefit for the different users should be the same. A well functioning market will provide price signals indicating the scarcity of the resource and ensuring that users derive equal marginal benefits from consuming it. Achieving such outcomes, however, has proved far more complex in practice than many realised when the water reform was launched, for several reasons:1

- Some benefits of water have public good characteristics (e.g. environmental flows, recreational uses).

- Water can be either a stock resource (groundwater, some surface water sources) or a replenishable flow resource. It is not always clear which is the case, making it particularly difficult to determine the environmentally sustainable level of extraction.

- It is hard to measure consumption: gross diversions are easiest to measure, but a good deal of diverted water returns to the system for reuse at a later time and a different location. Hence, it is much harder to assess net diversions, particularly where the extraction-and-return process affects the quality of the water returned to the system. Indeed, water is difficult to measure and identify, given the variety of forms that it takes and in which it may be extracted. This makes exclusive property rights more difficult to establish or enforce, and makes it much harder for agents to keep track of the good – a key concern in any market.

- Water consumption generates numerous externalities which are still only imperfectly understood. Some, such as pollution, exist even in the absence of water markets. Others arise in conjunction with the move to water trading. When water is consumed or traded, the impact on third parties may be significant and may depend not only on how much water is involved but also on when it is consumed or traded, where it goes and whether/how its use changes.2

Given these realities, government intervention is required to ensure sufficient environmental flows, and some regulation of water trading is required to address other externalities and achieve equality of marginal social benefits between users of water. The complexity of the regulatory challenge is compounded by the need to take proper account of the legacies of past policies. Thus, although it was widely believed that the scarcity value of water was rising and was probably far higher than service-delivery cost,3 it was feared that a rapid transition to scarcity pricing would do severe damage to some water users who had invested under the previous policies, while generating windfalls for others.

More generally, there is still much controversy about basic hydrological issues, such as the extent to which environmental flows mitigate or reverse environmental damage, the relationship between groundwater and surface water, and the extent to which recent changes in rainfall trends reflect climate change, as opposed to “long cycles” in patterns of rainfall.4 This makes valuing externalities all the more difficult. It also creates contradictory incentives for governments trying to design water markets. Environmental uncertainty gives the state an incentive not to create very strong private property rights, since stronger property rights raise the costs of any subsequent intervention to address environmental problems. However, if the state does not create strong property rights, policy uncertainty will raise transaction costs (Crase et al., 2001; Vourc’h and Price, 2001). Scientific uncertainty always makes policy–making more difficult, but its impact was made all the greater by awareness of the potentially tremendous distributional consequences of alternative hypotheses about issues like climate change, levels of sustainable extraction and the interactions between different water sources and uses. Thus, debates about investment in irrigation infrastructure upgrades in the MDB to a significant extent turn on highly contested evaluations of likely long-run rainfall there.

1. Many of the issues discussed here are addressed in Brennan (2001).
2. For example, some water trades may involve changes of use that reduce return flows to the system, adversely affecting downstream users. Trade also tends to reduce off-season water consumption and increase diversions during periods of peak demand, which can create congestion problems and local shortages. See especially Heaney and Beare (2001); MIL (2006); and OECD (2008).

3. Brennan (2001) notes that fees for service delivery, based on full cost recovery for water supplied to irrigated agriculture, were around AUD 20 per megalitre at the end of the 1990s, as against the value of water traded in the market, which was then about AUD 100/megalitre.

4. Recent observations by the Meteorology Bureau establish a link between the severity of the present drought and the effects of climate change (Garnaut, 2008). Those taking a sceptical viewpoint point out that the first half of the 20th century saw rising rainfall in many regions where it has declined since 1950 and that overall rainfall levels were higher (on the basis of ten-year moving averages) in the early 2000s than in the first decades of the last century.

Policy process: implementation

Agreement on the 1994 CoAG package had been relatively easy, both because the implications of some proposals were not well understood and because the package comprised a set of principles, targets and incentives and left the states to handle the details of implementation. However, as stakeholder groups became aware of the effects of reform, implementation became harder. There was considerable conflict over water-sharing plans, the allocation of water for environmental needs, changes in water prices and competition between agricultural and urban uses. National Competition Council (NCC) assessments found that water management plans tended to reflect irrigator interests, effectively defining “environmental water” as the water left over after supply to diverters. This was hardly surprising given the direct involvement of local stakeholders in the preparation of the plans.55

As supply constraints grew tighter, following the onset of a major drought in the early 2000s, there was renewed interest in “engineering solutions”, involving government support for large-scale investments in the efficiency of irrigation infrastructure; the “water savings” thus generated could be used for the environment.56 For the Liberal-National coalition, with its significant rural electorate, the appeal of this approach was obvious. However, though presented as a way to ease supply pressures, and thus mitigate distributional conflicts around water policy, such proposals came in for criticism on account of their own distributional consequences. While some forms of “increased efficiency” could indeed reduce wasteful evaporation, others effectively reduced the return flows from water diverted. This meant that upstream consumers could use more water, reducing supplies to those downstream.57 In addition, more recent projections suggesting that annual runoff in the MDB might fall by 20-40% over the next 50 years raise further questions about the long-term value of such investments.58

NCP payments from the Commonwealth to the states, which did much to advance reform in other domains, proved to be of only limited use, largely because the water framework was too general to commit the states to much that was really binding. The 1994 agreement was only about ten pages long, far shorter than the National Water Initiative, which succeeded it in 2004. The states could thus collect their NCP payments for water reform while continuing to pursue policies that were at odds with the water framework. Thus, neither the NCC nor other national-level bodies could penalise the states for much more serious reform failures, such as NSW’s repeated breaches of the cap or its failure to adopt the kind of compliance models recommended in successive MDB Ministerial Council reviews.59 These commitments were simply not part of the NCP process. The NCC did exercise what power it had, at one point withholding payments
from NSW following a sharply critical review of that state’s water management plans, but in general its role was limited.60

**Urban water**

The 1990s witnessed significant progress with respect to urban water reform: most urban water authorities introduced consumption-based charges for households and firms. The norm was usually some form of two-part tariff, the most common being “inclining block tariffs” that confronted consumers with rising per-unit charges as consumption increased.61 Consumption-based charges made up an increasing share of water utilities’ revenues.62 Free allowances, which had been large and widespread, were largely phased out in favour of charges starting from zero consumption. Assistance for low-income households was provided through clearly defined and separately funded community service obligations. Cross-subsidies from industrial and commercial users to households were reduced or eliminated. However, state and territorial authorities remained reluctant to apply full cost-recovery principles to small local authorities, because of limitations on the ability of many communities to pay.63

When addressing cost-recovery issues, the CoAG task force had defined upper- and lower-bound levels of cost-recovery: to reach the lower bound, water businesses had to recover operational, maintenance and administrative costs, taxes or their equivalent,64 and interest charges on debt, while making provision for asset maintenance and replacement. Upper-bound cost recovery – defined as a maximum, in order to avoid monopoly rents for water utilities – included all of the above plus the weighted average cost of capital. Full lower-bound cost recovery for water storage and delivery was largely achieved by 2004 – the minimum obligation under the 1994 framework – but most utilities were still some way from the upper bound of cost recovery.

Consumption-based tariffs resulted in falling individual water consumption. Water use in Brisbane, for example, was reduced by 20% between 1995-96 and 1997-98, after the adoption of metering and consumption-based charges. Across Australia, the cost per unit of water in metropolitan areas fell by more than 16% on average between 1992-93 and 1997-98.65 Most consumers, however, did not feel the benefit of this change, the unwinding of cross subsidies meant that they faced higher water charges, particularly in the largest cities: real prices for commercial water consumers fell sharply between 1991-92 and 1996-97 – by 40% on average in the two largest cities, Sydney and Melbourne. In many small towns, however, explicit community service obligations were imposed,66 with utilities supplying water below cost and the states making good the water companies’ losses. Moreover, as supply tightened owing to a protracted drought in the early 2000s, the use of price for purposes of demand management was limited in most urban areas. Non-price regulation of urban water use (usually prohibitions or restrictions on specific uses) remained a widely employed response to supply constraints, even where these were more or less permanent. Since supply is price-inelastic in the short run, quantitative restrictions arguably make sense in response to short-term shocks, but over the long run, prices are more likely to create incentives to reduce consumption and invest in water-saving technologies and/or new supply options, such as recycling.67

**Rural water**

*Problems administering the cap*

From the start, administration of the cap on extractions from the MDB proved highly problematic, owing to the poor quality of the historical data on water use and the
existence of very different management rules in the three states that were parties to the cap agreement, SA, Victoria and NSW. There were also very real problems simply monitoring the level of current diversions. With the exception of bulk diverters, the states had little capacity to regulate water supply to licence-holders when restrictions were imposed. Diversion estimates were neither reliable nor homogeneous across jurisdictions. These problems were never satisfactorily resolved, making it difficult to assess the scale of over-allocation. The effectiveness of the cap was further undermined by the fact that it did not regulate certain forms of diversion of surface water, such as on-farm dams and rainfall captured by tree crops before it could enter the system, nor did it cover groundwater. Water could thus be appropriated at many points in the water cycle at no cost, even though extraction at these points in the system affected its availability elsewhere. Landowners therefore rushed to capture water from such sources before controls could be imposed.

The creation of tradable water rights actually stimulated demand in some areas. Prior to the cap, water entitlements had been issued far in excess of what could actually be allocated, even in a relatively wet year. The problem was particularly acute in NSW; SA had historically been very conservative in assigning water rights and Victoria had only a modest over-allocation problem. Over-allocation of entitlements did not necessarily imply over-use, since physical allocations depended on flow levels, and there were in any case many water entitlements that had never been exercised (“sleepers”) or had been exercised only rarely (“dozers”). Under pre-cap arrangements, NSW had redistributed surplus water that was not used by sleepers and dozers to established irrigators at the same low cost as the latter paid for water diverted under their own entitlements. In making allocations under the cap, however, NSW treated inactive and rarely exercised entitlements as unconditional rights. This prompted large numbers of sleepers and dozers to “wake up” and claim their water, in order to trade it. Irrigators thus had to buy water from the former sleepers and dozers at market prices rather than being allocated it by the state at the usual low prices. This led many to complain that their water allocations had been cut by the cap. Problems also arose in SA, which in theory had most to gain from the cap and had pushed hard for a co-ordinated approach to basin management: the SA government came under fire after announcing a large additional allocation to irrigators that was clearly not compliant with the cap. This generated no externalities for the other states, since all were upstream of SA, but the decision undermined support for the cap further north.

Queensland never became part of the cap process at all: although its Labor government had agreed in principle to the adoption of the cap in 1995, there was considerable resistance to it in that state and it was never applied there. On the contrary, Queensland governments saw further opportunities to expand irrigated agriculture upstream and continued to issue new entitlements, arguing that the state’s farmers should not be penalised simply because it had been slower than the other states to assign water entitlements. Agrarian opposition to the cap came to the fore in the 1998 federal election, when the NP campaigned to “zap the cap”. Resentment of the cap also ran deep in NSW, in part because of the treatment of sleepers and dozers but also because successive governments there failed to explain the benefits of the cap as a means of protecting security of supply – a failure some observers believe was at least partly strategic, prompted by a desire to deflect farmers’ anger from the state government to the MDB Commission. If so, NSW was not alone: other states also allowed the cap to be seen as an imposition of the MDB Commission rather than something agreed by state governments.
In 2000, a five-year review of the cap concluded that the degradation of the MDB river system would have been worse without it but that the deterioration had not been arrested. In part, this was because environmental flows had not been maintained in dry years: Commonwealth policies to reallocate water for environmental purposes were substantially diluted by agricultural pressure groups, which generally dominated catchment management authorities and thus largely controlled the development of catchment management plans. The result was that environmental flows bore a disproportionate share of the losses in dry years, so that agricultural production would not suffer. The volumes of water allocated for diversion in some areas during dry years exceeded the volumes available for diversion – in other words, it was shortage of water, rather than the cap, that constrained consumption. The five-year review also highlighted the need to integrate management of groundwater and surface water and to introduce better systems for metering, monitoring and reporting data, agreed cap definitions and less restrictive trading rules within and between jurisdictions. Integrated management of groundwater and surface water is very difficult technically, because the relationship between the two is imperfectly understood and varies within and across catchments. Australia’s 340 surface and 367 groundwater management districts are connected hydrologically but continued to be managed by separate local authorities.

Obstacles to the development of water markets in the irrigation sector

The development of a water market was affected by many of the factors that complicated administration of the cap. Here, too, data issues were critical: the states had to develop water registers, better mechanisms for monitoring diversions and systems of water accounting to record trades, as well as the biophysical knowledge needed to make judgements about environmental flow thresholds, externalities arising from trade and so on.

The CoAG agreement did not specify the form that property rights to water should take. The Commonwealth’s Standing Committee on Agricultural and Resource Management took this task on, but significant differences still emerged across states, and intended property rights were substantially modified as the reform was implemented: water remained state property, private rights were limited to specific entitlements to divert and use water rather than ownership, and restrictions were imposed on who could possess rights and who could participate in which water markets. The water rights that were created differed across states but tended to share certain characteristics. Entitlements generally specified a volumetric allocation, but the actual allocation might be much smaller, due to the high variability of supply; the entitlement was really to a share of the consumptive pool of a specified water resource. In some states, the adjustments to entitlements in periods of low water supply represented a proportional re-allocation of entitlements, albeit with some regard for the potential injury to different water users. In others, a priority system was created, with “high security” rights for some (usually for permanent crops) and “low security” rights for others (mostly annual crops). Even in relatively wet years, adjustments might be required, because a variety of contingent entitlements to water had been issued without charge prior to the reform, often where it was clear that not all of them could be exercised.

Both temporary and permanent water trades were heavily regulated by state and local authorities. Permanent transfers, transfers between catchment areas and transfers of irrigation water to other uses were all subject to various limitations and often outright prohibition. While environmental externalities prompted some restrictions, regulation
also served to protect irrigators from the substantial rises in price that might otherwise have occurred.\textsuperscript{80} There was widespread fear that trading in water rights would lead to the emergence in rural areas of “water barons” – non-agricultural investors from outside the region who would buy up water rights for “speculative” purposes.\textsuperscript{81} The sale of water entitlements outside an irrigation district could impose costs on the remaining irrigators, who shared the fixed infrastructure costs of the irrigation network, which often accounted for a large proportion of total delivery costs.\textsuperscript{82} However, restrictions on such trade and fees on those exiting irrigation networks served to impede network rationalisation. In part, the determination to restrict exit may have reflected local governments’ own fears of the fiscal consequences of water trade: if a landowner sold his water entitlement, his land was then re-classified as dry land for the purposes of assessing property taxes. Restrictions on trading outside the irrigation district where the water was being used were thus retained virtually everywhere, and trades could be refused if they were inconsistent with local water plans. There were also limits on the share of individual rights that could be transferred on a permanent basis, even within a given district.

Water markets were thus more often than not inaccessible to urban users and reserved for irrigators; intermediaries were effectively excluded.\textsuperscript{83} Markets were also relatively thin: trading volumes were small and mainly reflected temporary transfers within irrigation districts. Altogether, temporary intra-state trades, usually between irrigators in the same region or the same hydrological management basin, tended to involve 5-8\% of total diversions from the MDB. Trade in permanent entitlements grew slowly: by 2003-04, the cumulative volume of rights transferred under intra-state permanent trades in the MDB was only around 7\% of total diversions. Trading also tended to be highly localised and there was little information available about trades, though water exchanges were eventually created in an effort to address this problem in the lower basin. The absence of intermediaries from the market impeded market deepening; they could, given the complexity of the market, have done much to lower search costs, increase transparency and spread risk.

Cross-border water trading was rendered far more difficult by the states’ adoption of different packages of water rights and different regulatory regimes, and interstate trades were only introduced towards the end of the 1990s, and even then on a trial basis and initially only for high-security entitlements on the lower reaches of the Murray. The volumes involved were small – less than 0.5\% of total diversions in the MDB in the case of temporary inter-state trades and only a tiny fraction of that amount in permanent entitlements – but the pattern of trade nevertheless generated friction among the participating states. Water was mostly traded downstream, contributing to the viticulture boom in SA. Victoria saw the pattern of water trading as a product of the salinity and drainage strategy. Developers in Victoria had to buy salinity credits in order to develop irrigation projects, or else undertake substantial engineering works to offset the salinity impact of their activities. In SA, they only had to promise to mitigate salinity impacts in the future, and when they emerged – and there was some doubt as to whether those commitments could be enforced. It was therefore more attractive to trade water over the border into SA than to buy salinity credits in Victoria. SA was eventually pressured to undertake engineering works to offset the salinity effects of developments already under way, but this was done at the expense of the state budget, not the investors. More generally, debates about how to limit the externalities created by cross-border trade, including “tagged” and “exchange-rate” trading, continued well into the 2000s.\textsuperscript{84}

Despite these difficulties, water markets did take shape. Volume charges were introduced in most irrigation systems by 2001. However, 17\% retained a land-area
component in the charge structure, and volume charges were based on average costs for
the irrigation group even where there was wide variation in the cost of providing water to
individual irrigators. Moreover, the charges levied on irrigators exercising their own
entitlements did not reflect the relative scarcity of water and most fell short of
“upper-bound” cost recovery as defined under the reform. Thus, although the use of price
signals to allocate rural water increased substantially, the water supplied to agriculture
remained much cheaper than that supplied to households and other urban users.

The urban-rural interface

Continued separation of urban and rural markets substantially limited the economic
benefits of water trading and created further distortions. For example, an irrigation user
who could cut leakage or waste for AUD 150 per megalitre had little incentive to do so
when irrigation water was trading for around AUD 100 per megalitre, even though the
cost of water in urban areas was around AUD 1000, making less cost-effective
water-saving options attractive there. Yet resistance to urban-rural trade remained
deeply ingrained in rural areas, even when the farmers involved were eager to sell.
Opposition to such trade was rooted in the same concerns as opposition to trade between
catchments: fear of “speculation”, asset stranding and the loss of productive capacity in
agriculture, as well as the perceived “wastefulness” of some forms of urban water
consumption and the conviction that permanent diversions of water out of irrigation
would undermine the vitality of rural communities. In fact, the scale of potential
diversion from rural areas was probably limited: urban use accounts for about 25% of
total consumption. If 20% of urban water demand were met by transfers from irrigation,
the amount transferred would be equal to about 5% of total consumption, or 7% of
existing irrigation use. Though small in relation to total consumption, this could have
significantly affected the availability of urban water, helping the major cities to reduce
their reliance on quantitative restrictions. For urban areas located inland, purchasing
water access entitlements from irrigators was one of the few attractive ways to meet
rising demand.

Conclusion

The 1994 Water Reform Framework can hardly be dismissed as a failure. It achieved
real progress in both urban and rural sectors, and brought about a paradigm shift in the
way water reform is viewed in Australia. Water policy debates are now framed in terms
of the principles of environmental and natural resource economics. Nevertheless, the
1994 framework fell far short of its original goals. The contrast between Australia’s water
and electricity reforms serves to highlight many of the reasons for this outcome:

• The technical and scientific challenges presented by water policy, particularly in
  the rural sector, were exceptional. As technically challenging as power-sector
  reform was, the physics of electricity generation and transmission were well
  understood. It did not involve the same level of scientific uncertainty as water
  reform.

• The distributional conflicts that arose, both between different economic sectors
  and between the states, were particularly acute in respect of water reform. In
electricity, the opportunity to realise potential gains through cooperation
prompted the states to work together and, in particular, to agree to the creation of
national-level institutions to govern the new power market, but this sort of
positive incentive was lacking in the case of the water framework, because reform was largely about the allocation of losses. Since every state was a *de facto* “veto player” and none was under pressure from a strong pro-reform lobby, co-operation proved difficult in the absence of a crisis or action-forcing event, and the disciplines of the NCP were simply too weak, given the limited nature of their reform commitments.

- Power-sector reform was driven during a critical early phase by the two largest states, working together with the Commonwealth. This constituted a powerful coalition for reform. When it came to water, however, Victoria and NSW were frequently in conflict.

In the rural sector, the need to respect acquired rights presented a further challenge: since the total volume of allocations was unsustainable, the conversion of limited and contingent entitlements into largely unconstrained and tradable rights should have been accompanied by a reduction in volumetric entitlements. However, this would have involved the withdrawal of rights that irrigators had long held and on the basis of which they had invested in good faith – often under policies explicitly designed to encourage rural settlement and irrigation development. Implementation of the 1994 reforms was further complicated by the power of the catchment management authorities, which were often dominated by local irrigator interests and which tended therefore to give priority to maintaining agricultural production rather than environmental needs when allocating scarce water. Finally, it should be borne in mind that many of these problems reflected the somewhat experimental nature of the reform, which really was at the “international frontier” in terms of water resource policy. Yet this in itself made it difficult to define the endpoint of reform with clarity: while the general objectives of efficiency and environmental sustainability were widely shared, the reform process was very much a case of “learning by doing”. It could hardly have been otherwise, but this was a further factor complicating the task of building and sustaining support for reform.

In recent years, growing concern about climate change and the impact of the drought have raised public awareness of water problems and strengthened the case for change. This has contributed to the re-launching of water reform with a number of new initiatives. The Living Murray programme, agreed in 2002 and launched in 2004, aimed at restoring 500 gigalitres per year to the environment (equivalent to almost 5% of extractions) and involving a higher level of inter-jurisdictional coordination than previous efforts. In 2004, CoAG approved a new National Water Initiative (NWI), which was considerably more ambitious than its 1994 predecessor and which – crucially – attempted to overcome the unanimity rule, which had thwarted previous reform attempts, and to move to true management of the MDB as an integrated whole. The NWI also encompassed groundwater in the reform framework. It was followed by a new national Water Act in 2007, which was introduced to address shortcomings in the operation and governance of the MDB. In addition, a broad programme of urban and rural water reforms was agreed between the Commonwealth and the states in 2008 for development within a CoAG framework. Building on the NWI, the new programme aims to address over-allocation and achieve environmental outcomes, to improve the operation of water markets and trading, and to tackle key urban water challenges.
Notes

2. For assessments of water reform in international comparison, see Isaac (2003); Allen Consulting (2004).
3. Quiggin (2006b). Australia ranks well above the world average in terms of per capita water availability: UNESCO (2002) ranks it 40th, ahead of such OECD members as Switzerland (77), France (104) and Spain (117).
6. On the basis of ten-year moving averages.
8. Whittington et al. (2000).
9. “The Commonwealth shall not, by any law or regulation of trade or commerce, abridge the right of a State or of the residents therein to the reasonable use of the waters of rivers for conservation or irrigation.”
11. On the NCP, see Box 17.1 of the case study of power-sector reform in Australia.
14. This has since changed, as prolonged drought and growing concern about climate change have raised public awareness of Australia’s water problems.
15. Connell (2007). As noted above, this changed over the years, owing to the increasing salience of concerns about climate change and the drought.
18. This was particularly true in view of indications that One Nation was drawing a disproportionate share of its support in areas where the NP was traditionally strong; see Charnock and Ellis (2004).
20. Connell (2007) traces these conflicts back to the 19th century, including, e.g. South Australia’s concern for navigability of the lower Murray and its tributaries, as against Victoria’s claimed right to divert water from the upper Murray and its tributaries within Victoria, and New South Wales’s claim to exclusive use of the waters of the Murray and its tributaries within its borders.
21. For details on this conflict, see Connell (2007:122-23).

