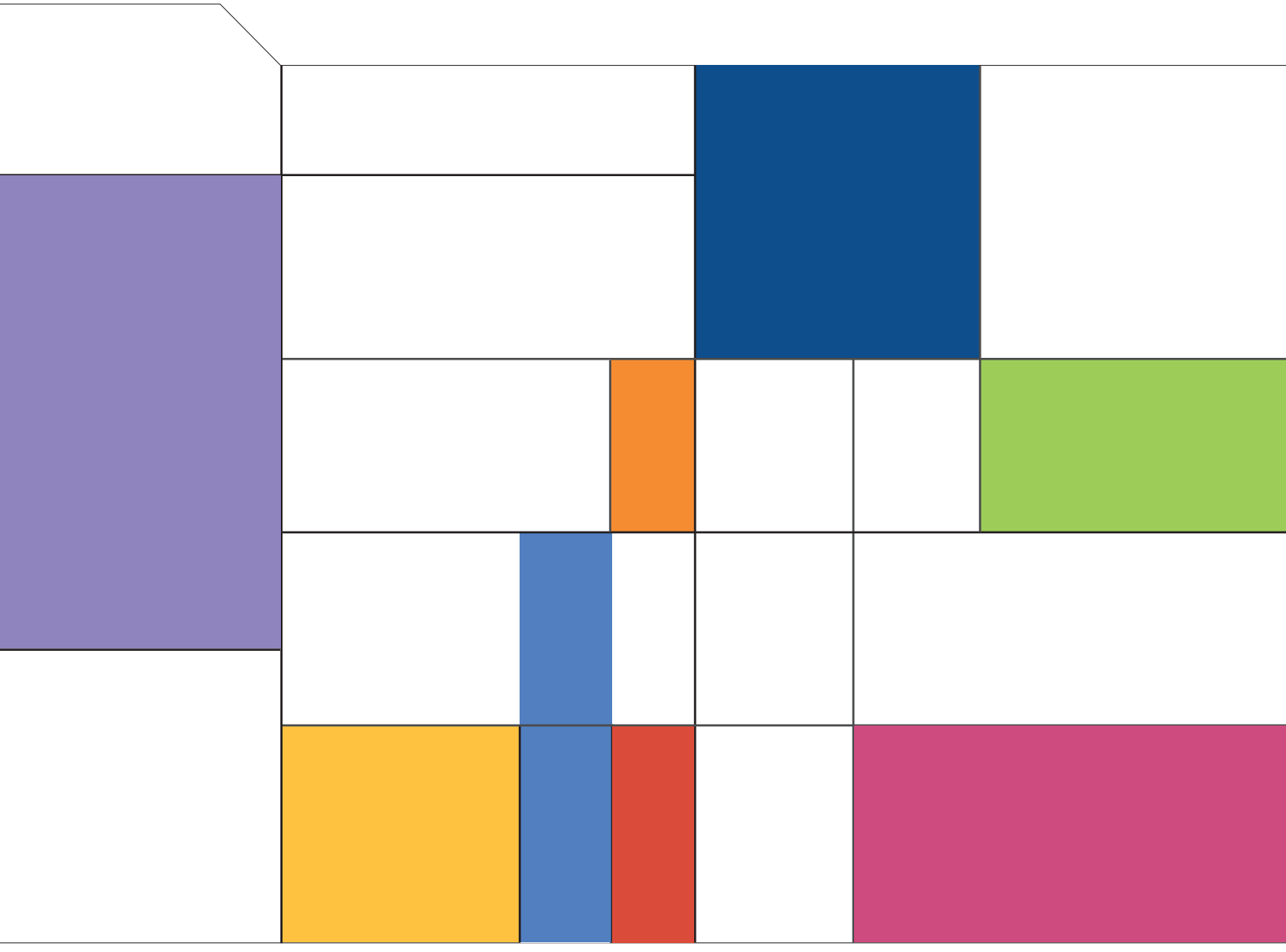


OECD Journal

Economic Studies



OECD

Volume 2013

OECD Journal

Economic Studies

Volume 2013

The OECD Journal collection comprises four periodicals: *OECD Journal on Budgeting*, *OECD Journal: Journal of Business Cycle Measurement and Analysis*, *OECD Journal: Economic Studies* and *OECD Journal: Financial Market Trends*. Each periodical within *OECD Journal* retains its own ISSN, volume and issue numbering. The volume numbering of each journal corresponds to the calendar year. There will be one issue of *OECD Journal: Economic Studies* this year.

OECD
BETTER POLICIES FOR BETTER LIVES

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.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

OECD Journal: Economic Studies, volume 2013
ISSN 1995-2848 (print)
ISSN 1995-2856 (online)
http://dx.doi.org/10.1787/eco_studies-v2013-1-en

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Note by Turkey: The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue".

Note by all the European Union member states of the OECD and the European Union: The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

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

© OECD 2014

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 the 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

OECD Journal: Economic Studies *is an official publication of the Organisation for Economic Co-operation and Development (OECD). It is addressed to a broad audience of policy-makers, academics and interested readers in the wider public.*

Economic Studies *publishes economic articles in various areas of policy analysis, generally based on empirical work with a cross-country dimension. However, beyond international applied research, theoretical papers, meta-analyses or country-specific pieces of empirical work are also welcome, inasmuch as they contribute to renewing the approach to an economic issue or they participate in important debates among economists or practitioners. As well, the submission of short contributions to such debates or of comments on papers previously published in* Economic Studies *is encouraged.*

While the journal was initially intended as an outlet for research done within the OECD, especially in the Economics Department, for several years now, Economic Studies *has been welcoming external submissions. Its scope has also been broadened to policy areas well beyond macro and microeconomic analysis. This volume, for example, includes articles on topics as different as the effects of trade on technologies, the international linkages