22. At one point, these efforts prompted NSW to adopt legislation to exclude the South Australian government from NSW courts.


24. See below for a discussion of South Australian policy on the salinity effects of development.

25. The Australian Conservation Foundation supported the inclusion of water in the National Competition Policy, stressing the need to correct unsustainable water use patterns.

26. If agents had compared the long-run consequences of reform and non-reform, many more might have identified themselves as potential winners. Most seem to have compared their positions under a status quo that was assumed to be sustainable and their positions under reform.

27. Brennan (2001); Connell (2007); Quiggin (2006a). Quiggin (2007:5) observes that water-supply investment was “rarely subject to economic appraisal and generally involved substantial government subsidies.”


29. Even in the late 1990s, some water authorities received over 80% of their revenues from government subsidies and other sources than user fees. See Tisdell et al. (2002); and Hussey and Dovers (2006).

30. This flat, property value-based fee was paid for a base or free allowance of water; however, as Musgrave (2000:304) observes, free allowances were generous, so few consumers exceeded them and incurred volumetric charges. See also Hall et al. (1994); Cox and Warner (2007).


32. Hall et al. (1994).


34. Tisdell et al. (2002). The last of these points reflects the properties of water economies: increasing water scarcity has also contributed to increase the externalities involved in water diversion and use.

35. OECD (2001); Tisdell et al. (2002); Quiggin (2007).


38. The others were a Murray-Darling Basin Ministerial Council and a Community Advisory Committee for the MDB.

39. There was a serious conflict over the benchmark date after which states were responsible for taking remedial action to counter the salinity impacts of irrigation projects. NSW wanted to limit such action to new developments, while SA wanted to cover most of the 20th century.


43. Hussey and Dovers (2006). A spectacular algal bloom extending 1000km along the Darling River during the summer of 1991-92 helped focus public attention on the issue.

44. MDBMC (1995). The audit found that increasing water use in the MDB had been encouraged in order to justify government investment in dams and infrastructure.


49. See the power-sector case study for detail on the Hilmer report and the launch of the NCP.


51. Including, to the extent practicable, positive rates of return on the written-down replacement costs of the assets.

52. More precisely, the cap for each year was not the volume of water used in 1993-94 but the volume that would have been used with the infrastructure existing in 1993-94, given climatic and hydrologic conditions similar to those experienced in the year in question. See Vouc’h and Price (2001).


54. Connell et al. (2005); Hussey and Dovers (2006).


58. See van Dijk et al. (2006).


60. In 2005, NCP payments to NSW were cut by 10% (AUD 26 million, later reduced to AUD 13 million).

61. In other words, the supply curve looked like a stair case, a series of horizontal curves that shifted vertically at specific quantities as consumption increased.

62. See Cox and Warner (2007) for detail. By the time of the 1999 NCC assessment, all but two major urban water authorities had adopted two-part tariffs with no base allowance.

63. NCP payments do not go to local authorities, so they had little incentive to pursue reform unless provision was made for sharing NCP payments with them. See Musgrave (2000).
64. State-owned utilities were not subject to tax but paid tax-equivalent payments to the governments that owned them.


66. A community service obligation arises when a government requires an enterprise to carry out activities relating to outputs or inputs that it would not do voluntarily except for higher prices. The Commonwealth maintains that the nature of such obligations should be explicit and public.


68. Metering the hundreds of direct pumpers in the MDB remained an issue a decade after the cap was imposed; see MIL (2006).


70. This is well described in Quiggin (2001); see also MDBMC (2000) and Connell (2007:125-26).


73. Whittington et al. (2000).

74. MDBMC (2000).

75. OECD (2008).


79. See Vourc’h and Price (2001:13). Transfers between catchments were largely limited to the MDB and a few places in Western Australia; in most other regions, such trade was physically impossible.


81. With this in mind, several states eventually moved to prohibit or cap the ownership of water rights by persons not owning or occupying land, or to restrict the proportion of entitlements in any given catchment that could be held by non-farm users.

82. Quiggin (2006b).


84. Under exchange-rate trading arrangements, a water entitlement within a state is cancelled, extinguished or suspended, and an equivalent – but not necessarily identical – entitlement is created in another state. Exchange rates are based on principles that minimise any third-party impacts of the transaction. In tagged trading, water transferred from one catchment or state to another always retains the characteristics of the entitlement to which it was originally allocated.


86. OECD (2008).

87. For an assessment of the potential benefits of greater urban-rural trade, see Dwyer et al. (2005).
89. Quiggin (2006b).
90. Quiggin (2007).
92. Quiggin (2007:8).
Bibliography


Chapter 19

Germany: Retail opening hours, 1999-2003

Summary and main conclusions

In late 1999, a combination of pressure from some of the federal states (the Länder) and the publication of a pair of conflicting studies commissioned by two federal ministries triggered renewed discussion of the need to liberalise Germany’s traditionally restrictive law on opening hours for retail shops (the Ladenschlussgesetz). The ensuing three-and-a-half years of debate resulted only in a modest further liberalisation, allowing Saturday opening until 20h00 and relaxing some other minor restrictions. Retail employees were guaranteed at least one Saturday per month off work. This compromise did not last long, as a June 2004 Constitutional Court ruling opened the way for the Länder to take over the regulation of opening hours, which they did from late 2006 following a broader reform of the Basic Law. The great majority of the Länder opted for extensive further liberalisation and some scrapped almost all restrictions on Monday-Saturday opening hours. The speed with which the Länder liberalised raises the question of why years of debate at federal level failed to produce a more significant result through normal political – rather than judicial – processes.

The inability to reach a federal-level agreement on a fundamental liberalisation seems to have reflected a number of factors:

- **Institutional factors.** Opponents of reform benefited from a federal policy process characterised by the presence of numerous “veto points”. Germany’s inclusive and consultative approach to policy making, coupled with the need to pass much legislation through two chambers of parliament that are often controlled by different parties, means that the policy process is characterised by the need for multiple negotiation processes. This can present obstacles to reform and ensure that change, when it occurs, is usually incremental and the product of multiple negotiation processes.

- **Alliances of vested interests.** Supporters of the Ladenschlussgesetz included a variety of groups motivated by different agendas: smaller retailers, especially in suburban and rural areas, who feared a substantial reduction in the number of smaller shops; trade unions, who wished to prevent what they viewed as undesirable changes in working conditions and check the trend towards greater reliance on part-time employment; the Ministry of Labour and Social Affairs, which focused on the Ladenschlussgesetz’s role as a law protecting retail employees and regulating working time; churches, which cited Sunday trading and the protection of family life; and so on. These disparate agendas intersected where the Ladenschlussgesetz was concerned, so each group could rely on allies in defending its particular interests.
• **Failure to sell the reform.** Divisions within the federal administration, with the Ministry of Labour and Social Affairs favouring retention of the law and the Ministry of Economics and Technology pressing for reform, made it harder for proponents of liberalisation to advance the case for reform. Studies commissioned by the two ministries failed to find consistent evidence that the previous (incremental) relaxation of the law had contributed to the growth of retail turnover and retail employment, as supporters of reform had argued it would. This apparent over-selling of the 1996 reform put reform advocates on the defensive at a time when employment conditions were deteriorating and concerns about the labour market consequences of reform were growing. The studies did point to the benefits of reform for consumers, but many consumers were unconvinced of the need for reform. Moreover, consumers were weakly organised and not well represented in the debate. Since the issue was not of major electoral significance, consumers’ voting power was not decisive, but consumer preferences were nevertheless a key factor, since they were crucial to the retail lobby’s own reassessment of its interests.

In the end, liberalisation came about as a result of a judicial action rather than the normal policy process. The regulation of opening hours was transferred to the Länder after the Federal Constitutional Court concluded that there was no case for federal regulation of the issue under the Basic Law. The interest-group coalitions opposed to reform were weaker in some of the Länder, and proponents of reform had already gained the upper hand in most of them. Even some of the more cautious Länder opted to open up, lest their retailers lose out in competition with their neighbours. The Land economics ministries were supported by a retail lobby that had once supported the law but that was increasingly in favour of reform. This shift in retailers’ perceived self-interest reflected changes in the structure of the retail sector, technological change and the emergence of new forms of retailing, as well as changes on the demand side in retail markets.

**The context for reform**

**Framework conditions for reform**

The German policy process is highly inclusive and consultative and is characterised by a multiplicity of actual or potential “veto points” that can present obstacles to reform. These include: the need to form multi-party coalitions, in order to command a majority in the lower house of parliament, the Bundestag; a federal system that involves extensive areas of joint federal-state jurisdiction and gives the Länder a very direct role in much national policy making via a fairly powerful upper house of the federal parliament, the Bundesrat, in which the government of the day often lacks a working majority; a strong role for judicial review; and well established corporatist procedures for interest intermediation involving tripartite negotiations among government, employers and trade unions. The heterogeneity of Germany’s major political parties means that intra-party disputes may further add to the obstacles to reform. Policy making has thus tended to be characterised by dense networks of negotiations among state institutions and some non-state actors, and substantial policy change generally takes time unless there is a strong consensus on the need for reform.

A centre-left coalition bringing together the Social Democratic Party (SPD) and the Greens was in power throughout the period under study. When first elected in 1998, the government was opposed to further reform of the Ladenschlussgesetz, but it was moving
towards some relaxation by the time it was re-elected in 2002. Nevertheless, the liberalisation of shopping hours was not a major election issue.

The economic context of the reform debate in the early 2000s was one of sluggish growth, high (and, from late 2001, rising) unemployment, and weak domestic consumption. The retail sector was growing more slowly than the rest of the economy, in part due to the weakness of consumption growth. Employment in retail was falling in the late 1990s by around 20-30 000 per year. This appears to have resulted largely from consolidation within the sector and restructuring in the eastern Länder, but it also reflected changes in business practices that create purely statistical effects. The sector’s lacklustre performance added to the urgency of the on-going debate over regulation of retail trade in general and of the Ladenschlussgesetz in particular. Advocates of liberalisation argued that reform would help stimulate consumption and retail-sector job-creation. Productivity performance would improve if consumption shifted to more efficient outlets and retailers had more freedom to optimise their hours and organisation of staff. The welfare gains arising from greater consumer choice were also emphasised. Opponents of reform predicted that it would lead to more job losses and a deterioration in the quality of retail employment.

**Actors and interests**

German consumers were by far the single biggest group affected by the Ladenschlussgesetz, and some polls showed broad and growing support for liberalisation. However, consumers as a group were not well organised and the consumer organisations that did exist tended to focus on issues like product safety and consumer protection against fraud; they did not support reform of opening hours. Since the liberalisation of retail hours was never a major election issue, consumers’ voting power was not decisive, and debates over the Ladenschlussgesetz were thus largely resolved via bargaining among state institutions and organised interests.

At governmental level, the Ministry of Economics and Technology and the Ministry of Labour and Social Affairs were the major players. While there were internal differences of opinion within the economics ministry, it was broadly in favour of further liberalisation. The labour ministry, by contrast, was opposed. Their divergent views of the issue stalled change for some time. After the 2002 elections, the two ministries were merged, a move that facilitated some progress on reform. Within the state administration, the federal dimension was also particularly relevant, as a number of Länder tended from a fairly early stage to push for more liberalisation than the federal government would concede. The Länder, in turn, were often responding to pressure from the municipalities, particularly the larger cities, which were initially inclined to push for ever wider exceptions to the Ladenschlussgesetz but in many cases came to favour general liberalisation instead. In the end, the determination of many sub-national governments to liberalise proved critical to the reform process.

Throughout the 1990s, both the independent German White-Collar Workers’ Union (DAG) and the service-sector unions affiliated to the German Federation of Trade Unions (DGB) were strongly opposed to liberalisation of the Ladenschlussgesetz. In late 2000, the DAG agreed to re-join the DGB in the context of the creation of a large new United Services Union, known as ver.di. Ver.di quickly took on the leading role in trade union opposition to reform of the Ladenschlussgesetz. On the employers’ side, the most prominent associations of retail-sector employers were the German Association of Branches and Self-Service Department Stores (BFS), the German Retail Federation
(HDE), and the Federal Association of Medium- and Large-scale Retail Enterprises (BAG). Their positions reflected the characteristics of their particular constituencies:

- The BFS primarily represented chains of supermarkets and super-stores and was thus a consistent proponent of liberalisation. In 2002, all but two of its members entered the HDE, a move that strengthened the pro-reform constituency within that organisation.

- Long dominated by small retailers who relied on the Ladenschlussgesetz for a measure of protection against competition from larger players, the HDE voted to support liberalisation only in 1998. Even then, conflicts of interest between its larger and smaller members meant that its position was sometimes ambivalent. The HDE remained opposed to Sunday opening, the issue that probably divided its members most sharply.

- The BAG, dominated by city-centre department stores, favoured a so-called “city-centre privilege” (City-Privileg), arguing for the liberalisation of hours for shops in town centres, while retaining restrictions for out-of-town shops. This was presented as a means of revitalising urban centres. The HDE opposed this, seeing it as clearly discriminatory.

Finally, both Roman Catholic and Protestant churches spoke up in defence of the Ladenschlussgesetz. They were primarily concerned to protect the requirement for Sunday closing, invoking article 139 of the Basic Law, which protects Sundays and official holidays as “days of rest from work and of spiritual improvement”. However, they also favoured restrictions during the rest of the week, citing the need to protect family life.

**Issue history**

Germany has traditionally had relatively strict regulations governing retail shop opening hours. The Ladenschlussgesetz was adopted in 1956. Its purposes included protecting retail employees from “anti-social” working hours and shielding small retailers from what was perceived to be the threat of unfair competition. For decades the retail lobby defended restrictions on opening hours. Until 1989, retail shops were permitted to open from 07h00 to 18h30 on weekdays and 07h00 to 14h00 on Saturdays. Although the first tentative step to liberalise retail opening was taken in 1989, and a further reform followed in 1996, Germany was still comparatively restrictive at the end of the 1990s (Table 19.1). Under the 1996 legislation, shops were permitted to open from 06h00 to 20h00 on weekdays and from 06h00 to 16h00 on Saturdays (18h00 during the four weeks before Christmas). Bakeries were allowed to open for three hours on Sundays. The 1996 amendments stipulated that an assessment of the impact of the new arrangements would be undertaken in 1999, as a prelude to possible further changes.

**Pressures for change**

By the late 1990s, changes in the nature of retailing were feeding pressures for reform. The consolidation of the retail sector was gradually strengthening the position of the large chains within the retail lobby, while the rise of new retail formats, such as shopping malls and inter-sectoral businesses, created new pressures for change and new problems in deciding how to apply opening-hours regulations. The growth of “factory outlets” led to a prolonged battle about how they were to be classified. Strictly speaking, they were part of industry, not the retail sector, but the retail lobby ultimately succeeded...
in having them subjected to the \textit{Ladenschlussgesetz}. There were also lengthy disputes about cash-and-carry outlets – formally “wholesale” distributors who sold “memberships” to private consumers. Running battles over definitions and exemptions highlighted the widespread demand from consumers for longer opening hours and the fact that the \textit{Ladenschlussgesetz} was based on a model of retailing that corresponded less and less to German realities. The expansion of mail-order businesses – Germany was one of the largest distance-selling markets in the world even prior to the rise of the internet – presented a further challenge to traditional retailers, as did the explosive growth of e-commerce from the end of the 1990s, and the regulation of conventional retail formats made it harder for them to respond to these emerging competitive threats.

Table 19.1. \textit{Shop opening hours in selected European countries, 1999}\textsuperscript{1}

<table>
<thead>
<tr>
<th>Country</th>
<th>Weekdays</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>6h-19.30h; Saturday -- 6h-17h</td>
<td>Exceptions for family business</td>
</tr>
<tr>
<td>Belgium</td>
<td>5h-20h; Friday -- 5h-21h</td>
<td>One rest day per week</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>No regulation</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0h-24h; Monday from 6h, Saturday -- 0h-17h</td>
<td>Closed</td>
</tr>
<tr>
<td>Finland</td>
<td>7h-21h; Saturday -- 7h-18h</td>
<td>12h-21h</td>
</tr>
<tr>
<td>France</td>
<td>7h-21h; Saturday -- 7h-18h</td>
<td>Closed</td>
</tr>
<tr>
<td>Germany</td>
<td>6h-20h; Saturday -- 6h-16h</td>
<td>Closed</td>
</tr>
<tr>
<td>Greece</td>
<td>6h-20h; summer -- 6h-21h</td>
<td>Closed</td>
</tr>
<tr>
<td>Hungary</td>
<td>0h-24h</td>
<td>As on weekdays</td>
</tr>
<tr>
<td>Ireland</td>
<td>7h-22, maximum -- 13 hours</td>
<td>Eight Sundays per year open</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>6h-20h; Once per week to 21h; Saturday -- 6h-18h</td>
<td>6h-13h</td>
</tr>
<tr>
<td>Netherlands</td>
<td>6h-22h</td>
<td>Twelve Sundays per year open</td>
</tr>
<tr>
<td>Portugal</td>
<td>6h-24h</td>
<td>8h-13h</td>
</tr>
<tr>
<td>Slovenia</td>
<td>7h-20h, Saturday -- 7h-13h</td>
<td>Regulated by regional authorities</td>
</tr>
<tr>
<td>Spain</td>
<td>0h-24h</td>
<td>As on weekdays</td>
</tr>
<tr>
<td>Sweden</td>
<td>7h-20h; Saturday -- 7h-13h</td>
<td>Regulated by regional authorities</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0h-24h</td>
<td>Canton regulation</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0h-24h</td>
<td>10h-16h</td>
</tr>
</tbody>
</table>

\textsuperscript{1}. In some countries, different rules applied to tourist centres.


Internationalisation affected actors’ interests in a number of ways. First, major international chains were increasingly entering the German market, and they were not content with the \textit{status quo}. Secondly, there was growing awareness among both consumers and retailers of how different things were abroad. Thirdly, there was growing interest in the development of “shopping tourism” among retailers in border areas and large cities. Retailers in many areas also faced growing competition from largely unregulated open-air markets, particularly in Berlin and around Germany’s eastern borders, as well as from retail outlets in neighbouring countries: by the late 1990s, eight of Germany’s nine neighbours had more liberal regulation of shopping hours than it did.\textsuperscript{15} As a result of these changes, even many small retailers, particularly in city centres and tourist areas, came to see opportunities in longer opening hours: what had largely been a conflict between large and small retailers increasingly pitted small retailers in suburban and rural areas against their competitors in the town centres.

Changing patterns of family life seem to have played a role in altering consumer preferences: rising female labour-force participation probably added to the pressure for change from the household sector, even as the \textit{Ladenschlussgesetz} served, at the margin,
to add to the disincentives to women’s employment. However, this issue did not figure prominently in public debates over the law. When the question of women in the workplace was debated in connection with the Ladenschlussgesetz, the focus was generally on women working in the retail sector rather than on women in the labour force as a whole. The great majority of workers in German retail were (and are) women, and the trade unions argued that longer opening hours would subject them to more “anti-social” hours. Many women’s organisations supported this view. The unions also feared that liberalisation would fuel the trend towards reliance on part-time work in the sector: by 2003, almost half of retail employees were part-time workers, and the share of part-timers in the sector was rising. Moreover, the vast majority of part-timers were in “mini-jobs” – that is, they were earning wages below the level at which they became liable for social security charges. While advocates of reform insisted that opportunities for part-time work might make it easier for many women to enter the labour force, ver.di and its allies argued that women who wanted full-time work in retail were increasingly being forced to settle for part-time positions.

Finally, the growing range of sometimes absurd anomalies thrown up by attempts to exploit the exemptions allowed by the law raised questions about both its equity and its necessity. Petrol station convenience stores, being exempt from the Ladenschlussgesetz under the provisions concerning “travel necessities”, grew larger and larger. Malls and other large stores organised “cultural” events to allow weekend opening, and the definition of exempt goods, like “travel necessities” and “tourist goods”, was stretched further and further, as was the definition of what constituted a “tourist area”. In some cases, municipal and Land authorities seemed to lend at least tacit support to such efforts to stretch the law, particularly in Berlin and the eastern Länder. Indeed, German unification itself seems to have contributed to the pressure for change: female labour-force participation rates were far higher in the Länder of the former German Democratic Republic, religious attitudes there were different, and the speed with which the larger western chains entered the East after 1989 ensured that the pro-liberalisation elements of the retail lobby were well represented there. Moreover, small entrepreneurs in the East, facing very difficult conditions (as well as competition from open-air markets) seem to have been very sensitive to any regulations that got in the way of their doing business. Under greater competitive pressure than their western counterparts, they probably lacked the kind of local market power that enabled small retailers in the West to profit from the Ladenschlussgesetz. Finally, the trade unions, including ver.di, seem to have been weaker in the eastern Länder. By the late 1990s, therefore, a number of eastern Länder were becoming increasingly brazen in allowing local business to circumvent the Ladenschlussgesetz.

The economics over opening-hours regulation

One reason that discussion of the Ladenschlussgesetz proved so inconclusive for so long was that there was no firm consensus concerning the rationale for, or the consequences of, regulating retail opening hours.

- **Effect on prices.** Supporters of the law argued that restrictions on hours helped hold down prices, since competition among retailers would ensure that the cost savings made possible by shorter hours were passed on to consumers. Critics of the Ladenschlussgesetz rejected this view, arguing that it enabled retailers with a degree of local market power (particularly in small towns and suburbs) to extract monopoly rents: lack of competition meant that such retailers did not have to lower prices to reflect lower costs and restricted hours meant that
time-constrained consumers were more likely to shop locally. The law also offered rents to “exempted merchants” (e.g. petrol station convenience stores), who could charge higher prices because they were allowed to sell after hours. Advocates of reform argued that, if liberalisation led to increased competition and consolidation in the sector, the result would be greater convenience and lower prices for consumers.

- **Differences among consumers.** Restricted opening hours discriminated against time-constrained consumers, who attached a high value to being able to shop when they wished. The law may have favoured “location-sensitive” consumers, who preferred to shop locally and who would have lost out if longer hours in some shops forced others, in preferred locations, to close. This line of argument suggests that the law might have protected groups like pensioners and students, as well as single-breadwinner households, at the expense of single workers and two-income households.

- **Differences among retailers.** The existing structure of the retail sector partly reflected the effects of past regulation: if reform were to change consumer preferences about where and when to shop (as surveys suggested it would), past investments would earn lower returns in some locations and higher returns in others. Reform was expected to hurt retailers whose scale of operations did not warrant hiring additional staff to work evening and weekend shifts, as well as those in locations that were profitable chiefly because they were convenient for time-constrained consumers. Larger retailers were expected to draw enough after-hours customers to justify extended opening.

- **Differences among workers.** Many retail workers (particularly women) had strong preferences for daytime hours on weekdays. At the same time, the law tended to discriminate against those, such as students, who might want part-time evening or weekend work. For full-time workers, then, reform was likely to be disruptive, while it might create new opportunities for those in (or seeking) part-time work.

- **Arguments over externalities.** A final set of arguments concerned potential externalities, such as the value of coordinated leisure – the main reasons that Sunday restrictions were strongly defended – and the value of maintaining vibrant city centres that might be “hollowed out” if consolidation in the retail sector shifted business from urban shops to out-of-town malls and superstores.

While the debate was unfolding in Germany, a growing body of empirical work began to cast doubt on many of the assumptions underlying the arguments in favour of the *Ladenschlussgesetz.* Evidence from OECD countries suggested that further liberalisation would increase employment as well as consumer choice. Moreover, there was little evidence to suggest that unregulated hours necessarily led to markets dominated exclusively by mega-stores. Empirical research in liberalised markets suggested that, except in 24-hour convenience stores, other criteria appeared to dwarf opening hours as a factor in determining prices.

**Policy process: formulation**

In late 1999, the economics and labour ministries, acting on a request made by the Bundestag in connection with the 1996 legislation, commissioned separate studies of the impact of the 1996 changes to opening hours. At the same time, a number of Länder began stepping up their own demands for the right to relax restrictions on opening hours.
in their jurisdictions. In August 1999, stores in three major eastern cities opened on Sundays in open defiance of the law. Taken together, the studies and the evident restiveness of some of the Länder led to renewed debate on the future of the Ladenschlussgesetz.

The studies came to conflicting conclusions.

- The survey commissioned by the economics ministry from the Ifo Institute for Economic Research found that around half of all customers were making use of the additional hours permitted under the 1996 legislation, though only a minority reported doing so weekly (most did so “now and then”). Around 45% of customers favoured Sunday opening, with a similar proportion supporting outright repeal of the Ladenschlussgesetz. Only a minority of retail shops were taking advantage of the extended hours, though many more had done so immediately after the reform, as retailers experimented to see whether longer hours were profitable or not. Extension of opening hours was closely linked to size and location, with larger stores overwhelmingly choosing to stay open longer on weekdays and Saturdays, and city-centre stores much more active than those in suburban or rural settings. Only about 25% of shops wanted the repeal of all legal restrictions on weekday hours, and 60% favoured retention of the general prohibition on Sunday opening.

- The labour ministry-commissioned study, prepared by the Dortmund Social Research Centre (SFS), focused on retail-sector employees. It drew attention to evidence of falling employment in retail since 1996, but it concluded that liberalisation of hours had neither caused the employment decline nor contributed to stabilising employment in the sector. Most shops that had adopted longer hours had not increased employment but had instead redeployed staff to cover the additional hours; indeed, more shops had cut employment while extending hours than had increased both. The SFS also highlighted survey results showing that employee assessments of the effect of longer opening hours on their quality of life were largely negative, as shift work and frequently changing schedules made it harder to balance work and family life. Only 4% of retail employees wanted to work after 18h00 on Saturdays and 61% favoured a return to 14h00 Saturday closing, despite the fact that retail workers had been paid a 20% premium for working 14h00-16h00 on Saturdays since 1996. Almost half favoured a return to pre-1996 hours on weekdays as well.

The Ifo study also looked at competition, concluding that liberalisation would accelerate structural change in the retail sector, but public debates over the law were focused largely on issues other than competition, including occupational health and safety, consumer convenience and the possible externalities associated with regulation, such as co-ordinated leisure time. Reactions to the Ifo and SFS studies largely reflected actors’ ex ante positions:

- The HDE and other retailers’ associations agreed with Ifo that further liberalisation would be in the interests of consumers, while acknowledging that longer hours did not necessarily lead to higher retail turnover and employment. While there was still some conflict within the retail lobby, only a minority of small retailers continued to oppose any reform. There was growing support for a substantial Monday-Saturday liberalisation, leaving mandatory closing on Sundays and bank holidays in place. This made it easier to avoid opposition from...
exempted merchants, who profited from the restrictions that the Ladenschlussgesetz imposed on other retailers.\textsuperscript{37} For many small exempt retailers, particularly in the food sector, Sundays were the critical concern. Other exempted groups, such as petrol station convenience stores, tended to be more focused on late evenings, but these were in many cases linked to large retail chains anyway, so they were not disposed to resist reform.

- The service-sector trade unions argued that the SFS study confirmed their opposition to liberalisation, stressing that the Ladenschlussgesetz was needed to protect workers from excessive and/or unsociable hours. Advocates of reform in the economics ministry argued that such issues were regulated by the Working Time Act and existing collective agreements in the retail sector, and that labour-market regulation was a more efficient instrument than product-market regulation for achieving labour-market policy objectives. However, the labour ministry supported the unions’ position, insisting that retail workers’ health and safety required the particular protections of the Ladenschlussgesetz and that it made enforcement of working-time legislation in the sector much easier.\textsuperscript{38} In an effort to broaden support for their position, the unions emphasised other issues, such as the need to protect small retailers, the threat to city centres posed by large out-of-town malls and stores, the importance of protecting Sundays, and the need to maintain high standards of service in retail outlets.

In a report to the Bundestag, the government concluded that the two studies taken together did not point to the need for immediate action and that further discussion with all the parties concerned would be desirable. While the two federal ministries continued to defend their respective positions, the Länder, most of which favoured some reform of the law, continued to push for change. In the summer of 2000, Land economics ministries approved and submitted to the Bundesrat a proposal to allow opening to 22h00 on weekdays and 20h00 on Saturdays, as well as permitting operation on four Sundays or public holidays per year (not in December) from 12h00 to 18h00. The Länder position was influenced in part by the experience of the “Expo 2000” world’s fair in Hannover, where extended Saturday opening had proved popular with retailers, visitors and local consumers. It may also have reflected the fact that ver.di was less influential in many Länder than it was at federal level.\textsuperscript{39} While the Minister of Economics expressed support for a closing time of 22h00 on weekdays (but no change on Saturdays or Sundays), the federal government did not move on the issue and the proposal was not followed up by the Bundesrat. The Bundesrat initiative nevertheless underscored the gap in perceptions of the issue that was opening up between the federal government and many of the Länder.

This inconclusive debate continued until after the 2002 elections, when the labour ministry and the economics ministry were merged, thereby making it possible to break the impasse between the two. The compromise proposed by the new Federal Ministry of Economics and Labour at the end of 2002 was limited in scope: weekday hours would remain unchanged but Saturday opening would be extended to 20h00, so that the Monday-Saturday rule would be uniform (06h00-20h00). In addition, barbers and hairdressers were to be exempt from the Shop Closing Hours Act,\textsuperscript{40} and the requirement that shops permitted to open on Sundays must close no later than 14h00 on the preceding Saturday was to be repealed. The bill did not address what was potentially the most controversial issue – Sundays and bank holidays. It focused instead on that part of the week that both surveys and behavioural indicators suggested both consumers and retailers most wished to see liberalised – Saturdays.
Policy process: adoption

The revision of the Ladenschlussgesetz proposed in December 2002 reached the statute books in March 2003 and entered into effect from June. Since the issue had been actively debated for years, there was little need for extensive consultations on the bill, nor was there any need for a major public relations campaign, since most voters appeared to favour at least some liberalisation. Yet even the limited further reform proposed by the ministry was resisted by the trade unions, the churches and those elements of the retail sector that feared the heightened competitive pressure that liberalisation would entail. The unions presented their case in terms of working-time regulation, the protection of small entrepreneurs and the need to prevent any “hollowing out” of German town centres. However, the reform had more direct implications for wages and union power as well:

- Employers gave notice that they would no longer be willing to offer bonus pay for “anti-social hours” to workers employed on Saturday afternoons, arguing that there was no longer a case for treating Saturdays differently from weekdays. The unions flatly rejected this. Since the Collective Agreements Act held that existing provisions must remain in force until a new agreement was reached, this meant that, once the new law entered into force, the entitlement to bonus pay would stand for 14h00 to 16h00 but not for 16h00 to 20h00. However, by 2003, a growing number of retailers were not covered by collective agreements, having either left the retailers’ associations or taken out so-called “non-coverage membership”.41

- The unions were extremely concerned about the growth of part-time work. If reform reduced shop owners’ reliance on full-time employees, the unions’ numerical strength and bargaining power would decline, not least because part-timers were harder to organise. Yet if part-time workers were not hired, then the extension of opening hours would be even more disruptive for full-time employees. Moreover, there was some evidence that wage scales for full-time employees in retail might exceed what a competitive market would have yielded.42 If longer opening hours led to more competition, these arrangements would come under pressure.

Though unable to halt the progress of the reform through the Bundestag, organised labour did win one important concession. At the request of the SPD and Green fractions, parliament added a provision to the draft bill guaranteeing workers in retail at least one Saturday a month off work.

Policy process: implementation

Implementation of the revised law was relatively straightforward. Since it involved the relaxation of a previously existing restriction, it made few demands on the public administration. It did, however, lead to an extended round of bargaining over working time and pay for the additional hours. In the end, ver.di, the BAG and the HDE agreed on the extension of the 20% supplement for late hours on Saturdays, although the extent of coverage varied across the different Länder. In some, the bonus applied from 14h00, while in others, it was paid only from 14h30 or 15h00. Under collective agreements concluded in 2006, the period covered by the supplement was further reduced, to 18h30-20h00 in most regions. The additional 20% was granted in the form of a
“working-time bonus”, effectively shortening the average working week of those who worked such hours.\textsuperscript{43}

The 2003 reform did not, in any case, last long. In June 2004, the Federal Constitutional Court ruled on a challenge to the constitutionality of the restrictions contained in the \textit{Ladenschlussgesetz} brought by the Berlin-based department store Galeria Kaufhof. The court did not find that the restrictions themselves violated the constitution,\textsuperscript{44} but it concluded that the \textit{Länder} should be given the authority to regulate opening hours. In areas of joint federal-state jurisdiction, the Basic Law declares that “the Federation shall have the right to legislate… if and to the extent that the establishment of equal living conditions throughout the federal territory or the maintenance of legal or economic unity renders federal regulation necessary in the national interest.” The court observed that, in view of the vast numbers of exemptions and exceptions already enshrined in the law, one could hardly argue that uniform national-level regulation was required. While it did not immediately annul the \textit{Ladenschlussgesetz}, the ruling did mean that it could no longer be revised at federal level. Further reform would require sub-national legislation.

Many \textit{Land} governments were eager to legislate quickly after the ruling, but the federal government did not wish to surrender competence to them immediately and opted instead to fold the issue into a broader federal reform then being prepared. Thus, it was not until August 2006 that the federal parliament adopted legislation shifting legal competence for the regulation of opening hours to the \textit{Länder}. Within months, all but Bavaria had liberalised to some extent, with some allowing 24-hour Monday-Saturday shopping.\textsuperscript{45} Proponents of reform had gained the upper hand in most \textit{Länder} and even some of the more cautious \textit{Länder} were pushed to open up, lest their retailers lose out in competition with rivals in neighbouring jurisdictions. Two issues remained in contention, however:

- Sundays and holidays continued to enjoy constitutional protection, raising questions about how far governments could go in liberalising Sunday/holiday opening. There were already numerous exemptions and exceptions to the rule on Sunday trading, so it could not be argued that a total ban was needed in order to meet the requirements of the constitution. However, it was not clear how many open Sundays were constitutionally acceptable. The city of Berlin’s attempt to allow up to 10 open Sundays a year provoked a constitutional challenge.

- When the \textit{Länder} took over responsibility for regulating opening hours, the employers’ organisations cancelled the existing framework collective agreements, indicating that they intended to scrap the bonus payments for late hours. The unions, by contrast, resisted any change, defending not only the 20\% extra paid for late Saturdays but insisting on retention of the 55\% bonus paid for work carried out between 20\text{o’}clock and 06\text{o’}clock – a rate previously applied only in exceptional situations (for example, when staff might work overnight to rescue stock from a fire or flood).

There were still no new collective agreements in place in early 2008. The stand-off may actually have provided some boost to \textit{ver.di}, whose membership grew significantly after the employers declared their desire to negotiate new agreements. \textit{Ver.di} was quick to remind workers that those who joined the union before the old collective agreements expired would continue to be covered by the bonus payments until new arrangements were agreed. However, this did not extend to the large number of retail workers in
enterprises not covered by collective agreements or employed in very small firms, where some provisions of collective agreements are not applied.\textsuperscript{46}

Conclusion

The difficulty of relaxing the *Ladenschlussgesetz* at federal level highlights some of the barriers to reform built into Germany’s policy process. There were important divisions between the labour ministry, which viewed the law as an important instrument for pursuing social policy objectives, and the economics ministry, which was pressing for liberalisation, as well as within the governing parties. These divisions prevented the federal government from moving decisively on the issue for several years. Even a relatively modest reform was possible only after the two ministries were merged and the government conceded the guarantee of one Saturday a month off work for retail workers. Some *Länder* were moving towards a more pro-reform position and even tried to introduce reform proposals through the *Bundesrat*.

Germany’s poor labour-market performance in the early 2000s probably reinforced opposition to reform from organised labour. Although proponents argued for the employment benefits of reform, the unions’ concern about the growth of part-time work and the job losses they feared as a result of liberalisation took on added force against the backdrop of high and rising unemployment, as well as falling employment in retail. The over-selling of the 1996 reform may have played a role here: the *ex post* studies of the 1996 legislation failed to find evidence that it had contributed to the growth of retail turnover and retail employment, as proponents of reform argued it would. While the decline in retail employment seems to have been a result of larger structural shifts, these findings put advocates of reform on the defensive for a time, despite the evidence that the reform had been good for consumers and that both consumers and retailers favoured further liberalisation. While consumers were by far the largest constituency affected, they were not well organised and the issue was not of sufficient electoral salience for their voting strength to decide the issue. Consumer preferences were, however, critical to the shift in the retail sector’s perception of its own self-interest.

The evolution of the retail lobby’s position forms a remarkable part of the story. While a significant body of small retailers, especially in suburban and rural areas, continued to oppose reform, increasing concentration in the sector, along with technological and economic change, led the retail lobby as a whole to adopt a pro-reform position. The original *Ladenschlussgesetz* can in some respects be viewed as a cartel arrangement.\textsuperscript{47} Since the retail sector was too large and diverse to maintain a cartel on its own, it relied on a federal statute to enforce restrictions on retail “supply”. On this view, the cartel broke down as the balance of power within the sector shifted towards agents who believed they were better off competing than sharing cartel-generated rents with their rivals. Yet although the *Ladenschlussgesetz* had been adopted largely in response to retailers’ demands for protection from competition, the change in their position was insufficient to bring about its reversal, not least because the unions opposed any change. For several years, then, retailers were caught in something of an institutional trap, unable to escape the constraints of a regulation for which they had once lobbied. The unions may have been resisting the loss of that portion of the rents which had effectively been captured by labour, but their major concern seems to have been to check the growth of part-time work.

Yet labour was not alone in resisting reform. One reason reform proved so difficult was that the coalition resisting change brought together a number of different players who
were motivated by different agendas. Many smaller retailers, especially in suburban and rural areas, continued to oppose what was perceived to be the threat of unfair competition; the trade unions, which were relatively weak in the retail sector, wished to limit pressures for longer working hours while checking the further growth of part-time employment, and the labour ministry, too, saw it as a means of regulating working time; churches cited Sunday trading and the protection of family life; and so on. These very disparate agendas intersected where the Ladenschlussgesetz was concerned, ensuring that each group could rely on allies in its fight against change.

Perhaps the most striking aspect of the Ladenschlussgesetz story is the contrast between the very limited, incremental steps taken at federal level and the speed with which the Länder liberalised once the issue was passed to them. In many Länder, the unions were relatively weak and advocates of reform had been in the ascendant for some years anyway. Border regions, eastern Länder and Länder dominated by large cities all tended to favour reform. Inter-jurisdictional competition also appears to have played an important role – liberalisation in Berlin, for example, confronted the surrounding state of Brandenburg with a choice between reforming the law and seeing its citizens doing still more of their shopping in Berlin.48

Notes

1. The Federal Republic of Germany has never had a one-party federal government.

2. There are, it should be observed, significant exceptions to this rule, including the labour-market reforms of the early 2000s, which are addressed in the case study on the Hartz reforms.

3. Since 1990, the outsourcing of things like cleaning services has led to the transfer of many jobs out of the retail sector for statistical purposes. The German Retail Federation believes that, if this factor is taken into account, the number of employees in retail has changed little, if at all, over the period.


5. See the survey data cited in, e.g., Scheele (1999); Financial Times, 6 June 2004; and The Business, 1 August 2004. Kosfeld (2002) observes that pensioners were the only large group of consumers who were substantially opposed (24%) or indifferent (52%) to reform. However, the survey conducted for the IFO study discussed below found that only 45% of respondents favoured Sunday opening and the abolition of statutory opening hours on weekdays.

6. Indeed, many of the 40 or so organisations united by the Federation of German Consumer Organisations (VZBV) had traditionally maintained close ties to the trade unions.

7. Officials responsible for the SME sector tended to favour restrictions on opening hours and those concerned with competition to oppose them.
8. With around 3 million members at the moment of its formation, ver.di was thought to be the largest single trade union in the democratic world; see Sadowski et al. (2001); and Annesley (2006).

9. Altogether, there were seven such associations, including one focused on craftsmen and small self-employed retailers (the ZGV), which tended to adopt a very conservative position.

10. Small retailers enjoy other protections, owing largely to zoning regulations and restrictions on activities like promotional pricing. In some cases, entry of large retail outlets into an area can be blocked on account of their potential impact on existing retailers nearby; see OECD (2006:143-45).

11. Saturday opening was extended to 18h00 on the first Saturday of the month and the four Saturdays before Christmas.

12. The 1989 amendment allowed for late opening, until 20h30, on Thursdays.

13. This concentration process was particularly marked in food-retail, a sub-sector once dominated largely by small shops. The market share of Germany’s ten leading food retailers rose from an estimated 45% in 1990 to around 84% in 2002. The number of small food stores and supermarkets (under 400 square meters) fell by 42% over the period, while the number of hypermarkets increased by 242%. See Dribbusch (2004); and Boeri et al. (2007:67).

14. For example, food retailers could by law devote 10% of store space to non-food goods, but this limit was relaxed in city centres, with the result that more and more cafés and food outlets were selling other goods as well.

15. It is impossible to be categorical about the ninth, Switzerland, since opening hours were regulated at canton level.


17. Over 70% of retail workers earning above EUR 400 per month (i.e. those subject to social insurance contributions) in 2003, according to the Federal Statistical Office. Most mini-jobs in retail were also filled by women.

18. Implicit in much of this argumentation was the assumption that women were responsible for childcare and basic household tasks and thus needed to be protected from long hours.

19. Dribbusch (2004), based on data from ver.di and the Federal Statistical Office. However, there was little evidence that the 1996 relaxation of the Ladenschlussgesetz triggered an acceleration of the shift towards part-time work; see Boeri et al. (2007:70).


21. Title inflation was also employed to support later closing times: the Dussmann book, computer and music shop in Berlin gave one-quarter of its 100-strong staff senior management titles, allowing them to serve customers until 22h00, and in 1999, it put in a small swimming pool and palms in an effort to become a “spa resort” and thus open on Sundays.

22. Even the definition of “railway station” became the subject of litigation, since shops in railway stations were not subject to the opening hours law.

23. See Coles (1997), especially Table 3.
24. As one East German entrepreneur put it, “West German businesses very often have a cushion so that they can ride out the bad times. In the east, where unemployment is so high, we have nothing to fall back on when times are bad, so the pressure to change the law is greater.” International Herald Tribune, 29 September 2000.

25. See Annesley (2006); and Sadowski et al. (2001).

26. The overview presented in this box draws on Kirchner and Painter (2000); Thum and Weichenrieder (1997); Burda (2000); Burda and Weil (2005); and Wenzel (2007).

27. This is not to suggest that they raised prices after hours: rather, their earnings from after-hours business were sufficient to allow them to charge higher prices at all times, even if this implied a loss of potential business during ordinary daytime hours.

28. Some analyses also pointed to problems for small family businesses, which account for a large share of German shops: extended hours were expected to make it more difficult for them to coordinate leisure and family life without incurring high costs from hiring non-family members.


30. Berlin also petitioned the federal authorities for permission to allow shop opening to 22h00. For more on the pressure from the eastern Länder, see Scheele (1999).

31. For overviews of the studies in question, see Kosfeld (2002) and Scheele (1999).

32. Other surveys also found broad public support for relaxation of the law.

33. 33% on weekdays and 29% on Saturdays.

34. Kosfeld (2002) cites data from several studies showing that shops in residential areas were the most likely to have found longer hours unprofitable – the number of stores in residential areas with extended hours fell by around 30% from its post-reform peak. Likewise, some 32% of stores with five employees or fewer experimented with longer hours and then reverted to pre-reform patterns. By contrast, some 96% of stores with over 100 employees opted for longer hours, a figure that remained more or less unchanged: one survey found 100% of hypermarkets, 97% of department stores and 67% of larger specialty stores making use of longer opening hours.

35. The study found declines in full-time employment and in part-time work subject to social security charges, but growth in the number of people working in part-time jobs whose earnings and hours fell below the threshold for social security charges. The number of people working in retail was actually rising steadily but the average number of hours worked was falling; see Boeri et al. (2007:70-71).

36. Scholarly analyses, too, tended to focus on possible externalities and market failures rather than competition; see, for examples, Thum and Weichenrieder (1997), Burda (2000) and Burda and Weil (2005). See Wenzel (2007), however, for a more recent analysis with a focus on competition issues.

37. It is significant that in 1998, the distance sellers had voted against liberalisation at the HDE congress, knowing full well that reform would benefit their conventional rivals.

38. Shops testing the limits of the Ladenschlussgesetz were indeed sometimes fined under the Working Time Act. See, for example, the case of the Berlin department store Galeria Kaufhof in the International Herald Tribune, 5 August 1999.
39. While some Länder controlled by the SPD and the Greens abstained at the meeting of state finance ministers in Berlin, the Länder did not split along straight party lines – the so-called “retail rebels” included some ministers from SPD-governed Länder, including the SPD Economics Minister of Chancellor Gerhard Schröder’s home state of Lower Saxony.

40. This exemption was not arbitrarily decided: in 2000, a Thuringian hairdresser had become a cause célèbre after she was arrested and imprisoned for cutting hair after hours in order to raise funds for a children’s cancer clinic.

41. Dribbusch (2003); Behrens (2002).

42. Kirchner and Painter (2000).

43. Since almost 30% of retail employees were part-timers on so-called “mini-jobs” that paid less than the EUR 400 threshold for social security payments, employers would have found it cheaper to opt for the shorter work week, and many part-timers also preferred shorter working time to bonus pay that would be subject to social charges.

44. It split 4-4 on the issue, and the tied vote meant that the measures in question were not judged to be in violation of the Basic Law.

45. Bavaria opted to retain the existing federal regulations for the time being and to study the experiences of other Länder before proceeding with reform.

46. In 2003, ver.di estimated that around 60-65% of retail employees in western Germany were covered by collective agreements, as against 40-50% in eastern Länder. However, these are rough estimates, as disaggregated data on bargaining coverage in retail are not available; the only available data cover the larger sector “retail and wholesale trade and repair services”. See Dribbusch (2004).

47. The problem is discussed in these terms by Kirchner and Painter (2000).

48. Länder bordering Hamburg and Bremen faced a similar dilemma.
Bibliography


Summary and main conclusions

Around 95% of Dutch rents are regulated, accounting for almost 44% of the housing market. In 2004, the government proposed reducing the share of regulated rents in stages to 75% by 2010, while allowing faster rent rises in the regulated rental sector, in an effort to ease the shortage of rental accommodation on the market. At the same time, Dutch housing associations were to make additional payments to the budget to offset the increasing public expenditure on rent subsidies to low-income households following the reform. The reform proposals also included measures to reduce the regulatory and other barriers to investment in new rented accommodation. Although the rent bill eventually passed the lower house of the States-General, it was still before the upper chamber at the time of the November 2006 general election. The coalition agreement that defined the new government programme after the election stipulated that it be withdrawn and that rent increases be limited to CPI inflation rather than the previous formula of “CPI plus”. Three years’ debate over partial liberalisation thus ended with a shift towards tighter regulation. Among the factors contributing to this outcome were:

- **Lack of consensus on the urgency or need for reform.** While there was widespread awareness of distortions in both owner-occupied and rental housing markets, there was no sense that the situation was intolerable or unsustainable. Opponents of reform argued that rent regulation generated long-term social benefits by preventing the social segregation that would result from liberalisation.

- **Lack of a broad constituency of potential winners.** Apart from landlords and developers, there were few clearly identifiable potential winners: landlords were a small constituency, while developers were more interested in building homes for sale. The status quo did impose costs on those whose access to affordable housing was blocked by insiders enjoying low rents, but housing shortages were not severe enough for outsiders to try to change the system; in general, they wanted to find rent-controlled accommodation themselves rather than to relax rent regulation. The absence of a strong pro-reform constituency, when combined with the lack of any urgent pressure for action, undermined the commitment to reform of even those political parties backing it.

- **Increasing complexity.** The reform proposals became steadily more complicated as the government tried to accommodate the demands of the various stakeholder interests involved. This made the reform more cumbersome and less likely to achieve its aims, and it also made it much harder to communicate its purpose to the electorate. The government was never able to define the endpoint of the
reform clearly, and its opponents were thus able to argue that it was the first step towards total deregulation of the rental market.

- **The need to compromise acquired rights.** There was widespread support for the idea that the new rules should apply to properties only after a change of tenant – i.e., that the acquired rights of those occupying dwellings that were to be liberalised should be respected. However, this would have postponed the positive effects of reform for decades while exacerbating, rather than mitigating, the barriers to mobility created by the existing arrangements.

- **Fear of a limited supply response.** There was much scepticism about the likely effectiveness of the government’s efforts to ensure that supply on the rental market would respond to higher rents. Constraints on supply were largely the product of sub-national governments’ planning policies, and many provincial and municipal governments did not want new rental stock.

- **Partial focus.** The trade unions, the tenants’ union and the parties of the left objected to any attempt to reform the rental market without tackling the owner-occupied market, which was also distorted by various government policies. This, however, was unacceptable to the parties of the governing centre-right coalition, whose constituents benefited from such policies as generous mortgage-interest tax relief and very low imputed rent income.

### The context for reform

**Framework conditions for reform**

The rent reform proposal was initiated under the centre-right “Balkenende II” cabinet, comprising the centrist Christian-Democratic Appeal (CDA), the economically liberal but socially conservative People’s Party for Freedom and Democracy (VVD) and the “progressive liberal” Democrats 66 (D66), which governed from 2003 until 2006. The government that emerged following the 2006 elections was significantly to the left of Balkenende II, bringing together the CDA, the Labour Party (PvdA) and the “social Christian” Christian Union, which combined social conservatism with centre-left positions on economic and environmental issues.

Two features of Dutch policy making significantly affected the shape of the rent deregulation legislation. First, the policy process in the Netherlands is highly consultative and inclusive; Dutch society is highly organised and interest groups are important actors in policy making. Key social stakeholders were thus deeply involved in discussions with the Ministry of Housing, Spatial Planning and the Environment (VROM) before and after publication of the reform proposals and their input led to important modifications. Secondly, Dutch ministries operate under medium-term spending caps fixed in the government’s coalition agreement and must offset unexpected financial setbacks on the spending side from their own budgets; this rule put VROM under considerable pressure as expenditure on housing allowances (rent subsidies to low-income households) rose during the economic slowdown of the early 2000s.

During the period under study, the Dutch economy was recovering from the sharp slowdown of 2002-03, which had contributed to a significant deterioration in the fiscal position. The general government balance moved from a surplus of 1.5% of GDP in 2000 to a deficit of 3.2% of GDP three years later; this was not merely the result of the
economic conjuncture, as the cyclically adjusted balance deteriorated by around 2.2% of GDP over the period.5

**Actors and interests**

Although VROM discussed the reform proposals with around six other ministries, the two most deeply involved were the VROM itself and the Ministry of Finance. The housing ministry’s primary objective was to address the increasing shortage of rental housing on the market, by relaxing rent regulation while simultaneously reducing constraints on the supply side; however, it was important for VROM to open up the market in a way that did not trigger an explosion in expenditure on housing allowances. Although housing allowances accounted for only 0.8% of general government spending, they consumed over half the VROM budget, and unexpectedly high levels of housing allowance expenditure were already creating serious problems for the ministry, since it had to keep spending under its medium-term cap. During the early 2000s, VROM had been forced to reduce the generosity of housing allowances at a time of rising unemployment, and continued increases in housing allowance expenditure threatened its ability to finance its urban renewal policies. The finance ministry was also concerned about rising expenditure on housing allowances, which were an open-ended entitlement – the government was obliged to pay them to qualifying households, even if this entailed expenditure in excess of what had been forecast.6 The other important state actors involved were municipal governments: some of them simply opposed deregulation, fearing the impact of higher rents, while others were more interested in determining themselves which parts of their local markets might be deregulated and in controlling new supply.

The biggest players on the Dutch housing market are the housing associations (HAs). Some 36% of all housing (77% of the rental stock) consists of social housing – a very large share by international standards.7 It is owned and administered not by central or local government but by the HAs, which are autonomous, not-for-profit institutions. The HAs have been financially independent of the state since 1995, and they build and sell housing, as well as operating on the rental market, which is their main occupation. The HAs “own themselves” (they have no shareholders), but use of their money is regulated by the state and their surpluses are mainly invested in urban renewal and new housing stock. The authorities aim to ensure that the HAs fulfil their primary responsibility of providing social housing but do not prevent them from undertaking commercial activities to support their social function. In 2003, their total rental income amounted to over EUR 10 billion and their reported net income reached EUR 725 million.8

The status of the HAs is controversial. The government has at times been concerned that, as “hybrid organisations”, combining social and commercial activities, they might create distortions in some segments of the market and that market risks might have a negative impact on their performance of their public duties.9 Questions have also been raised about implicit and explicit public subsidies for the HAs, which their private competitors argue are incompatible with European competition law.10 Questions about the increasingly anomalous position of the HAs under EU law thus added a European dimension to the motives for reform. The HAs’ umbrella organisation, Aedes, has long argued for gradual deregulation of rents, in order to increase investment in the social housing sector, but it favours retention of rent control for housing allowance recipients, who are unable to cover the cost of even their current regulated rents.11 However, Aedes insisted during the early 2000s that planning restrictions rather than rent levels were the main obstacle to increased investment.
There are two main organisations representing private landlords. The Association of Investors and Property Owners (VB) represents roughly 45,000 small landlords, accounting for around 15% of rented dwellings. Roughly 97% of VB rental properties fall under the rent regulations, and a large share are at the low end of the market. The Association of Institutional Property Investors in the Netherlands (IVBN) represents private commercial investors in the sector. Its 30 or so members own around 4.7% of the dwellings on the rental market. About 43% of these are in the unregulated sector. The rest tend to be at the upper end of the regulated market (rents of EUR 500 per month or more). IVBN members' holdings were falling in the early 2000s, as they sold properties when dwellings became available, and most of their new construction was for sale; a small proportion was for rental in the free sector. VB members, too, were selling properties and leaving the rental sector, so the private share of the total rental market was in decline. Though not in conflict, the IVBN and the VB approached rent reform with different interests. Both wanted more freedom to raise rents, but the IVBN was chiefly concerned with removing as much of its stock from the regulated sector as possible and with relaxing constraints on new construction. New construction opportunities were not an issue for VB members, who generally were not in a position to build. Since their holdings were largely at the lower end of the market, they were more concerned with rent regulation and the reform of housing allowances.

The Dutch Union of Tenants (Nederlandse Woonbond) brings together 700 local tenants' organisations, representing an estimated 1.2m tenants. Its chief priority was, not surprisingly, to hold down rents and press for other solutions to the shortage of rental accommodation, which the Woonbond saw as primarily a product of the elimination of state subsidies for new construction in 1995. This emphasis on supply enabled the Woonbond to claim that it represented the interests of both insiders (sitting tenants) and outsiders (those seeking accommodation) on the rental market. The Woonbond claimed that the households most likely to suffer adversely from the reform were those on lower-middle incomes: around one-third of tenants were estimated to fall above the ceiling for housing allowance but below the average household income for the Netherlands. One estimate suggested that, if proposals for faster increases in regulated rents were adopted, this group might see its expenditure on rent rise from an average of 26.5% of disposable income in 2004 to just over 30% in 2009.

**Issue history**

Rent regulation and rent subsidies in the Netherlands

The basic elements of rent policy (Box 20.1), including the valuation system, have been in place since the 1970s. The first major innovation after that time was the creation after 1989 of a small free rental sector, where properties with a rental value that exceeds a certain “liberalisation threshold” may be let at whatever rents the market will bear. In the early 2000s, there were somewhat over 200,000 properties on this market, and the number was declining. The second major policy change towards the rental sector was the grossing and balancing operation of the mid-1990s. Until 1 January 1995, HAs received annual subsidies for their rented housing stock, to offset deficits in their operations; these subsidies enabled the HAs to invest in new supply. At the start of 1995, the HAs repaid all their outstanding government loans, substituting private loans where necessary. These private loans were guaranteed by the Guarantee Fund for Social Housing, with 50% backing from the central government and 50% from the municipalities. On the same date, the HAs received a lump sum corresponding to the
discounted value of all future state-approved property subsidies; these pledges covered a period of 50 years, based on estimates of long-term trends in rents, interest rates and inflation. All further renovation subsidies, new building subsidies and exploitation subsidies were eliminated, making the HAs financially independent of central government. The case for providing such subsidies had in any case been undercut by above-inflation rent increases over a period of years prior to that time. Since rent control remained but bricks-and-mortar subsidies disappeared, the burden of subsidy effectively shifted from the state to the HAs, except for housing allowances.

Box 20.1. Rent regulation in the Netherlands

Both rent levels and rates of increase are regulated in both the social and private rental sectors, as long as the value of the rent under the rules falls below the “liberalisation threshold” (EUR 597.54 in 2004/05) at the start of the rental agreement. The rules for the two sectors are almost identical. The rent for each dwelling must fall within the “maximum reasonable rent” determined on the basis of the Housing Valuation System, a points system that assesses each dwelling in terms of size, quality and location. Regulated rents may be reviewed once a year, usually on 1 July. If the current level of the rent is in line with the points assigned to a dwelling, then the rent may rise at the trend rate of increase determined by the housing minister. Before 2004, this was usually defined as a moving average of CPI inflation over the recent past plus a specific percentage. If the rent was greater than the points system would award, the allowable increase would be below trend – a rent freeze or reduction might even be possible. If the current rent fell below that defined by the points system, then an above-trend increase was allowed, but only up to a maximum rate of increase, which was set annually by the minister. The application of these rules in respect of specific dwellings can be affected by the fact that HAs (but not private landlords) are also subject to a maximum average rent increase for their entire stock. This is usually lower than the maximum increase allowed for any given dwelling.

Increases over and above the annual increments allowed under the rules may be adopted when there is a change of tenant, though even then the point-based valuation of the property will determine the maximum rent unless it is found to exceed the liberalisation threshold. Landlords’ ability to profit from turnover adjustments depends on tenants’ decisions, since Dutch law does not provide for temporary tenancy agreements, and any contracts drawn up as temporary will generally be treated as indefinite by the courts. This means that low-income households which qualify for social housing continue to enjoy below-market rents even if their incomes rise substantially, provided they do not move: the only means test is at the point of entry. In practice, rents in both social and private regulated sectors are well below the maximum rents – 70% in the former and 83% in the latter in 2003. The lower ratio for the HAs reflects the fact that they were less profit-driven than private landlords. The ceiling on average rent increases for HAs may also have slowed down the rise in rents in the social sector. Priemus (2006a:226).

Households that cannot afford to pay regulated rents are eligible for housing allowances. Housing allowances are paid to households with incomes below certain thresholds and are based on the relationship of rent to income. Up to a certain limit, the state pays 100% of the difference between the “standard rent” defined for a particular
household as a proportion of income and the actual rent it pays. The state pays 75% of the additional rent between this ceiling and the actual rent up to a cap, above which it pays only 50%.\textsuperscript{19} The introduction of the housing allowance in 1970 was intended to make it easier to allow rents to become more market-determined, while relying on a means-tested benefit to protect the poorest, but rent regulation remained untouched.\textsuperscript{20} As regulated rents rose relatively rapidly during the 1990s (though less rapidly than market rents and house prices), expenditure on housing allowances rose sharply. By 2000, one-third of all tenants received housing allowances, up from just over one-quarter in the early 1990s. These were not particularly high proportions by international standards, but they were unusual in a country where almost all rents were regulated.\textsuperscript{21}

The open-ended nature of the benefit was increasingly a concern for the finance ministry, as were the perverse incentives it created. Sharp withdrawal rates meant that housing allowances were a major disincentive to increased earnings, while encouraging actual or fake division of households. Moreover, the allowance system generated no incentives to seek more affordable accommodation, since a reduction in housing costs would result in a reduction in the allowance. Indeed, they generated the reverse incentive: households moving to more expensive accommodation bore only a part of the additional cost involved. By pushing up rents in a market characterised by low price elasticity of supply, housing allowances also increased the share of households in need of rent subsidies.\textsuperscript{22}

Problems in the housing sector

The rent regulation regime has come in for criticism on numerous grounds. For the economy as a whole, its most costly side effect may be the barriers to labour mobility it creates: middle and upper middle income households renting in the social sector face extremely high costs if they move, and housing shortages in high-growth regions can make it impossible for workers to move to where there are job opportunities.\textsuperscript{23} These shortages were growing palpably worse in the early 2000s: on the government’s estimate, the shortfall of rental accommodation nation-wide\textsuperscript{24} rose from around 1.5% in 1998 to 2.5% in 2002, and it was expected to surpass 3% in 2003-04.\textsuperscript{25} This national figure masked significant regional disparities. There was surplus stock in much of the country but the large cities faced serious shortages. Thus, in 2004, the average waiting time for social housing was 3.4 years nation-wide, rising to 8-10 years in the larger cities.\textsuperscript{26}

Rent regulation was also socially regressive: an estimated 30-40% of the HAs’ rental stock was occupied by households with above-average incomes, whereas those searching for accommodation in the social sector tended to be on lower incomes.\textsuperscript{27} This was because the only means test for social housing was at the point of entry: households renting in the regulated sector continued to enjoy protection even if their incomes rose substantially. On one recent estimate, only around 43% of the financial benefits of rental policy was going to lower-income households.\textsuperscript{28} While this strengthened the case for reform, it also made reform harder, since better off tenants, who stood to lose, were more likely to organise and lobby effectively. Many of them were also among the coalition parties’ supporters. Yet although housing shortages were growing worse, there was no sense that the system was in crisis: the vast majority of households in the rental sector benefited from the system, and the efforts of rental-market outsiders were focused on finding rent-controlled accommodation rather than on altering the system.

Shortages gave rise to other distortions, particularly in the major cities. These included illegal sub-letting (as much as 10% of the rental stock in Amsterdam\textsuperscript{29}), payment
of “key money”\textsuperscript{30} and very wide variations in quality/price ratios. It was increasingly clear that the points-based valuation system did not properly reflect differences in housing quality or relative scarcities on the market, and the gap between regulated rents and estimated market rents was large, with the latter a multiple of the former in Amsterdam and other large cities.\textsuperscript{31} Prior to 1995, successive governments had tried to limit this gap by providing multi-year subsidies for new construction. After that time, the only fiscal instruments used to stimulate supply were modest land-cost subsidies.

The distortions created by rent policy were interacted with two other factors: policy towards the owner-occupied housing market and policy towards new construction. Dutch tax policies have historically been extremely generous to owner-occupiers, providing them with \textit{de facto} subsidies that by the early 2000s were roughly equal to those provided to the rental sector (Table 20.1). These policies contributed to the very high level of house prices, which benefited home-owners and encouraged the diversion of supply to the owner-occupied market. Faced with inflated house prices and artificially low rents, many tenants with above-average incomes ended up consuming less housing than they otherwise would have, while impeding entry into the rental market for starters. At the same time, a number of factors constrained the supply response to rising prices on the Dutch housing market.\textsuperscript{32} Spatial planning policy, which is largely in the hands of sub-national governments, has tended to be very restrictive, while planning procedures, building regulations and other regulatory policies further limit the supply response to rising prices.\textsuperscript{33}

Table 20.1. \textit{Housing subsidisation in the Netherlands, 2001}

<table>
<thead>
<tr>
<th>EUR billion, unless otherwise indicated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Owner-occupied sector</strong></td>
</tr>
<tr>
<td>Tax saving from deductibility of mortgage interest payments</td>
</tr>
<tr>
<td>Tax saving from investing equity in house (the avoided taxation on alternative investments)</td>
</tr>
<tr>
<td>Tax on imputed rent</td>
</tr>
<tr>
<td>Local property tax and transfer tax</td>
</tr>
<tr>
<td>Total owner-occupied</td>
</tr>
<tr>
<td><strong>Rental sector</strong></td>
</tr>
<tr>
<td>Housing allowances</td>
</tr>
<tr>
<td>Local tax</td>
</tr>
<tr>
<td>Government subsidy (narrow definition)</td>
</tr>
<tr>
<td>Rent reduction by corporations and through regulation</td>
</tr>
<tr>
<td>Total rental (narrow definition)\textsuperscript{1}</td>
</tr>
<tr>
<td>Total rental (broad definition)\textsuperscript{1}</td>
</tr>
<tr>
<td><strong>Total of both sectors (narrow definition)\textsuperscript{1}</strong></td>
</tr>
<tr>
<td><strong>Total of both sectors (broad definition)\textsuperscript{1}</strong></td>
</tr>
<tr>
<td>\textit{As a per cent of GDP}</td>
</tr>
</tbody>
</table>

1. The broad definition includes the benefit provided by social and private landlords via regulated rents.


Policy process: formulation

\textit{The formulation of the reform}

By 2003, it was clear that the housing shortage was worsening. Aedes was pressing for some relaxation of rent controls as well as measures to ease the constraints on supply, arguing that the net present value of future rent streams was insufficient to justify
investment in many places where housing was at a premium, but VROM was concerned that such a policy would trigger even faster growth in expenditure on housing allowances, which would force cuts elsewhere in its budget and which the finance ministry could not countenance. Discussions between Aedes and VROM thus moved towards a two-prong strategy, involving some adjustment of existing rent regulations together with the creation of a separate regulated rent system for low-income households. Initially, they focused on adjusting the rules governing rent levels and increases, but when the Balkenende II cabinet took office in 2003, the new minister, who represented the economically liberal VVD, raised the issue of liberalising some portion of the market and commissioned research to assess its impact. The results suggested that the shock of a partial liberalisation might be less than had been feared, and the VROM started to explore options for increasing the size of the free sector from 5 to 25% of the market. From the summer of 2003, the ministry began informal consultations with a half-dozen other ministries, as well as with Aedes, the other landlords’ organisations, the Woonbond and other stakeholders.

The initial idea of freeing 20% of the market evolved significantly during 2003-04, while the ministry was formulating its legislative proposals. The most intensive discussions involved Aedes, which agreed, in return for greater freedom in rent policy, to compensate recipients of housing allowances via a so-called “target-group discount”. There was concern, however, that this would limit the investment capacity of HAs, as housing allowance deficits increased from year to year, and that it was in any case unrealistic to expect landlords, whether social or private, to administer the relevant means tests, given that they would not have direct access to tax information and other relevant data. Moreover, if landlords were obliged to offer such discounts to members of the targeted group, then landlords would have strong incentives to reject such applicants. Instead, therefore, it was agreed that landlords would make additional payments to the government to offset the increased cost of housing allowances arising as a result of deregulation. Aedes agreed to such payments, but the IVBN wanted income support through housing allowances to come from general taxation, arguing that a dedicated tax would deter investment in housing. It was also agreed that the liberalisation must proceed slowly and housing allowance recipients must be protected. Thus, tenants in properties to be liberalised would be given five more years of regulated rents, although their regulated rent increases would be faster than those of others, and housing allowance recipients in dwellings that might otherwise be liberalised would continue to enjoy regulated rents under the pre-reform rules. These compromises then formed the basis for a set of government proposals published in November 2004 in the form of a letter from the minister to the lower chamber of the States General.

The “rent letter”

The so-called “rent letter” proposed a large number of inter-related changes to rent policy:

- **Partial liberalisation of rents.** The regulated share of the rental sector was to be reduced to 75% by 2010; in the interim, rents for the affected dwellings would be allowed to rise 2-4% faster than other regulated rents, which would remain tied to inflation. As a result, there would be roughly 750 000 rented dwellings in the liberalised sector, of which around 500 000 would be owned by HAs. Tenants not on housing allowance would pass through a four-year transition period, during which the maximum annual rent increase per dwelling would rise from
inflation+2.0% to inflation+4.0%. Housing allowance recipients would retain their allowance entitlements and would also continue to enjoy regulated rents under the old rules.

- **Faster regulated rent increases.** In an effort to strengthen the role of market forces in the remaining regulated market, the letter provided for faster growth in regulated rents, subject to investment in sufficient new supply to create more slack in the rental market.36

- **Modernisation of the valuation system.** The valuation system was to be revised to take better account of the way in which location and living environment affected the maximum rent and (albeit indirectly) the permitted rent increases. The letter proposed relying on the method set out in the legislation on real estate valuation for the real estate tax, the so-called WOZ Act. The government argued that this offered a better indication of the market value of real property. Use of the WOZ system, which was itself to be modified, was also intended to make the housing market more transparent, since the values of owner-occupied, regulated-rental and unregulated-rental dwellings would be more readily comparable.

- **New arrangements for housing allowances.** Changes were proposed in the formulae for calculating the housing allowances and maximum increases on the lowest-priced rental accommodation. At the same time, landlords were to contribute EUR 250 million a year (indexed) to offset increasing public expenditure on housing allowances. This was most likely to be raised via a tax related to the earning capacity of the housing stock, applied to both housing associations and private landlords. The government argued that this would have recouped some of the costs of reform from its principal beneficiaries.

- **Stimulation of new construction.** The above changes were contingent on sufficient loosening up of the supply side. To that end, the government was to conclude multi-year housing construction agreements with the provinces and municipalities, and landlords were to conclude obligatory performance agreements. Aedes had expressed its willingness to do so. The Housing Minister was to assess the situation as of 1 January 2008 to see if the housing market had loosened up enough to permit the proposed rent increases in 2008 and 2009. It would then be decided whether to deregulate these rents once and for all, based on an assessment of regional housing construction: deregulation would proceed only if enough new building was taking place. The overall aim was to reduce the housing shortage in each urban region to a maximum of 1.5% by the start of 2010. This figure was to be no higher than 2% by the start of 2008 if the deregulation process was to move forward uninterrupted. The policy rested on the targets shown in Table 20.2, which were relatively modest: production in 2005-06 was to remain at about the historically low level of 2004, while the targets for 2007-09 were still well below the annual increases of over 40 000 typical of the early 1990s.37
Table 20.2. Indicative targets for housing production by housing associations

<table>
<thead>
<tr>
<th>Year</th>
<th>Association dwellings for rent</th>
<th>Association dwellings for sale</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>10 000</td>
<td>5 000</td>
<td>15 000</td>
</tr>
<tr>
<td>2006</td>
<td>10 000</td>
<td>5 000</td>
<td>15 000</td>
</tr>
<tr>
<td>2007</td>
<td>15 000</td>
<td>6 000</td>
<td>21 000</td>
</tr>
<tr>
<td>2008</td>
<td>20 000</td>
<td>7 500</td>
<td>27 500</td>
</tr>
<tr>
<td>2009</td>
<td>25 000</td>
<td>7 500</td>
<td>32 500</td>
</tr>
</tbody>
</table>


Policy process: adoption

Reaction to the proposals

By the time the VROM proposals were presented, there was a basic consensus on a number of issues. First, although rents had been rising faster than incomes, raising the average net-rent-to-income ratio from 19.7% in 1990 to 23.9% in 2002, they needed to rise further to bring the housing market into equilibrium, eliminating illegal sub-letting, key money and waiting lists. Secondly, housing allowances targeted assistance better than property subsidies and rent controls. Thirdly, some relaxation of the constraints on new supply had to occur if rent reform were not simply to deliver windfalls to landlords. Finally, it was widely accepted that historical rent trends appeared to be irrational, in that they were very hard to explain in terms of differences in quality or social/physical environment. Nevertheless, the proposals generated significant criticism from the Woonbond, some of the parliamentary parties and a number of experts, who questioned both the timing of the reform and the likelihood that the measures proposed would deliver the desired outcome. Successive adjustments to the proposals in an effort to address the concerns of various stakeholder groups at different stages of the policy process (Figure 20.1) had the effect of making the reform proposal increasingly cumbersome.

Figure 20.1. Stages in the rent reform debate, 2003-07

Even some who were sympathetic to reform in principle questioned the ministry’s timing. When the proposals were mooted, the economy was just emerging from the stagnation of 2001-03, and the social partners’ agreement on wage moderation made for declining real disposable incomes for many. The Woonbond and others thus argued that the reform threatened to raise rents at a time when households were already feeling squeezed and that the provisions covering the transitional period did not provide them with sufficient protection. Moreover, house-building had fallen sharply since 2000, at a time when there was a trend away from buying and towards renting, which aggravated the shortage of affordable rented accommodation. Critics thus expressed doubt about the
government’s ability to stimulate sufficient new supply of rental accommodation, arguing that private investors’ decisions were driven far more by the prospects for selling homes than by rent policy. The commercial developers affiliated to the IVBN had little interest in the rental market and even less in the low-to-middle range rental market. Thus, any increase in supply would have to come from the HAs. However, it was not clear that the HAs had made their building plans dependent on rental policy: the HAs since the early 1990s had kept rents below regulated maxima while building chiefly to sell, which hardly suggested that the level of rents was the decisive factor in their investment plans. These doubts about the supply response were reinforced by the evident lack of enthusiasm of many municipal authorities for the reform: to a great extent, it was the municipalities who controlled the availability of land, and many of them preferred to restrict construction altogether or to focus on high-end rentals and owner-occupied dwellings. Other municipalities expressed the opposite concern: they were worried about the impact of higher rents on tenants in their jurisdictions.

The Woonbond, which organised substantial demonstrations against the proposals in Amsterdam and The Hague, focused much of its opposition to the reform on the adoption of the WOZ system for property valuation. While the tenants’ union acknowledged that the old system’s subjective assignment of points concerning things like location and social environment was hard to defend, they argued that modernisation of the valuation system, when combined with partial liberalisation and higher regulated rent increases, would lead to forced mobility, as low-income tenants were compelled to leave more desirable areas, and thus to the emergence of ghettos. The concern that many neighbourhoods might move out of the reach of lower-income households proved a potent argument with many on the left, who feared that reform would lead to social segregation.

Much of the criticism focused on the compromises reached in order to secure stakeholder acceptance of the plan. Thus, both tenants’ groups and many landlords challenged the requirement that landlords compensate the government for any increase in housing allowance expenditures resulting from above-inflation rent increases. This measure had great appeal for the finance ministry, as it effectively ensured the affordability of the housing allowance scheme while protecting the state budget, but it also implied that housing allowances would increasingly be financed by transfers within the rental sector: middle-income tenants in rented accommodation would help finance a social policy that exerted upward pressure on their own rents, while owner-occupiers would make no contribution. In addition to their traditional responsibility to provide affordable housing below the market price, the HAs would effectively assume much of the government’s responsibility for social protection. The attempt to protect recipients of housing allowances also came in for attack, since maintaining rent controls for those receiving housing allowances would lead to a widening gap between what they paid and what others paid for the same standard of accommodation. Moreover, they only retained such protection if they remained where they were. This, it was argued, would impede the mobility of housing allowance recipients, who might otherwise improve their position in the labour market by moving, and would create new opportunities for abuse. As a result, it was argued, shortages at the low end of the market would not be reduced by rent reform and might even get worse.

The parliamentary process

Although the coalition commanded a majority in parliament, the rent deregulation bill encountered significant resistance there. The Labour Party (PvdA) and other parties on
the left opposed it, seeing it as an attack on tenants and arguing that it was inequitable to deregulate the rental sector without touching generous de facto subsidies for owner-occupiers, who constituted an important part of the coalition parties’ constituency. Even some of the government parties expressed significant misgivings, with a number of CDA deputies arguing that sitting tenants’ rights should not be affected and that the new rules should take effect only when there was a change of tenant. This, however, would have postponed the positive effects of the reform for decades, while exacerbating the barriers to labour mobility that the reformers hoped to reduce. Concern about social segregation prompted deputies to adjust the valuation mechanism in the legislation so that the “WOZ system” would form the basis for only 20% of the valuation, a move that greatly reduced the significance of location and environment in valuing regulated rental properties; the rest of the valuation would be determined by a slightly modified version of the existing points system. Many CDA parliamentarians were also sceptical of the government’s ability to ensure sufficient loosening on the supply side. Much of the CDA was concerned about the policy, particularly provincial and municipal elected officials, who feared the electoral consequences of higher rents. Such fears gained substance when the CDA lost 15% of its local councillors in the municipal elections of 2006: a large proportion of these losses came in Amsterdam and other large cities, where rent policy was a salient issue.44

The rent bill finally passed the lower house at the end of October 2006, nearly two years after the rent letter was published. However, it then met resistance in the upper house, whose members were elected by provincial councils. The “intra-rental” transfers that were to finance housing allowances after the reform came in for criticism, and technical concerns were raised about issues like valuation, where it was suggested that even partial reliance on the WOZ system could open up rent valuations to judicial review. However, before these issues could be resolved, the general election took place, resulting in losses for all the governing parties. When the next coalition was formed, in February 2007, the CDA found itself in government with the PvdA and the Christian Union. The coalition agreement stated explicitly that the rent deregulation bill would be withdrawn and that rent increases would be linked to inflation – an even more restrictive policy than that prevailing prior to the debate over the rental letter. For the CDA, this was not a huge concession: rent reform had been more of a priority for the VVD, and the Christian Democrats themselves had always been divided over it. While it had been part of the party’s platform for the 2006 general election, many CDA parliamentarians had worked to block its passage in the previous parliament.

The government’s communications difficulties

From the beginning, the reformers found it hard to persuade parliament or the public of the urgency of reform or the desirability of the proposed solutions. While there was widespread awareness that the rental market was not functioning well, it was not perceived to be in crisis or in danger of imminent collapse, and public awareness of the costs of the status quo was limited – the issue had not been prominent in the general election campaigns of 2002-03 and the government’s main economic analysis unit, the influential CPB, had done relatively little work on it.45 The CPB thus had less impact on debates over rent regulation than it did on many other issues, including disability benefit reform. At the same time, doubts about the solutions on offer were fuelled not only by the kind of criticisms outlined above, but also by uncertainty about where the reform was headed: since the endpoint towards which the government was working was never defined, opponents were able to argue that deregulation, once begun, would proceed further and further. Moreover, despite the fact that the status quo undoubtedly
disadvantaged many low-income households and despite the measures included in the reform in an effort to protect low-income households, government found it hard to make the case for reform in terms of equity: there was much discussion of “forced mobility” and the emergence of ghettos, feeding the impression that the proposals were socially regressive. Even critics’ use of the word “liberalisation” to describe a very cautious partial deregulation proposal served to reinforce fears that the reform would prove very radical and far-reaching.

Conclusion

Building support for a reform that confronted the vast majority of tenants with the prospect of higher rents was always bound to be difficult. The only obviously identifiable winners in the short term would have been landlords, and if the anticipated supply response to reform were weak or slow to emerge, they could have captured enormous rents. Here the problem was the tension between demand-side policies regulated largely by the central government and the decentralisation of a great deal of authority over spatial planning to provincial and municipal governments that in some cases seem to have had an interest in restricting supply. The reformers’ inability to persuade parliamentarians or the public that the supply-side measures contained in the reform were adequate was thus a significant weakness, since it contributed to the perception that the main immediate effect of the reform would be to redistribute wealth from tenant households to landlords. If the political costs of a reform are associated with the amount of income redistributed away from loser groups, while the economic benefits are understood in terms of the efficiency gains that reform makes possible, then rent reform may be said to have had a high political cost – it would have involved a great deal of redistribution up-front, which would only have been offset by efficiency and other gains over a long period.

Proposals to avoid large-scale redistribution in the short term by “grandfathering” the rents of current incumbents offered no real solution, since this would have ensured that the reform had little positive effect for decades and would, in the meantime, have strengthened rather than weakened the barriers to mobility created by the status quo. Acquired rights had to be compromised for the reform to proceed. This meant that the reform would have created a large number of losers in the first instance, even if some households were broadly unaffected. Moreover, the households at the upper end of the social rented sector, which stood to lose most from liberalisation, were also, on average, better off and better organised than those whose access to affordable accommodation might be improved as a result of the reform. Finally, the identities of the potential losers were fairly clear ex ante, while the identities of the winners were not and might never have been. While reform might have made it easier for some low-income households to find more affordable housing in the social rented sector, many of them would eventually have done so even in the absence of reform, and they would have done so at lower cost. Thus, not all those disadvantaged by the existing arrangements would have ended up as beneficiaries of reform and those who did would not have known that they were winners, either ex ante or ex post.

These problems made effective communication of the case for reform, on both equity and efficiency grounds, all the more important, but in the end, the government failed to establish a consensus on either the urgency of reform or the attractiveness of its proposed solutions. On the contrary, the status quo, for all its defects, was both well institutionalised and popular; in that sense, it did not look like a policy regime that was yet “ripe” for reform. Moreover, lack of clarity about where the reform was ultimately
intended to lead – VROM failed to present a clear long-term vision of how it thought the rental market should be structured – left the government vulnerable to charges that the November 2004 proposals were merely the first step towards a potentially radical deregulation of the market, while the increasing complexity of the reform plan, as concessions were made to stakeholder groups, made it increasingly difficult to explain how it would achieve its aims.

The defeat of the November 2004 rent proposals notwithstanding, rent deregulation and other aspects of housing policy remain very much on the Dutch policy agenda. In October 2007, the Council of the VROM issued a major report on housing policy and recommendations for change to both rental and owner-occupied sectors, and a report from the Social and Economic Council is due in spring 2009. In addition, a number of CPB studies of the housing market appeared in 2007-08, furthering highlighting the case for reform. Any reform initiative that emerges from these discussions will need to overcome the major impediments to reform encountered in 2004-06.

Notes

1. Owner occupation rates, though rising in recent decades, are relatively low by OECD standards. They rose from 28% in 1947 to 54% in 2004; see Toussaint et al. (2006).
2. The position of the D66 is difficult to crystallise, not least because it is the subject of much debate within that party. In general, it is associated with socially liberal positions; it has tended to support market-oriented reforms in fields like healthcare and labour markets.
4. Priemus (2006a:231-32). It is possible to override the medium-term spending caps, but this requires an agreement among the parties to revise the coalition agreement during the government’s term of office; see Blondal and Kristensen (2002); and Anderson and Minarik (2006).
7. For the EU as a whole, social housing accounts for just 23% of the rental stock; see REA (2006). In non-European OECD countries, this share is typically smaller, though Canada, with 31% of all housing in the social sector, is an exception. For comparative data on OECD countries, see OECD (2006); and Norris and Shiels (2007).
8. Data from the Ministry of Housing, Spatial Planning and the Environment.
10. For details, see Priemus (2008).
13. The VB represents small landlords for whom rental activity is a significant business: around 100 000 rental units are owned by individual landlords who have purchased one or two properties to let, or by expatriates and others who have temporarily let their own homes.


15. Haffner et al. (2007).


17. The first supplementary allowance for housing was introduced in 1970, but it became a general instrument only in 1975; Priemus (2004).

18. These vary according to age and household composition. Allowances are permitted only for dwellings below the liberalisation limit.


24. Measured as the balance of people seeking new accommodation and the houses coming onto the market (new-builds or those released by households) as a percentage of the total stock.

25. See Boelhouwer (2005) on the evolution of the shortage and the factors underlying it.


27. Priemus (2006a:227). Around 700 000 HA tenant households had gross incomes in excess of EUR 33 000 per annum.


30. “Key money” is a one-off payment made by tenant to landlord at the start of a lease; it is generally not an issue for HAs but is widely reckoned to be a common, albeit illegal, practice with private landlords.


32. Swank et al. (2002) estimate a price elasticity of supply of 0.3 for Dutch housing, far lower than they estimate for other countries under study. Vermeulen and Rouwendal (2007) find investment in housing, measured in terms of both value and units built, to be highly inelastic to price, as are measures of structure and location quality; they conclude that this is chiefly the result of policy interventions by the authorities. See also Boelhouwer and Hoekstra (2008).
33. For a detailed look at restrictions on supply, see Boelhouwer (2005); and Boelhouwer and Hoekstra (2008).


35. Priemus (2006a:232)

36. See Priemus (2006a) for details


40. Boelhouwer (2006); Priemus (2006a).

41. As noted above, members of the VB generally had small portfolios of rented dwellings and rarely built new housing for rent.

42. This was entirely in line with government policy at the time, as the government had concluded in 1991 that there were too many social rented dwellings and insufficient supply on the owner-occupied market.


44. The impression that rent policy played a role here was reinforced by polls suggesting that the largest defections from the CDA were by voters on below-average incomes who backed the Socialist Party.

45. The major exceptions were ter Rele and van Steen (2001); and Hakfoort et al. (2002).

46. This is the conceptualisation of the problem found in Rodrik (1994).

47. On the concept of “ripeness for reform”, see Ross (2007).
Bibliography


Chapter 21

Spain: Retail entry and opening hours, 1995-2004

Summary and main conclusions

The regulation of retail hours and entry has been a controversial issue since the mid-1990s, when the government partially reversed an earlier retail-sector deregulation, allowing regional authorities to regulate both opening hours and Sunday/holiday opening, and to grant or withhold permission to open new hypermarkets. The centre-right Popular Party (PP) government sought to liberalise opening hours and holidays again after the 2000 elections but achieved only very limited progress. In 2004, a new Socialist government reversed the PP measures in respect of large stores, although it did not re-regulate small retailers. Numerous barriers to entry and competition thus remain in the retail sector, largely due to the policies of regional authorities. Studies suggest that the cumulative impact of these barriers is significant, retarding productivity growth and increasing prices for consumers. The restrictions on opening hours also impose non-pecuniary costs on consumers trying to reconcile work and family life. The politics of the issue thus involve both tensions between central and local governments and conflicts between producer and consumer interests.

Efforts to open up the retail market were frustrated by a number of factors, of which the most important included:

- **The power of small retailers.** Small retailers are exceptionally numerous in Spain, which has one of the most fragmented retail sectors in the OECD. They are well organised at provincial and regional levels, and are particularly prominent among the supporters of nationalist/autonomist parties in places like Catalonia and the Basque country, which at times provided crucial parliamentary support to minority governments at the national level. The small retailers were supported in their campaign to restrict competition from hypermarkets by trade unions and, in some regions, church organisations.

- **Support for restriction from many incumbents.** Entry and opening hours for small independent shops were fully liberalised throughout Spain, and a number of significant retail categories, including convenience stores and shops in tourist areas, were exempt from the restrictions. Restrictions on hours and Sunday/holiday opening thus applied only to larger shops outside of tourist areas, while the licensing regime affected only potential entrants above a given size. A large number of incumbents therefore profited from the restrictions. During the late 1990s, restrictions on very large hypermarkets tended to benefit national supermarket chains, which might otherwise have pressed for change. Regulation of entry really started to affect middle-sized supermarkets only later, as regions...
began lowering the size threshold above which retail outlets were subject to regional licensing requirements.

- **The lack of a strong pro-reform coalition.** The welfare gains from reform, in terms of efficiency and productivity, were more diffuse than the rents being appropriated by the protected sector, and the central government failed to communicate a strong case for reform when introducing changes in 2000. While the large retail sector supported liberalisation and there were proponents of reform in the Tribunal for the Defence of Competition and some of the national ministries, there was no strong coalition pressing for change. The large retailers enjoyed far less public sympathy than the proprietors of traditional small outlets, and consumer interests were weakly represented in policy processes, though when regulations were relaxed, consumers took advantage of the new opportunities.

- **Limited scope for cross-border competition.** In contrast to Germany, where sub-national governments pressed for liberalisation and reform was long blocked at federal level, it was the regional governments in Spain that tended to favour restrictive regulation. There are few places in Spain where relatively liberal trading conditions might draw significant numbers of consumers from more restrictive neighbouring jurisdictions. Consequently, the kind of cross-border pressures that contributed to reform in many German Länder were lacking in Spain.

**The context for reform**

**Framework conditions for reform**

Constitutionally, Spain is one of the most decentralised countries in Europe, with a wide range of powers devolved from the centre to the governments of its 17 regions, known as “Autonomous Communities” (Comunidades autónomas), and to local governments. Central governments have been subject to persistent regional pressure for more autonomy in a wide range of policy domains, particularly from Comunidades with strong nationalist or regionalist parties, such as Catalonia and the Basque Country. The regions’ influence was enhanced during 1993-2000, when the central governments of Felipe González and Jose Maria Aznar lacked congressional majorities and relied on support from representatives of small nationalist and regionalist parties in the national parliament, most notably the Catalan Convergence and Unity (CiU) party and the Basque Nationalist Party (PNV). This had a significant effect on the politics of retail regulation. During 2000-04, the second Aznar government commanded a majority in Congress, but its successor, the Socialist government of José Luis Rodríguez Zapatero, did not and had to rely on the support of minor parties.

The impact of economic drivers on reform was limited and ambivalent. Although the initial liberalisation of 1985 was adopted in an effort to stimulate growth, there is little subsequent evidence of a link between growth performance and reform effort: restrictive policies were adopted in 1993 in the midst of a severe downturn and then confirmed during the ensuing strong recovery. Unemployment, however, was an issue: though trending downwards after 1994, it remained well above EU and OECD averages, and high levels of joblessness reinforced concerns about job losses in the event of retail liberalisation. Inflation was also a factor: reform advocates argued that greater retail
competition would help to reduce inflation, which was consistently above EU and euro-area averages.

**Actors and interests**

Regionalist and nationalist parties in Spain tended to be strongly in favour of regulating retail opening hours and entry, particularly the CiU, which argued that Catalonia’s strong small business tradition was a part of the Catalan national identity that must be protected. The national political parties, by contrast, adopted somewhat ambivalent positions with respect to the issue. Thus, the Socialist Party (PSOE), despite its relationship with the trade unions, adopted a significant liberalisation of the retail opening hours in 1985, before partially reversing this reform in the 1990s. The centre-right Popular Party (PP), for its part, was also subject to cross-cutting pressures. Supported by the bulk of the Spanish business community, the PP was broadly in favour of liberal policies, but elements of the party were hesitant about opening up the retail sector, owing to the concerns of small retailers, many of whom supported the PP. The party thus supported the PSOE’s partial re-regulation of the sector but then sought to relax those restrictions when it came to power. At times, moreover, local branches of both major parties adopted more restrictive positions when in power at regional level. In general, regional governments of whatever political stripe wanted the power to regulate retail businesses and were eager to use that power to protect local businesses and retail employment.²

The trade unions in the retail sector – the sectoral unions affiliated to the Trade Confederation of Workers’ Commissions (CC.OO) and the General Union of Workers (UGT),³ as well as the Federation of Trade Union Associations (FASGA) and the Federation of Independent Retail Workers (FETICO) – all opposed longer opening hours and Sunday/holiday opening. They argued that liberalisation would lead to a deterioration in working conditions for over 2m retail workers, a majority of whom were women and many of whom worked part-time.⁴ The unions claimed that liberalisation of hours would lead to more precarious employment and greater difficulty balancing the needs of work and family life. On other issues, the unions’ positions diverged: FASGA and FETICO, which were not affiliated to the UGT or CC.OO, represented workers in the large retail chains (mainly at the level of middle management, in the case of FASGA) and were therefore not inclined to support restrictions on the entry of large-format retailers into new markets. The CC.OO and UGT unions supported regulation of entry, arguing that the sector would become less labour-intensive if small retailers lost market share.⁵

The business lobby was broadly divided into three groups, owing in part to different views of the three issues at stake: opening hours, Sunday/holiday opening and market entry for large retailers.

- The Spanish Federation of Small and Self-Employed Retailers, the Spanish Retailing Federation and several regional entrepreneurial associations, which represented small independent retailers, all defended regulation of entry and opening hours, insisting that liberalisation would force small shopkeepers into “wage slavery” and subject them to “unfair” competition from larger players.⁶

- By contrast, the Spanish Association of Shopping Centres (AECC), which represented small shopkeepers in shopping centres, and the National Association of Large Retail Establishments (ANGED), which represented the hypermarkets, department stores and other very large-format retailers, strongly favoured liberalisation of both hours and entry, as did the Confederation of Employers’
Organisations (CEOE). Virtually all the restrictions concerned ANGED, since they were, in one way or another, aimed at protecting other retailers from the big hypermarkets and department stores. The AECC was particularly concerned about opening hours and Sundays/holidays, since most large shopping centres also contained cinemas, restaurants and other leisure services, and developers wished for the shops to be able to operate alongside them. For the CEOE, market unity was the key concern: whatever restrictions were implemented, the CEOE argued for uniformity of regulation across Spain, citing the cost to firms of complying with different regulatory requirements and procedures in every region. In general, though, the large players enjoyed far less public sympathy than the proprietors of traditional small outlets.

The political power of the first of these three groups was greater than might be expected, because Spain’s retail sector was and remains much more fragmented than those of most OECD members. The number of retail establishments per head of population in the early 2000s was roughly twice the EU-15 average, and the number of employees per outlet just over half the EU average. Only Italy’s retail sector was similarly fragmented. Though its share in total employment was only slightly higher than that of Germany, for example, Spain had more than twice as many retail companies relative to population as Germany, and just under half of workers in Spanish retail were self-employed. Small retailers are well organised at provincial and regional levels, and are prominent among the supporters of nationalist/autonomist parties in places like Catalonia and the Basque Country, as well as among local PP and PSOE activists in various regions. Their public image is broadly favourable and they have been supported in their campaigns to restrict competition from larger entrants by trade unions and, in some regions, church organisations.

Spanish consumers were, of course, the largest constituency affected by retail regulation. However, consumers were poorly organised and the issue was not of sufficient electoral salience for most consumers for their voting power to be decisive. Opponents of reform argued that liberalisation of hours would lead to higher costs, which would be passed on to consumers in the form of higher prices, while reform advocates insisted that the increased competition and consolidation in the sector resulting from the relaxation of restrictions on opening hours would more than offset this effect, leading to lower prices overall as a result of greater efficiency in retail. Consumer surveys did not find most households dissatisfied with the status quo: successive Ministry of Agriculture, Fisheries and Food surveys consistently found that over 80% of respondents were satisfied with the opening-hours regime. On the issue of market entry, by contrast, the issue was clear-cut: prices tended to fall when restrictions on entry were relaxed and competition increased. When small traders sought to keep large hypermarkets out of their markets, voters were often inclined to support the former. Yet actual consumer behaviour pointed to a different conclusion: new entrants attracted custom even where they had encountered political
resistance. Supporters of reform also argued that a liberal regime with respect to hours and entry would offer consumers more choice and more flexibility in the use of their time, which would make it easier to reconcile work and family life.13

Issue history

In 1985, the PSOE government introduced complete liberalisation of opening hours as part of a package of measures aimed at stimulating growth.14 Nevertheless, a number of regions either maintained existing restrictions on operating hours and holiday/Sunday opening or introduced them in the years that followed. In addition, some regions began adopting more restrictive conditions for the entry of very large retail outlets. Where large retailers had previously required licences only from the municipalities in which they wished to open, a number of the Comunidades imposed regional licensing requirements as well, arguing that municipal authorities did not take proper account of the impact of their decisions on neighbouring jurisdictions and that they were too easily swayed by the potential tax revenues such development would generate and by the hypermarkets’ promises to create employment and invest in local infrastructure.15 In July 1993, the Constitutional Tribunal struck down a number of these regional laws, prompting the regions to demand that the central government delegate such issues to them: the first bill submitted in the new parliament was a CiU proposal to give the regions the authority to regulate retail trade, which was loosely modelled on France’s 1973 legislation on commercial zoning regulation.16

Such pressure proved effective. Having lost its majority at the June 1993 general election, the PSOE government was reliant on CiU support in Congress. Moreover, small retailers, who were an important part of the PSOE electorate in some regions, were pressing for change, even to the point of organising public demonstrations. They had been losing market share to supermarkets and hypermarkets for over a decade, and as the recession bit, they stepped up their demands for protection, including a reduction in permitted opening hours to 60 per week with mandatory closing on Sundays and holidays.17 The government’s response was to adopt a decree-law as an interim measure, invoking the severity of the on-going recession and the consequent need to ease the pressure on small businesses to justify its retreat from the liberalisation adopted in 1985.18 The decree-law gave the regions the power to limit weekly opening hours, provided the weekly limit was not below 72 hours, and to restrict Sunday/holiday opening, provided that retailers were able to open on at least 8 such days per year. Retailers were given freedom to distribute their 72 hours throughout the week as they saw fit, but the regions were authorised to choose the Sundays and holidays on which retailers could open. Exclusions were allowed for tourist areas, shops in railway stations and airports, bakeries and some other food outlets and small convenience stores.19 These provisions were confirmed in statute in early 1996,20 at which time the government also confirmed the regions’ power to licence the entry of large-format (at least 2500m2) retail outlets.21 The new law’s stated intent was liberal: it confirmed the principle of freedom of opening hours. However, it declared that this principle could not be applied immediately. The regime established by the law was described as transitional, but it would not be open to review before 2001 at the earliest.

In practice, most regions opted for the tightest restrictions allowed: 72 hours a week and eight Sundays a year for most shops. The sole exception was Madrid, which initially allowed 12 Sundays or holidays, rising to 21 in the early 2000s. By 1999, most regions had stricter regulation of opening hours than were found elsewhere in Europe. For most retailers, the number of hours mattered less than the question of Sunday and holiday
opening. All but the largest retailers and some shopping-centre developers found 72 hours sufficient, and the law in any case left merchants free to distribute their hours as they saw fit. Such flexibility also affected the unions: while they supported restrictions on opening hours for fear that freedom of opening would lead to more part-time employment and more anti-social hours, the failure to define specific opening hours in the legislation ensured that the link between working hours for employees and opening hours was less clear-cut in Spanish retail than it was in most jurisdictions with strict opening-hours regulations. To some extent, therefore, the issue had more to do with small retailers’ interests than with protecting their employees from “anti-social” hours. However, since around half of all employment in the sector consisted of self-employed persons (as against an EU15 average of about 30%), the issues were linked. A very large proportion of Spanish retail outlets had no employees, so opening hours were by definition the proprietors’ working hours.

For very large retailers, the implementation of the regional licence for new hypermarkets proved even more of a problem than the regulation of opening hours. Such licences were frequently denied, especially in Catalonia and the Basque Country. Licensing decisions were taken by local commissions, which considered the impact of entry on local incumbents. Consumers were represented on these commissions, but elected politicians and local retailers predominated. While the region was required to obtain a report from the Tribunal for the Defence of Competition (TDC) before approving or denying an application to open, the Tribunal’s views were not binding; the decision was highly discretionary and there was no appeal to the central government in the event of a denial. In theory, the new licence was meant to address congestion and other potential externalities that might extend beyond a given municipality, but the TDC concluded in 2003 that many regions simply used the law to suppress competition. However, unions, small retailers and regional authorities argued that restrictions on competition among retail formats were needed to prevent the “hollowing out” of urban centres, to cater to those with mobility restrictions and to prevent staff from having to work anti-social hours. Large retailers were also accused of employing “unfair” competitive strategies against their smaller rivals.

Policy process: formulation

Although the PP had voted for the 1996 trade law along with all the other major parties, the PP government elected in 1996 was – in principle, at least – committed to the liberalisation of opening hours. During its first term in office, however, it made no move in this direction: any legislation would have required the consent of the CiU and other small parties, since the PP lacked a congressional majority. Since the 1996 law provided for a review no earlier than 2001, there was an argument for waiting and, in any case, the issue did not appear to be urgent: the economy was growing strongly, large retail chains seemed to be adapting to the regulatory environment by opening more medium-sized outlets rather than hypermarkets, and the government’s economic policy makers were preoccupied with privatisations, labour-market reforms, an overhaul of competition law and other issues. In such circumstances, the PP had little incentive to antagonise not only its regional party allies in Congress but also those small retailers who supported it. The issue of retail regulation reappeared on the policy agenda only after the March 2000 elections. Returned with an absolute majority in Congress, the government began its second term with a package of regulatory reforms, covering a large number of sectors, including retail, pharmaceuticals, publishing, energy and telecommunications.
The aim of the deregulation package was to stimulate growth, which was beginning to flag after four very strong years, without increasing inflation, which remained stubbornly above the euro-area average, thereby eroding Spanish competitiveness. Retail trade was included in the package at least in part because the 1996 legislation was approaching its expiry, and the government would have to renew it as it was or to modify it in some way. The case for change rested largely on the fact that productivity performance in retail trade was poor in comparison both with the economy as a whole and with less regulated retail sectors in other countries, while profit margins in the sector were growing faster than in the rest of the economy. This suggested that protected retailers were indeed appropriating rents. Large retailers were pressing for change in any case, as it was they who were most constrained by the existing regulations. However, the PP could hardly ignore the interests of small retailers, many of whom were among its supporters, nor did it have an interest in provoking conflict with the regions. Though aware of the need to spur competition and productivity growth, therefore, the government had little incentive to do anything very radical, and the proposals that emerged were fairly modest.

In April 2000, the government announced plans to grant small shops (under 300m\(^2\) and not belonging to a chain) total freedom to set their own opening hours, to increase the 72-hour minimum in the 1996 law to 90 hours for all other retail outlets and to allow shops to open on at least 16-20 Sundays or public holidays per year. This latter figure was reduced to 12 in the final version of the government proposal: there would be a minimum of 9 such days in 2001, with one further Sunday or holiday added each year up to a total of 12 in 2004. These arrangements were to have effect until 1 January 2005. Even with its absolute majority in Congress, the government remained sensitive to the regions’ prerogatives in this field: the draft decree law stated explicitly that “full freedom of shop opening hours and trading days shall not apply until the government jointly with the governments of the Autonomous Communities grants it in their respective territories and in no event before 1 January 2005.” Moreover, the government made no move to abolish, or even to impose any limitations on, the regions’ licensing powers in respect of new entrants. Yet for many in the pro-reform lobby, this was by far the most important restriction on competition.

Policy process: adoption

Despite their modest aims, the government proposals met fierce opposition from the unions and many retailers. The Spanish Federation of Small and Self-Employed Retailers, the Spanish Retailing Federation and a number of the regional associations of small retailers announced lock-outs in protest, claiming that they would be forced to extend their working hours and expressing the fear that total freedom of opening hours would be extended to large establishments in the near term. They also accused the government of acting to benefit foreign distribution companies, since increased Sunday/holiday openings would tend to transfer market share to the big hypermarket chains. A study commissioned by the supermarkets’ lobbying arm ASEDAS estimated the market-share transfer resulting from the change in Sunday/holiday opening at USD 1.67 billion and the resulting job losses at between 34 000 and 65 000. The retailers’ associations accused the government of electoral “fraud”, because the measure had not been included in the PP’s 2000 electoral programme: although the government had never disavowed its 1996 promise of further retail liberalisation, it had not made the issue a priority during the campaign, and the Minister of Industry had even promised that the legislation would not be modified during the parliament elected in 2000. Critics contrasted the unilateralism of the newly re-elected PP with the consensual approach taken by the PSOE in 1996, when...
virtually all major parties in Congress had backed the new trade law. Nevertheless, a package comprising some 70 liberalising measures across all the affected sectors was approved by the government on 23 June in the form of five decree laws and ratified by Congress a week later.  

The regions regarded the new law as a violation of their constitutional prerogatives, and the governments of Catalonia, the Basque Country, Andalusia, Asturias, the Balearic Islands and Castilla-La Mancha commissioned studies to analyse its possible unconstitutionality; several launched appeals against it on constitutional grounds. The unions also mobilised against the reform, with the CC.OO, UGT, FASGA and FETICO issuing a joint manifesto opposing the measure and announcing the beginning of a campaign to raise workers’ awareness of the issue. Protests were co-ordinated with small shopkeepers in all regions, and a strike of all retail-sector workers was called for 10 October. The government’s opponents also objected to its use of the decree-law mechanism to adopt the reform. The decree-law is a “fast-track” procedure that requires parliamentary action but does not permit political debate. Since decree-laws are theoretically reserved for occasions of “extraordinary and urgent need”, the government’s critics argued that this was an abuse of the mechanism intended to force through the changes with a minimum of discussion. If so, the government’s manoeuvre may have been a mistake, since it meant that it was never compelled to present a strong, clear case for deregulation to Congress or the public.

Policy process: implementation

Regional resistance to liberalisation

In applying the new legislation, which entered into force on 1 January 2001, the regions largely stuck to the strictest regulation allowed by the national law, just as they had done during 1996-2000. None allowed opening for more than 90 hours a week and the Balearic Islands stuck to 72 in defiance of the central government. Only Madrid opted for substantially more Sunday/holiday open days than required by the new legislation. Most regions adhered to the national minima, which rose year by year, and a few held below it – including, again, the Balearics. The most significant regulatory shifts in this period, however, concerned the licensing of large-scale retail outlets, which was proving to have some unintended consequences. It had become evident by 2000 or so that entry restrictions were imposing costs on consumers without arresting the decline of traditional retail outlets. Restrictions on hypermarkets tended to benefit not the small retailers but the national supermarket chains, which might otherwise have pressed for liberalisation of entry. Some of these had strong regional bases, which put them in a good position to lobby regional governments. The hypermarket chains, in turn, largely followed their lead, opening fewer very large hypermarkets and more medium-to-large supermarkets. Thus, in food retailing, the market share of supermarkets, which had been roughly stable for some years prior to the 1993 legislation, surged from 42 to 62% over the ensuing decade, and the number of supermarkets in the 1000-2500m$^2$ size class more than doubled. Small retailers’ market share continued to decline steadily and the number of small independent outlets fell by roughly a quarter, while the hypermarkets’ share peaked and began to fall.

Yet mounting evidence that regulation was ineffective in protecting small merchants did not prompt the regions to deregulate. Instead, they tended to regulate more, particularly in respect of entry. Licences were frequently refused. By 2003, ten regions had introduced moratoria on the entry of new hypermarkets, and around 50 hypermarket
applications were on hold.\textsuperscript{42} Some regions went for over a decade without awarding a single licence to open one. Licensing rules were also extended beyond their original scope (Table 21.1). In national legislation, the regional licence was envisaged for retail outlets in excess of 2500m\textsuperscript{2}, but beginning in the late 1990s, the regions revised this threshold downwards to as little as 250m\textsuperscript{2}, at times applying additional criteria, such as population in a given area. In some regions, licensing was extended to particular retail formats, like hard discounters. Other formats were sometimes exempted, such as furniture and “Do-It-Yourself” stores, which by definition required fairly large spaces.\textsuperscript{43} Numerous other, essentially ad hoc, measures were adopted, such as definitions of “large” based on ownership,\textsuperscript{44} which prevented the large chains from simply opening more small outlets, and restrictions on outlet growth or ownership change.\textsuperscript{45} Finally, in a further effort to protect small outlets, four regions imposed special taxes on large hypermarkets.\textsuperscript{46} Overall, 14 of 17 regions increased barriers to entry in the retail sector after 1996, with regions that were more retail-friendly \textit{ex ante} converging with the more restrictive Comunidades. Over half imposed outright bans on the entry of large retail outlets for at least part of the period.\textsuperscript{47}

In general, regional governments adopted policies explicitly aimed at ensuring what they regarded as a proper “balance” among retail formats – \textit{i.e.}, at limiting competition among different formats – and there were surprisingly few attempts by regions to exploit potential loopholes. With few exceptions, the notion of “tourist areas” was interpreted fairly strictly – in most places, it was limited to municipalities with more visitors than inhabitants. This meant that many small coastal villages qualified, while large cities like Barcelona did not, despite the number of tourists they attracted. Madrid was something of an exception: its decision to classify the city centre as a tourist area provoked significant controversy. There was also remarkably little evidence of the kind of inter-jurisdictional competition that affected German retail reform. Geography seems to have played a role here: the coverage area of a hypermarket is usually estimated at roughly 30 minutes’ driving time, rising to as much as 60km for very large specialty stores.\textsuperscript{48} There are relatively few places in Spain where a hypermarket or very large store in one region would be in a position to draw large numbers of consumers from other regions or across international frontiers. However, this would not seem to be a complete explanation, since this sort of cross-border competition was possible around Madrid, but its liberalising measures did not prompt any neighbouring regions to follow suit.

The greater fragmentation of the Spanish retail sector probably also played a role – Spain had more than twice as many retail companies relative to population as Germany, and average employment per retail outlet was just under one-third that of Germany. Small retailers were thus a larger voting bloc in Spain than in Germany. Finally, there is the question of heterogeneity among regions, which cannot easily be assessed. Both German \textit{Länder} and Spanish regions vary widely in terms of economic and social conditions. However, no Spanish Comunidades generated the kind of pressures for liberalisation that emanated from Germany’s eastern \textit{Länder}, which approached the issue of retail regulation from a very different historical experience from their western neighbours and which faced a very different set of economic challenges.\textsuperscript{49} The contrasting Spanish and German experiences highlight the difficulty in making categorical statements about the desirability of decentralised decision-making when it comes to product-market regulation. Much may depend on the specific institutional, geographic, economic and socio-political context.
### Table 21.1. Regional barriers to retail competition, 1997-2004

<table>
<thead>
<tr>
<th>Region</th>
<th>Hours/week</th>
<th>Sunday/holiday openings per year</th>
<th>Licensing</th>
<th>Other restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalucía</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aragón</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asturias</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balearic Islands</td>
<td>72</td>
<td>72</td>
<td>8</td>
<td>5 (2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basque Country</td>
<td>72</td>
<td>90</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canary Islands</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cantabria</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Castilla-La Mancha</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Castilla y León</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catalonía</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extremadura</td>
<td>72</td>
<td>90</td>
<td>9 (1997)</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Madrid</td>
<td>72</td>
<td>90</td>
<td>13-14</td>
<td>18-21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murcia</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navarre</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rioja</td>
<td>72</td>
<td>90</td>
<td>8</td>
<td>9-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Ranges reflect change over time as well as differential thresholds applying to different formats.
2. “Hard discounters” refers to specific licensing requirements for hard discount stores; “ownership criteria” refers to defining a firm to be “large” if it is at least 25% owned by a large firm; “expansion/ownership change” refers to restrictions on expansion or changes to a firm’s ownership.

Partial reversal

The approaching expiry of the transitional period legislated under the decree-law of 23 June 2000 meant that the PSOE government elected in early 2004 had to act on the issue relatively quickly. Elected with a minority in Congress, it was, like the governments of 1993-2000, reliant on the support of small nationalist and regionalist parties in Congress. With the unions pressing for change and the regions still contesting the legality of the PP measures (their constitutional challenge remained unresolved), the new government had strong incentives to accommodate demands for a rollback of the PP liberalisation measures. In late 2004, it adopted legislation stipulating that, as from 2005, the national law would return to a minimum allowable opening time of 72 hours a week, with the regions free to impose restrictions on opening, provided that they did not go below that floor. The minimum number of Sundays and public holidays was cut from 12 back to 8. Exemptions for small retailers and other exempt merchants were preserved, but here again, the regions were granted a little more power to adapt their regulations: small retailers were still defined in the national law as independent outlets of no more than 300m², but regional governments were authorised to adjust this definition upwards or downwards, as long as the threshold was not below 150m².

Conclusion

The 1993-96 legislation had left Spain less restrictive on opening hours than many continental countries at the time, largely because of the flexibility it left to entrepreneurs when it came to choosing their hours of work. However, Spain was re-regulating, even as many other European countries were liberalising. By the end of the decade, most Spanish regions had adopted regulations that looked exceptionally strict in a Continental context. More striking still was the trend towards the creation of new barriers to entry, which stood in stark contrast to European policies, including in recent years the Services Directive, the transposition of which into national law could become an important opportunity to liberalise and harmonise regulation of entry. A growing number of Spanish and international studies suggest that Spain has paid a price for such regulation. The productivity growth from economies of scale and scope could be substantial if regulations were relaxed, and some assessments find that less restrictive regulation would have positive employment effects and help to reduce prices.

One of the main reasons resistance to reform proved so successful in the early 2000s was that welfare gains such as these were more diffuse than the rents being appropriated by the protected sector, which was well organised and enjoyed a high degree of public sympathy. In any case, since the most severe restrictions concerned entry rather than day-to-day operations, regulation tended to benefit large numbers of incumbents who might otherwise have favoured liberalisation. The measures adopted also had significant unintended effects, which served in part to reduce the pressure for a more liberal regulatory regime: in particular, partial re-regulation after 1996 appears to have benefited the supermarket chains, in the first instance, rather than the small retailers that it was meant to protect. The chains thus had no reason to press for reform until the increasingly strict entry barriers established by the regions began to affect them, too. Moreover, the PP government, which attempted to relax retail regulation somewhat, made little effort to present the case for reform in terms of its growth or inflation objectives, nor did it persuade the general public that regional regulations were costly to households and often failed to correct the kind of externalities they were – in theory, at least – intended to mitigate. Because the costs of the status quo were principally opportunity costs, they were
largely invisible to the electorate: households saw little evident need for change. Making
the case for reform would have required significant investment in analysis, consultation
and communication, but no serious attempt was undertaken to tap expert advice or to
raise public awareness of the cost of entry restrictions. Public debate over the issue was
instead dominated by the interest groups involved.

Notes

1. On the Ministry of Industry’s assessment of the productivity costs of regulation in the
   early 2000s, see Cinco Días, 13 July 2007.
3. Specifically the National Federation of Trade, Hotels and Tourism of the Workers
   Commissions (FECOHT, affiliated to CC.OO) and the National Federation of
   Workers in Trade, Hotels, Tourism and Gaming (FETCHTJ, affiliated to the UGT).
5. Llorens (2000) cites estimates that 277 employees in a large supermarket can replace
   1 557 traditional shopkeepers.
7. For a recent statement of the CEOE’s position, see the Wall Street Journal Europe,
   2 August 2007.
10. In fact, empirical research in other jurisdictions suggested that other variables
    appeared to dwarf opening hours as a factor in determining prices See, inter alia,
    Kirchner and Painter (2000), Boylaud (2000); OECD (2001:79) and (2006:145); Høj
    et al. (2007); Bertrand and Kramarz (2002); and Schivardi and Viviano (2007).
11. Data from the Ministerio de Agricultura, Pesca y Alimentación.
13. OECD (2007). See, inter alia, Boylaud (2000); Høj et al. (2007); Bertrand and
    Kramarz (2002); and Schivardi and Viviano (2007).
15. Regional licences were introduced in the Basque Country (1983), Valencia (1986),
    Catalonia (1987), Galicia and Navarre (1988), Aragon (1989) and the Canary Islands
    (1994).
16. On the so-called “loi Royer” in France, see Bertrand and Kramarz (2002).


19. Stores of up to 500m² of useful floorspace for display and sale, open at least 18 hours a day and stocking, in similar measure, books, magazines, newspapers, records, videos, toys, gifts and sundry items.


22. Some regions did, however, impose a limit of 12 hours’ opening in any 24-hour period.

23. See, for example, the case study on retail hours in Germany.


26. OECD (2000b). In this respect, the Spanish regional licensing regime differed from the French system on which it was ostensibly modelled; see Bertrand and Kramarz (2002).


30. Including adjustments to permitted profit margins on different products, with a view to lowering the prices of some products and encouraging the sale of generics, and the removal of restrictions on discounts for over-the-counter medicines.

31. Giving bookshops and distributors freedom to give unlimited discounts on textbooks.

32. Including measures to foster competition in the downstream oil sector, to make access to oil infrastructure more transparent, further steps to liberalise the natural gas market in line with EC directives, and an acceleration of the on-going liberalisation of the electricity market.

33. Measures aimed primarily at easing market access in an already liberalised telecoms sector and ending Telefonica’s monopoly over local calls.

34. In addition to the sectors discussed above, the package affected such diverse areas as technical inspection of vehicles, governance of the internet and sick leave.


37. Effectively, the executive action has the status and force of law, provided that the lower house of the Cortes (the Congress of Deputies) ratifies it within 30 days. Decree-laws cannot be used in areas governed by “organic” laws, which affect the exercise of fundamental rights and public liberties. See Chuliá (2007).

38. For an overview of regional legislative changes over time see Banco de España (2007).

41. AC Nielsen data. Other studies also find falling market shares for both hypermarkets and small outlets and growth among supermarkets; see de Jorge Moreno (2008a, 2008b).
42. OECD (2003, 2007). All but one of these moratoria had been ended by mid-2006, but the end of the moratoria did not always lead to much easier entry.
44. In many regions, an outlet was defined as “large” if it was at least 25% owned by a large firm.
45. These are described in some detail in Hoffmaister (2006) and Orea (2008).
46. Aragón, Asturias, Catalonia and Navarre; Aragón’s levy was imposed only in 2006 and is therefore not shown in Table 2.1.
49. For details, see the case study of Germany’s Ladenschlussgesetz.
50. In fact, article 4.1 of the law stipulated that the minimum number of Sunday/holiday openings per year would remain twelve, but article 4.2 allowed the regions to various this number downwards, to a minimum of 8 such days, which thus became the de facto minimum.
52. See ICE (2006); Cruz and Oubina (2006); Múgica (2006); van Ark et al. (2002); Høj et al. (2007); Orea (2008).
Bibliography

Banco de España (2007), “Una aproximación a la regulación del comercio al por menor a partir de indicadores sintéticos”, Boletín Económico, October.


CHAPTER 22. SWEDEN: POSTAL LIBERALISATION, 1992-99 – 477

Chapter 22

Sweden: Postal liberalisation, 1992-99

In January 1993, parliament abolished the state-owned postal operator’s monopoly on the collection and delivery of letter mail items. Legislation adopted at the beginning of the following year transformed the Postal Administration (Postverket) into a joint-stock company, Posten Sverige AB (hereafter simply “Posten”), and created a new legal and regulatory framework for the postal sector, with a view to fostering competition. Since the parcel business had never been subject to a public monopoly, these changes created the world’s first wholly liberalised postal market. Postal liberalisation followed several years of study and discussion, as well as the appearance of a new entrant determined to challenge the monopoly. Although the authorities moved rapidly to open up the market, the initial liberalisation was followed by several years of regulatory innovation and adjustment, as Posten and its emerging rivals adapted to the new conditions and the authorities sought to create a genuinely level competitive playing field in the postal sector. A number of factors contributed to a relatively smooth liberalisation of postal services:

- A policy context open to reform. Two decades of disappointing economic performance, culminating in the severe economic contraction of the early 1990s, gave rise to a willingness to reconsider a wide range of established institutions and practices in an effort to increase potential output. Sweden began to restructure and open up network monopolies like rail, telecoms and power, relying on generous social protection to reduce the associated adjustment costs. As a result, there was already a generally favourable disposition to reforming network sectors when events pushed postal services onto the policy agenda. At the same time, technological and regulatory changes abroad were prompting a large number of countries to begin examining options for postal reform.

- Pressure for reform from the incumbent. The management of the Postal Administration had long wanted – and had gradually been winning – greater freedom to operate on a commercial basis and to move into new sectors and markets. However, postal managers believed that, sooner or later, they would have to surrender the legal monopoly in order to gain the freedom they sought. The value of the statutory monopoly was in any case limited: it covered only a fraction of the postal services market, its limits were ill defined, and its value was being eroded by technological change and the emergence of new forms of information transfer, as well as by changes in the composition of the mail market itself.

- The entry of a new competitor. While postal reform had been under discussion for several years, it was the emergence in early 1991 of a new entrant, CityMail – in apparent violation of the monopoly – that forced the authorities to confront the issue. The Postal Administration was prepared to give up the legal monopoly,
though it did protect its markets from CityMail by other means, apparently confident that it could see off new rivals without statutory protection. The unions, for their part, were concerned about the potential effects of postal reform, but since they represented CityMail’s workers as well as those of the Postal Administration, they could not support enforcement of the monopoly.

- **An adaptive approach.** The initial reform was relatively simple, as most contentious issues were left to be resolved by the market participants, the newly created Competition Authority or the sectoral regulator in response to events as they unfolded. The reform legislation underwent several revisions in the late 1990s, sometimes to address issues that critics argued could have been better regulated beforehand. However, this approach made initial adoption of the reform easier: the incumbent might well have resisted reform if the initial plan had included some of the requirements that were subsequently imposed on it with respect to issues like infrastructure access. Moreover, the reform was without precedent, so any attempt to regulate all the relevant issues *ex ante* would probably have failed: many problems that received much discussion *prior* to liberalisation never materialised, while the issues that actually did emerge were sometimes unanticipated.

**The context for reform**

*Framework conditions for reform*

When events pushed postal reform onto the political agenda in 1991-92, Sweden was in the midst of its deepest recession since the 1930s: real GDP fell by a cumulative 5% from 1990 to 1993 and gross fixed investment by about a third, while open unemployment jumped from 2.1 to 11.3%, despite a sharp drop in the participation rate. While the recession, moreover, struck at a time when Swedish policy makers were already worried about long-term growth performance: having grown considerably faster than the OECD average prior to 1970, Sweden had fallen below it in the 20 years since. While postal liberalisation was hardly seen as the solution to Sweden’s growth problems, this disappointing growth performance had fostered a willingness to reconsider a wide range of established institutions and practices in an effort to increase potential output. Growth concerns also reinforced a trend towards deregulation that had begun in the 1980s with the relaxation of banking and foreign-exchange regulations. This was followed by vertical separation in the railways sector and the liberalisation of taxis and air transport. Telecoms liberalisation was already under way and was perceived to be going well, and work had begun on a new competition law, which was adopted in July 1993. The prevailing model of combining openness to competition with generous social protection provided a broadly favourable background to postal liberalisation: rather than using regulation to forestall structural change, the Swedish authorities increasingly preferred to open markets up and rely on social policies to reduce the ensuing adjustment costs. The trend towards reform of network sectors, which had begun under governments led by the Social Democratic Party (SAP), was reinforced by the election in late 1991 of a centre-right coalition committed to competition and deregulation.

At the same time, changes in the postal market itself were generating pressure for change. First, the rise of new forms of communication, such as fax and e-mail, raised the prospect of increasing competition from outside the postal sector. Secondly, technological change was creating new possibilities for the organisation of postal production. Thirdly, demand for postal services was shifting from correspondence...
towards mass communication: volumes of addressed and unaddressed bulk mail dwarfed those of individual letters. This changed the nature of Postverket’s core business: the monopolist’s principal product was no longer the processing of individual correspondence but its ability to offer large senders access to every address in Sweden.19 It was this market segment that attracted would-be entrants. The shift towards mass mail also meant that the competitive pressure on Postverket came less from courier services than from other advertising media. Further pressure for change stemmed from the exploding volume of international re-mail operations,10 which put liberalisation on the agenda of many states, as well as the European Community (EC). EC states began exploring the feasibility of opening up postal services to greater competition, while agreeing some common regulations. Sweden was committed to entering the EC, so the prospect of postal reform in Europe constituted a further strong argument for reform. In the end, Sweden actually went much further than Europe, particularly in choosing not to retain any “reserved area” of the market for the universal service provider.

**Actors and interests**

Discussion of postal liberalisation began under an SAP government, and it was implemented under governments of both centre-right and centre-left. Thus, while there were parliamentary clashes over the details and timing of postal reform, it was never a major point of contention between the SAP and its rivals on the centre-right, or even an electorally salient issue; both left and right accepted the need for productivity increases in the public sector.11 However, postal reform was not seen as a top priority by any major party.

From an early stage, the most important agent pushing for change was the monopolist itself. The Postmaster General had called the letter monopoly an anachronism as early as 1982, and by the mid-1980s, Postverket’s management was pressing for greater freedom to operate the post as a business, free of political interference, and to restructure it as a corporation. Postal managers argued that Postverket would face increasing competitive pressure regardless of what the government did, as a result of technological developments and changes in the postal market, and they wished to restructure Postverket in response to these challenges. Corporatisation would also give them greater authority over salaries, which they did not control as long as postal managers and employees were civil servants. They feared that the statutory letter monopoly, which was of limited value, might get in the way of other business options, including entry into foreign markets and the possible expansion of postal banking.12 Yet it was the appearance of the privately owned CityMail that really forced the issue. While CityMail did not have anything like the political weight of Postverket, its challenge to the monopoly in 1991 put postal reform on the political agenda and was thus welcomed by some Postverk managers, even if the monopolist did its best to squeeze its new rival out of the market.13 Moreover, CityMail’s very business model – relying on new technology to undercut Postverket in the bulk mail market – appeared to confirm the arguments postal managers had been making for a decade about the advent of competition and the consequent need for reform.

Postal liberalisation was viewed with suspicion by the Swedish Trade Union Confederation (LO) and the Union of Service and Communications Employees (SF), both of which feared the consequences of liberalisation for the pay and working conditions of postal workers. They were worried, in particular, about the potential loss of civil service status by Postverk employees. However, the SF’s position was complicated by the fact that it also represented CityMail employees, so it could not easily oppose the end of the
letter monopoly, which, if enforced, would put a small but significant number of its members out of work immediately.

Although advocates of reform insisted that competition would bring greater efficiency and would thus benefit consumers, consumer interests had relatively little impact on the early stages of the reform. Households and small businesses were the most numerous consumer groups, but their share in postal demand was modest and declining, the relative weight of postal services in their expenditure was generally small, and they were not organised. In any case, both households and business clients rated Posten’s services highly, tariffs were fairly low by European standards, and Postverket did not appear to be making “super-profits”, so consumer discontent was not really a factor. The government’s commitment to maintaining universal service further allayed concerns they might otherwise have had. However, the issue was of potentially enormous importance to bulk mailers, for whom postal services accounted for a large share of costs. Thus, the most active consumer lobby was the association of direct-mail merchants; other sectors for which postal services were a first-order concern, such as direct marketing firms, were less involved but still supportive of reform.

The Swedish Employers’ Confederation was in favour of opening the market but did not lobby strongly for reform.

Finally, two key players came into being during – and, in one case, as a result of – the implementation of the reform: the Swedish Competition Authority, formed under the July 1993 competition law, and the Post and Telecom Agency (PTS), the new sectoral regulator created under the Postal Services Act in 1994. Both subsequently played important roles in shaping the reform as it was implemented over the course of the years that followed.

**Issue history**

By the late 1980s, Postverket was making progress with respect to its commercialisation agenda: quality and profitability objectives were introduced, and Postverket was given greater freedom to adjust postal rates and to finance investment from its own funds rather than the state budget. Although a proposal to open up the postal sector was rejected by the government in 1987, due in part to trade-union objections, the government began studying the issue seriously. The management of Postverket continued restructuring and began shedding staff in the expectation that it would eventually face direct competition. Its attitude to the letter monopoly, which was meant to compensate it for the provision of universal service at uniform tariffs, owed much to the peculiarities of Swedish legislation. The monopoly was limited to daily collection and delivery of personal written correspondence. This was unusual: most statutory postal monopolies were defined in terms of size or value. It was also vague, since the definition of “personal written correspondence” had never been clarified. No one knew whether it covered addressed bulk mail, which accounted for around half of the addressed mail market. Postverket itself claimed in 1990 that the legal monopoly covered as little as 8% of all letters. Other estimates suggested that it protected anywhere from 25 to 60% of Postverket’s business. However interpreted, the share of the postal market that the law reserved to Postverket was declining. Yet Postverket saw the monopoly as a restraint on its commercial freedom and a potential obstacle to its international ambitions, believing that few governments would be keen to open markets to a foreign state-owned enterprise that enjoyed a legally entrenched monopoly at home.

One reason why the 1987 liberalisation proposal had so little chance of success was that it was not supported by the preparatory studies and consultations that are normal in
the Swedish policy process. However, during 1989-92, a series of official reports appeared, exploring Postverket’s status, the competitive pressures arising as a result of technological and economic change, and the prospects for opening up at least some segments of the market to greater competition. These studies did not propose anything very radical. A government commission on telephone and postal deregulation argued that the regional and social policy requirements imposed on Postverket should remain unchanged and that Sweden should simply await further developments in the EC. The Ministry of Communications concluded that the provision of a universal postal service was indeed a natural monopoly and estimated that costs could rise by up to 20% in a market with two players. It also argued that increased competition in those sub-sectors that were not characterised by a high degree of natural monopoly was dangerous, since competitors could “cream-skim” the most lucrative segments of the market, making it harder for Postverket to finance its universal service obligation (USO). The report therefore concluded that Postverket should maintain daily nationwide service with cross-subsidies to finance delivery to remote areas. Nevertheless, the ministry, perhaps reflecting the views of Postverket, also recommended ending the letter monopoly. Other studies argued for revising it to create a clearer (weight- or value-defined) “reserved area” for Postverket, but the monopolist itself never expressed any interest in reformulating the statutory monopoly.

Despite the “business-as-usual” conclusions advanced in most of these studies, the SAP government showed increasing interest in liberalisation from 1990. In its 1991/92 budget bill, it proposed a review of the mail monopoly, indicating that a new postal law, together with a more stringent fair trade law, could stimulate competition and efficiency to the benefit of consumers. The government also favoured corporatising Postverket. However, before any of this could happen, CityMail entered the market. In May 1991, it began offering delivery of pre-sorted bulk mail in and around Stockholm. CityMail relied on the fact that customers could use new technology to print bulk mail in postcode sequence to carrier-route level – in other words, to produce addressed bulk mail already sorted and ready for delivery. This enabled it to offer lower prices than Postverket, which did not yet offer a pre-sort discount; the discount for non-pre-sorted bulk mail was just 4%. CityMail made deliveries to any given area only once every three days, so it avoided the fixed costs associated with daily delivery, and it also had greater freedom than the monopolist to adapt its operations to customer demand. Postverket reported CityMail for suspected violations of the letter monopoly and a preliminary investigation was launched to determine whether or not it had violated the law. Before the issue was resolved, CityMail was in bankruptcy and a new government, committed to deregulation, competition and reducing the size of the public sector, had come into office.

Policy process: formulation

The search for solutions

The new centre-right coalition took the view that it was not reform that required a clear rationale but the status quo: the statutory monopoly should go unless there was a clear public interest rationale for restricting competition in this way. The coalition appointed its own commission to prepare for deregulation and to examine, in particular, the scope and cost of the USO, and the ability (or not) of Postverket to compete “on equal terms” if certain segments of the market were opened. The cost of maintaining universal service at uniform tariffs was expected to be high, given Sweden’s large size and low population density, as well as the problems posed by delivery to some particularly
difficult-to-reach locales. Deliveries to these areas were financed by cross-subsidies within Postverket, and there was a fear that opening the market might send the monopolist into a downward spiral of lost volumes that would – unless direct subsidies were forthcoming – eventually force it to begin curtailing services to rural areas. More specifically, new entrants could capture the most profitable segments of the business by pricing below the incumbent in low-cost markets. The loss of these market segments would then force Postverket to raise prices for some of its remaining activities, thereby making previously unprofitable volumes in those activities attractive to entrants. The newcomers would then capture more volumes in these markets by pricing below the monopolist, and so on – a process that critics of liberalisation referred to as the “graveyard spiral”. However, Postverket insisted that it could compete effectively and that leaving the monopoly in place would achieve little.

The new commission was due to report in early 1993. In the meantime, however, the plight of CityMail constituted a source of pressure on the government to act quickly. The company entered bankruptcy in October 1992, while the legal dispute over its alleged infringement of the statutory monopoly was still unresolved. By December, it was in a position to re-enter the market with new capital, provided the legal position could be clarified and its right to conduct its bulk-mail operations confirmed. Investors were understandably wary of investing in conditions of legal uncertainty, and many commercial consumers were unwilling to risk using CityMail’s services if they might later be found illegal. Given press sympathy for CityMail, the situation threatened to become an embarrassment for the government, particularly in view of the fact that that deregulation was a central theme of its first budget bill. Without waiting for the commission to report, therefore, the government shifted its position and called for ending the monopoly – even if the question of the USO remained to be resolved later – in effort to give a clear signal to the market that CityMail would be free to enter.

The reform

The government’s decision to move rapidly resulted in a two-stage reform. The first effectively ended the post office’s legal monopoly but provided no other legislation to replace it. This was followed by a consultative procedure leading to the second phase of liberalisation, implemented in the March 1994 Postal Services Act. Under the Act, Postverket, a public agency, was to become Posten Sverige AB, a wholly state-owned corporation, on 1 March 1994. There were no plans to privatise it, nor did any government subsequently put such plans forward. The Telecommunications Agency was to become the Post and Telecom Agency (PTS), taking over supervisory responsibilities in the postal market. Guidelines for the PTS would be set by parliament and communicated to the PTS via annual “regulation letters” from the government.

The legislation gave the PTS a general obligation to promote the smooth functioning of the postal market, as well as to supervise compliance with the Postal Services Act and regulations adopted under it. It was also assigned a consumer-protection mandate and the role of competition “watchdog” for the sector: the newly formed Competition Authority was to be responsible for the enforcement of the Competition Act, but the PTS was charged with informing the Competition Authority of any suspected violations it observed in the performance of its supervision activities. The PTS was not, however, given a competition-promotion role in respect of postal services similar to that assigned to it in the field of telecommunications; the PTS lobbied for such a role but Postverket was opposed. The VAT exemption on Post Office deliveries was ended, with the full 25%
VAT being phased in over two years. In an effort to protect small consumers, the Act provided for a price cap for individual domestic letters (up to 500g), a market segment in which little or no competition was expected. The government took the view that Posten’s dominant position in the market for individual mail necessitated some degree of regulation.

Under the Postal Services Act, the government, rather than Posten, assumed responsibility for the USO. The government then contracted with Posten to provide this service, but it was free to extend this arrangement to other postal operators if they achieved sufficient scope of operations. Under the USO contract, Posten was obliged to provide nationwide delivery “from all to all”, Monday through Friday (with some exceptions), to convey single letters at uniform and reasonable rates, and to maintain a network of physical postal counters. Posten was not required to maintain a specific number of retail outlets, but the PTS was charged with monitoring closures. This issue was in any case less sensitive in Sweden than in many countries, because there were no post offices at all in many very remote areas: postmen in such areas did not merely deliver the mail, they also provided, on a mobile basis, the kind of counter services that would have otherwise required the maintenance of a large number of very small post offices in rural districts. The bill also provided for compensating Posten for such social services as the free distribution of electoral mail and free delivery in rural areas for the elderly and disabled, activities regarded as distinct from the USO.

Policy process: adoption

The first phase of reform took little time: on 22 December 1992, the Riksdag approved a brief amendment to a lengthy communications bill that abolished the state mail delivery monopoly as of 1 January 1993. By amending a bill already under consideration in parliament, the government short-circuited the usual legislative process, which would ordinarily have involved extensive study and consultation with both parliamentarians and stakeholders. The government wished to resolve the issue before the legal investigation of CityMail could resume, and there was little resistance to the idea in parliament. The main source of uncertainty stemmed from the position of the Centre Party: though a member of the governing coalition, this predominantly agrarian party was very concerned about rural services and employment. In the end, however, it voted with the government rather than risk the unity of the first non-socialist coalition in nearly a decade over an issue that it did not regard as being of primary importance. The SAP, now in opposition, voted against the December bill but did not in principle oppose deregulation, which it had initiated two years earlier. It focused its attacks on the manner and timing of the government’s move rather than its direction; this was critical, since it meant that the SAP was not committed to reversing the reform when it returned to office in 1994. With the opposition Left Party staunchly opposed to market opening, it was the New Democrats who, though not part of the coalition, provided the minority government with the votes it needed to pass the bill, taking a straightforward position in opposition to monopoly.

There was remarkably little opposition outside parliament to the first stage of the reform. Since it represented CityMail workers, the SF could hardly defend the statutory monopoly, and the unions focused their energies instead on shaping the subsequent legislation governing the de-monopolised market and on negotiating directly with Posten to ensure that market opening did not damage the interests of Postverk employees. The SF and the LO therefore lobbied hard against the corporatisation of Postverket, and some
demonstrations against it were organised in Stockholm, but in the absence of any privatisation proposal, the unions could not arouse much public interest. For the unions, corporatisation was a critical issue in its own right, since it threatened the status of postal employees as civil servants, but for most Swedes the question of whether the post was delivered by a state agency or a wholly state-owned company was hardly of great import. The unions also challenged the claim that competition would necessarily increase efficiency and argued that the monopolist must be given a substantial “reserved area” in order to offset the costs of the USO.40

On the latter issue, however, the unions were isolated, with the governing parties and the monopolist all convinced that no such reserved area was needed. The former issue was largely resolved via bargaining between the unions and management: Posten was able to negotiate increased flexibility with respect to working times in return for very generous arrangements for incumbent employees and a commitment to maintaining full-time employment wherever possible – a shift to greater reliance on part-timers was perhaps the unions’ greatest single concern. CityMail’s presence on the market strengthened Posten’s bargaining position at this point: the challenger actually paid slightly higher rates than Posten for many jobs but it had negotiated much more flexible arrangements for the deployment of its workforce, and Posten was able to follow suit.41

The second stage of the reform took far longer to reach the statute books, so the postal market operated in a legal vacuum from 1 January 1993 until 1 March 1994, when the Postal Services Act entered into force. The report of the government commission on post and telecommunications reform formed the basis for consultations on this bill. Posten itself endorsed most of the report, which it had helped fashion. Some 43 other organisations submitted opinions on the report, ranging from CityMail and the unions to the philatelists’ association.42 The most contentious issues concerned access to physical and informational infrastructure:

- CityMail pressed for regulation of access to physical infrastructure, but Postverket strongly resisted legal guarantees of third-party access – particularly to post office boxes.
- CityMail and the Competition Authority expressed concern about Posten’s control over the system of postcodes, arguing that all operators on the market should share in this responsibility. Other state agencies suggested that the PTS should handle it.
- Historically, Postverket had provided a one-stop service for changes of address: anyone moving simply notified the post office, which was responsible for informing a number of other state agencies. In a competitive postal market, it was critical to ensure that all postal operators had access to change-of-address data.

CityMail also raised concerns about the proposed solutions regarding management of issues like mail forwarding. In the end, Posten’s views prevailed on all these issues, largely because the PTS and the government chose to leave such matters to market participants to resolve among themselves and to regulate only as and when it became apparent that they could not agree on satisfactory solutions. Since the reform was without precedent, the PTS argued against over-regulation at the outset and in favour of a wait-and-see policy. This hands-off approach also applied to entry: proposals for licensing were dropped and new entrants needed only to report to the regulator.

Parliamentary consideration of the Postal Services Act in December 1993 revealed that the positions of the parliamentary parties were essentially unchanged. The SAP did
not oppose the reform in principle but neither would it support the government: it accepted corporatisation of Postverket and made no attempt to restore the monopoly, focusing its opposition instead on specific details, such as the lack of licensing. The Centre Party wavered but voted with the government following assurances about rural services, and the New Democrats once again provided the government with the votes needed to assure passage of the measure, following a strong joint appeal by Postmaster General Ulf Dahlsten and CityMail CEO Bror-Anders Månsson, both of whom had emerged as strong public advocates of reform.

**Policy process: implementation**

When the market was opened, the authorities did not expect more than a handful of new entrants, but by 1997 there were over 100 operators registered, of which close to 80 were active. By 2002, fewer than 40 remained. Nevertheless, the impact of reform on market structure was limited: Posten’s market share has remained above 90% since liberalisation, and CityMail was the only new entrant to emerge as a significant player (Figure 22.1). The other new companies on the market were very small operators engaged in local activities; altogether, they accounted for around 0.3% of the market in 2007. Around 60% of Swedish addresses are still served by Posten alone. There were no obvious losers in the early years of the reform, since, contrary to expectations, postal volumes grew until the end of the decade, when they finally levelled off and then began to decline slightly. This ensured that there was no backlash against the reform after it was adopted. However, implementation of the reform was not entirely smooth. On the contrary, passage of the Postal Services Act was followed by a period of experimentation, conflict and adjustment as the postal operators and the authorities adapted to the new conditions.

![Figure 22.1. The growth of CityMail](image)

1. Market share percentages are shown above corresponding volume bars.

Posten’s defence of its markets

The widely predicted “graveyard spiral” anticipated by opponents of de-monopolisation never materialised. Posten recognised that preserving both its monopolistic position and continued high margins was not an option. It opted instead to defend market share, while accepting lower margins and raising productivity. Moreover, it turned out that the unit cost of delivery to rural areas was not necessarily costlier than in many urban areas, so the USO did not prove to be as great a burden as anticipated. Indeed, Posten later argued that the competitive advantage conferred by being the only postal firm that served all addresses in Sweden was worth the cost. It never received a direct subsidy or other compensation for the USO.

The central problem thus turned out to be Posten’s dominance rather than its inability to compete. Indeed, CityMail’s subsequent history – the company went through two bankruptcies and several changes of owner in the 1990s – highlights the enormous difficulty that new entrants faced in sustaining the losses of their early years before they became established. While CityMail’s difficulties were partly the product of industrial customers’ reluctance to use a new and little known supplier, Posten’s use of its dominant position to defend market share also played a very important role. Posten quickly began offering loyalty rebates, exclusivity clauses, selective pricing and annual bonus discounts. This led to legal challenges and a number of interventions by the Competition Authority, which dealt with just over 100 cases involving Posten during the first years after liberalisation. Posten also offered price cuts in competitive segments of the market. In 1996, it introduced a four-zone pricing model for mail not subject to the USO, which was rejected by both the Competition Authority and the Market Court as an abuse of dominant position. A subsequent two-zone model was eventually allowed by the court to stand, subject to specific criteria governing the determination of the discount size.

To some extent, the various competition-related rulings against Posten may have been inevitable consequences of Posten’s testing of the limits of legitimate competition in the first years after liberalisation; the Competition Authority has not made a finding against the monopolist since 1998. However, critics argued that Posten fought competition disputes that it knew it would lose under Swedish or EU law, simply in order to disadvantage its rivals, who were less able to bear the cost of protracted legal battles. Competition disputes may thus have served to reduce new entrants’ staying power regardless of the outcomes.

Posten tried to finance price cuts in competitive market segments by raising prices in non-competitive segments. Despite the imposition of the price cap on single domestic letters, the price of domestic overnight 20g letters rose by 90% over ten years, far above the rate of CPI inflation. Posten achieved this thanks to a loophole in the initial formula for the price cap. The monopolist was obliged to issue a certain number of discounted stamps to households, and their weight in the calculation of the price level subject to the cap was 90%. By introducing a small cut in the price of the discounted stamps, Posten could raise the price of all other single-letter stamps very substantially while holding the average rate of growth below the growth of the net price index. When the sale of discounted stamps was scrapped in 1998, households suddenly ended up paying the higher “standard” tariffs. Posten’s exploitation of the loophole attracted fierce criticism, but the company argued that without it, the necessary adjustment of relative prices between individual and bulk mail could not have taken place, and larger customers would have continued to subsidise smaller ones. In segments where Posten faced neither price
caps nor competition, prices rose even faster than in the domestic letter market. Posten discouraged rivals from entering those markets, or customers from leaving them, by offering significant discounts to key customers. For example, non-pre-sorted bulk mail was offered to large customers at a significant discount; this discouraged them from sorting their own mail, which would have opened them up to CityMail’s competing offers.

Feedback and adjustment

The Postal Services Act was amended three times in the late 1990s, in response to changes in EU legislation, demands from the PTS and the Competition Authority, and pressure from CityMail and its large bulk-mail consumers. In January 1997, the previous registration procedure for businesses wishing to engage in postal operations was changed to a licensing procedure administered by the PTS. This was largely in order to bring Sweden into line with the rest of the EU but it was also a response to the unexpectedly large number of small new entrants: while there had been no major abuses arising from their appearance, policy makers wanted better monitoring in future, particularly given the issues of confidentiality involved in handling private correspondence. However, the new arrangements did little to restrict entry: licensing fees were kept low and the other requirements were not onerous. Indeed, the PTS argued that, far from being an entry barrier, licensing was a form of “entry assistance”, since it gave the regulator a chance to educate potential entrants about the market and the regulatory regime – a form of free start-up consultancy.

A proposal to levy a small fee per letter on all postal companies to finance the USO was rejected, largely because, given Posten’s continued dominance, it would raise too little revenue to be worth administering.

A second set of amendments in 1998 followed a ministry report examining the state’s role in the postal sector. The new legislation revised the price cap on single letters: discounted stamps for households were phased out and price increases were limited to the three-year moving average of CPI inflation – a measure Posten strongly protested, pointing out that its own cost structure bore little resemblance to the CPI. This made the cap harder to evade and it also extended its effective coverage: the weight of discount stamps in the price formula meant that the old price cap had protected only households, but the new cap protected all senders of single letters. This, in turn, led to conflicts between Posten and the PTS over the definition of “bulk mail”, as the tighter price regulation gave the monopolist an incentive to stretch the definition of “bulk” as far as possible. Most of the disputes between Posten and the regulator over the definition of bulk mail concerned discounts that the PTS regarded as discriminatory. The PTS’s concern was to protect small operators from unfair competition. It insisted that the price cap was needed not only to ensure cheap postal services for small consumers but also to prevent Posten from cross-subsidising competitive segments of its business by raising tariffs on single items, where it had a monopoly.

The third set of amendments, in July 1999, aimed to make access to postal infrastructure fairer to all parties. Initially, many of these issues appeared to have been resolved among the market participants themselves, as the PTS had hoped: CityMail and Posten had concluded an infrastructure contract agreement in 1993 and had gone on to form a jointly owned address file company, SvAAB. By the end of the decade, however, CityMail’s new owners were at odds with Posten over infrastructure access and the monopolist had taken full control of SvAAB. Moreover, conflicts had arisen over the postcode system, particularly after Posten introduced changes to the system in 1996, prompting complaints from competitors, addressees and other private and public bodies,
which made use of postcodes for administrative purposes and which had experienced
difficulties obtaining the number series for their own needs. The 1999 amendments
required a licence-holder who owned facilities for the transfer of mail to make them
available to other licence-holders on terms that were not less advantageous than those
available to the infrastructure owner. In other words, Posten had to show the regulator
that it was not charging other postal operators more for access to post office boxes and
other infrastructure than it was “charging itself”.70 While a proposal to give control of the
postcode system to the PTS was dropped, Posten was obliged to consult other operators
before changing the system and to give them the required series of postcodes.71

**Public communication and postal reform**

Although postal reform was covered in the press, it was not a highly salient issue at
the time of its adoption, and public awareness of what was happening was limited. No
real attempt was made to explain the reform to the electorate, and many ordinary Swedes
became aware of it only when they began to see postmen in unfamiliar uniforms making
their rounds. Since a price cap was in place and there was no competition for individual
letters, the reform had little immediate impact on households anyway, and there was not
much public reaction to the reform during its first decade. The early 2000s, however, saw
a reaction against it: survey data in 2005 suggested that almost 90% of the population
“wanted the old Post Office back”.72 This dissatisfaction had little to do with postal
services. Rather, it reflected the closure of local post offices in connection with the
elimination of Posten’s cashier services, on which Swedes had traditionally relied.73 The
1994 legislation had required Posten to maintain basic cashier services nation-wide, but it
was allowed to shed this line of business after 2002. Postal services *per se* were in many
respects more readily available than ever: service counters offering basic stamp and mail
services appeared in small shops, service stations and kiosks, as well as outlets offering a
wider range of services in larger retail stores and supermarkets.74 Many of the new outlets
had longer opening hours than the old post offices.75 Nevertheless, the disappearance of
Posten’s cashier services, which had little to do with postal operations, proved deeply
unpopular and coloured attitudes to postal reform.

**Conclusion**

The opening of the Swedish postal market has so far had remarkably little impact on
its structure: Posten’s dominance has never really been in doubt, prompting some
observers to question whether the reform has made a difference. However, the reform had
a big impact on Posten’s structure, operations and pricing, all of which show the impact
of competition.76 The initial liberalisation triggered a sharp increase in its productivity: in
1993-94, the volume of addressed mail delivered per employee rose by around 20%.77
Productivity measured in these terms rose by 32% during 1993-2000, suggesting that
productivity growth reverted to rates typical of the economy as a whole after the initial
adjustment.78 Posten’s introduction of its own pre-sorting discounts led rapidly to the
stage where most bulk mail was pre-sorted, and the monopolist moved to outsource some
activities, including transportation, while entering new segments of the market, such as
running other firms’ internal mail departments.79 Overall, postal prices appear to have
become more closely linked to costs after the reform. While single-piece letter rates rose
sharply during the first decade after market opening, the standard price of sorted bulk
mail – the market segment where Posten faced competition from CityMail – rose only
marginally, and sorted bulk mail with individual discounts was actually cheaper than it
had been ten years earlier.80 Cohen *et al.* (2007) estimate the increase in consumer surplus
during 1990-2004, based only on Posten’s letter prices and volumes, at SEK 1.75 billion at 2004 prices. However, this decomposes into a loss of SEK 623 million for senders of single letters and a gain of SEK 2.4 billion for bulk mail – a reflection of the elimination of cross-subsidies and a further reason, perhaps, why Swedish households have not been enthusiastic about postal reform.

The structural factors driving reform in Sweden were common to many countries in the early 1990s: technological change and shifts in the structure of demand for postal services created new business opportunities for postal operators and exposed existing public monopolies to new competitive pressures. Even if regulatory frameworks remained unchanged, they would face growing competition from other mass media and from new technologies for information transfer. At the same time, the explosion of re-mail reduced the ability of national governments to shield public postal monopolies from emerging competitive pressures. In many other countries, however, these factors did not result in swift or radical postal reform. That they did so in Sweden reflects the impact of a number of factors more specific to the Swedish case: a policy context that was unusually open to competition-enhancing regulatory reform, thanks both to the immediate crisis and to the prominent role played by economists in public policy debates, the aggressive promotion of reform by the management of Postverket itself, the entry of CityMail and the peculiar character of the pre-existing statutory monopoly, which was both limited and vague. The expectation of reform in the EC also played an important role in shaping the Swedish debate, though in the end, European action on the issue was both later and more modest than expected, and Sweden in the end went much further than Europe required.

The simplicity of the initial reform also contributed to its smooth adoption: many of the most contentious issues were left to be resolved later. This approach later prompted some criticism, since a more elaborate regulatory framework and a stronger sectoral regulator at the start might have prevented some problems and allowed competition to develop more quickly. However, a more prescriptive approach at the outset would have generated more resistance, particularly from Postverket, which preferred the “hands-off” approach taken by the early legislation in respect of issues like infrastructure access. It might also have strengthened opposition from the unions, which could not really oppose the abolition of the monopoly but which did resist other aspects of the reform. The initial liberalisation, moreover, was conducive to further reform, since much subsequent pressure for change came from the regulator and the competition authority, together with the new entrants in the market – actors who would not have been engaged in the process had the initial opening not taken place. The reform was in any case without precedent, so an attempt to regulate all the relevant issues ex ante would probably have failed: many problems that received much discussion prior to liberalisation never materialised, while the issues that actually did emerge were sometimes unanticipated.

Notes

1. The monopoly did not cover parcels, unaddressed items or daily newspaper distribution, three markets that are related to, but distinct from, the mail market. Nor
did it encompass “printed” (mass duplicated) material, as opposed to “written” material. See Andersson et al. (2006b:1-2) for details.


3. By 1970, GDP per capita stood at about 90% of the US level, before a long relative decline set in. See OECD (2008) for more on Sweden’s long-term growth record.

4. The importance of the overall economic context is emphasised in numerous Swedish analyses of the reform, including Andersson (2001); Andersson and Thörnqvist (2006); and Brandt (2007).

5. For an overview of the Swedish model, see OECD (2008:20-23).


7. Though e-mail was still in its infancy, Postverket estimated that the fax had, by the early 1990s, taken over 35% of letter volumes.

8. Computerised production and automatic sorting, in particular, raised the prospect of the dis-integration of postal operations, raising the question of the extent to which it remained (if it ever had been) a natural monopoly.

9. Private letters sent by households constituted just 12% of the addressed mail market; see Bladh (2001).

10. Re-mail is mail from one country that is posted in another country by a private operator or a foreign postal administration instead of the originating country’s own postal administration. The primary beneficiaries of re-mail services tend to be large-volume users like financial institutions, catalogue companies, advertisers and publishers. Re-mail opportunities can arise due to the differences in compensation arrangements between postal administrations for delivering inbound international mail. In 1986, a shift in US policy effectively ended decades of attempts to suppress re-mail under the Universal Postal Convention, which had repeatedly authorised interception and/or surcharging of re-mail. See Ghosal (1999); and Boeri et al. (2007).


12. Andersson (2006b). Prior to liberalisation, the Postal Administration was considering merger with a Swedish bank; in the event, Posten AB decided in 1999 to minimise its financial services activities and focus on messages and logistics; in 2001, it sold off Postgirot Bank AB (giro payment services) and spun off its cashier services into a wholly owned subsidiary as a prelude to withdrawing from cashier services entirely.


15. Other sectors, like banking, accounted for a large proportion of the demand for postal services but did not mobilise over the issue, because such services constituted a relatively small part of their costs.

16. Up to 80% of the “net price index” (i.e. the CPI excluding price changes arising from changes in indirect taxes and subsidies).


21. It is impossible to be precise, because the definition of “bulk mail” has never been precisely specified and because there was little incentive to track specific trends concerning individual items, addressed bulk mail, unaddressed bulk mail, etc, as long as Postverket faced no competition. See Andersson (2001), chapter 3.

22. If the law was construed as narrowly as possible, excluding not just bulk mail but all printed correspondence.

23. This variation in part reflected different interpretations of the law, but it also reflected the dearth of good data on the structure of the market prior to reform.

24. Studies were commissioned by the Riksdag (parliament), the Price and Competition Board, the Post and Telecommunications Commission, the Ministry of Finance and the Ministry of Communications.

25. See Falkenhall and Kolmodin (2005:6-7) or Andersson (2001:121-3) for brief overviews of these studies.

26. Jonsson and Selander (2006); CPC (2008). At the time, the USO was referred to as “regional and social responsibilities”.

27. The apparent contradiction was not explained. Andersson (2006b:19) observes that the argument cannot rest on the alleged natural monopoly character of the postal business; otherwise, a new entrant, facing the same cost structure as the Post Office but without its “regional and social responsibilities” would quickly take over the market. The argument rested rather on the high cost of entry.

28. Postverket was arguably lagging in this respect: postal services in many other countries had been offering substantial bulk-mail and work-sharing discounts for decades.

29. CityMail’s personnel belonged to the same union as the Post Office workers and were paid equivalent wages and benefits, but their work rules were less restrictive. See Cohen et al. (2007).

30. As noted above, Postverket had been arguing for some time that the statutory monopoly was useless. However, it argued that, as a state agency, it was legally bound to report the violation and that clarification of the monopoly was in any case needed.

31. The coalition comprised four parties – the Moderates, the Liberals, the Centre Party and the Christian Democrats (in order of size) – and relied on a fifth, New Democracy, for a parliamentary majority.

32. Chiefly the archipelagos, which are iced in during the winter, and the isolated mountain communities along the Norwegian border.

33. For a detailed look at the “graveyard spiral” model, see Crew and Kleindorfer (2005) and Jonsson and Selander (2006).

34. Via either the abolition of the monopoly or a ruling that confirmed that CityMail’s activities fell outside it.

35. This was particularly relevant in the case of bulk mail that was itself of legal significance, such as bank statements and bills.
37. “The government or an authority appointed by the government.”
38. The Basic Counter Service Act imposed this requirement on the universal service provider.
39. Ambrosini et al. (2006); Andersson (2001:126). In theory, these services could be provided by another company: they are not part of the USO and the PTS has contracted them to Posten under normal public procurement procedures.
40. The unions were the only party to the consultation to commission and submit an academic report in which the reform was analysed in terms of the theories of natural monopoly and arguments for and against the greater efficiency of a liberalised postal sector were analysed; Andersson (2001:128).
41. Perhaps the most significant change in this respect was the requirement that Posten employees work the hours for which they were paid; previously, if a carrier finished his route ahead of time, he could simply take the rest of the day off. After liberalisation, this practice was ended, a move that helped Posten rationalise its routes and its use of staff over the years that followed.
42. The Swedish Philatelists’ Federation (SFF) opposed liberalising the right to issue stamps, to which Postverket did not object. Since stamped mail is an expensive and probably unprofitable undertaking for a new entrant, the most likely motive for entry would be philatelic, and the SFF feared that interest in, and the value of, Swedish stamps would decline in the event of unrestricted entry.
43. Most of these issued their own stamps at one time or another, to the consternation of the SFF.
44. There are an additional 260-70 small unlicensed firms on the courier, express and parcel market. See Brandt (2007).
45. See the data in PTS (2008).
46. In many urban areas, factors such as complex handling of codes and keys to buildings and difficult parking and traffic conditions raised the cost of delivery. In addition, postal staff turnover is far higher in urban areas, as is resident mobility, which results in more frequent errors in addresses. See Jonsson and Selander (2006).
47. This claim is entirely plausible: the cost of the USO has recently been estimated at just SEK 100 million per year (Cohen et al., 2007).
48. Posten did receive state aid until 2007 for counter services. This was instituted when Posten withdrew from its giro business and had more to do with financial reform than postal reform.
49. Oddly enough, prior to the second bankruptcy, in 1997, Posten itself, under pressure from the PTS, purchased and operated CityMail for a time.
51. Some customers did not wish to accept the service lag implied by delivery every third day, and many bulk mailings were not sufficiently dense in their coverage to make presorting to carrier-route level attractive.
52. Both European and Swedish law generally forbid the use of such instruments by dominant players on a market.

53. This figure includes both applications for negative clearance by Posten itself as well as complaints brought by private persons and companies.

54. The Market Court (Marknadsdomstolen) is a specialised court that tries cases related to the Competition Act, the Marketing Act and other consumer and marketing legislation.

55. In essence, the court found that Posten wished to cut prices in areas where it faced competition from CityMail, while maintaining higher prices for the same services elsewhere.

56. Not all areas to receive discounts were served by CityMail. The discount was fixed at SEK 0.40.

57. The PTS, however, has ruled against some of Posten’s discount regimes; see PTS (2000:8-9); and Cohen et al. (2007).


59. Cumulative inflation over the period amounted to just 14%. The imposition of 25% VAT on postal services accounted for only a part of the more rapid increase in postal rates.

60. PTS (2000:4).

61. Jonsson and Selander (2006) report that the price of a 20g letter to another Nordic country rose 129% in the first decade after liberalisation, while rates for some domestic parcels rose by up to 189%.

62. All three packages of amendments included numerous technical and definitional changes that will not be discussed here; for details, see Andersson (2001), chapter 6.

63. The licensing requirement applied to companies delivering addressed letters and parcels on a regular basis according to a fixed distribution schedule.

64. These amendments also brought Swedish legislation into line with the EU postal services directive.

65. Posten claims that labour accounts for about 60% of total costs and transport for 20%. The complaint is not without foundation: Cohen et al. (2007) independently estimate that the aggregate cost of all inputs used to deflate letter prices rose by 59% during 1990-2004, while the CPI rose only 34%.


67. In part, such disputes arose because the PTS was reluctant to issue any precise definition of bulk mail. Precision on this point was difficult, because a number of factors besides the number of items might come into play (format, weight, etc.) and because technical change could render a ruling on definitional issues obsolete.

68. See PTS (2000:8-9).

69. See Andersson (2001), chapter 6 for details.

70. In essence, this enshrined in statute the criteria set out in an October 1997 ruling by the Competition Authority.
71. Andersson (2006b). The legislation did not address the future of SvAAB or forwarding arrangements: these were resolved by negotiations among postal operators, with the PTS mediating.


73. Andersson (2006a); Brandt (2007).

74. See Postcomm (2005) for data on the evolution of Posten’s network. By the early 2000s, Posten had more outlets than it had prior to liberalisation, but only 15% of these were staffed by Posten employees. Cohen et al. (2007).

75. See KKV (2004) for details.

76. Cohen et al. (2007).

77. Cohen et al. (2007). This is a somewhat simplistic measure of productivity, but more conventional measures would be difficult to estimate given the reorganisation of Posten, the liberalisation of the market and the fact that pricelists give a misleading picture of changes in the market, because most mail since the reform has been priced through agreements with individual customers that often involve significant discounts.

78. According to PTS (2007), Posten AB in 2006 employed roughly 33% fewer employees than in 1993. Geddes (2003, 2005) reports that this was done mainly through attrition.

79. In 2006, Posten bought the largest upstream firm in Sweden, Strålfors, which prints and envelopes presorted bulk mail. This transaction has since led to friction with the Competition Authority, which has accused the monopolist of failing to observe the conditions under which the sale was allowed to go ahead.

80. Between 1991 and 2006, prices for most forms of bulk mail rose modestly in constant SEK, and prices for many actually fell in real terms, particularly presorted bulk mail addressed to large towns. The reform thus accentuated, as well as reflected, the pre-existing trend towards use of the post for mass communications rather than correspondence. See Andersson (2006b:40); and Cohen et al. (2007).
Bibliography


KKV (2004), Monopolmarknader i förändring, Konkurrenserverket rapportserie 3, Stockholm, July.


The Political Economy of Reform
LESSONS FROM PENSIONS, PRODUCT MARKETS AND LABOUR MARKETS IN TEN OECD COUNTRIES

Why are some policy reforms implemented while others languish? This report aims to answer this important question by looking backwards – at 20 structural reform efforts in 10 OECD countries, during the past two decades. The case studies cover a wide variety of reform attempts in three key areas: pensions, labour- and product-market regulation. Key factors in the political, economic and reform-specific arenas are identified as helping or hindering reform, and these findings are cross-checked using a relatively simple set of Spearman rank correlations. The report’s two-pronged analytical approach – quantitative and qualitative – results in unique insights for policy makers designing, adopting and implementing structural policy reforms.

Further reading
Economic Policy Reforms 2009: Going for Growth

The full text of this book is available online via this link:
www.sourceoecd.org/generaleconomics/9789264073067
Those with access to all OECD books online should use this link:
www.sourceoecd.org/9789264073067
SourceOECD is the OECD online library of books, periodicals and statistical databases. For more information about this award-winning service and free trials, ask your librarian, or write to us at SourceOECD@oecd.org.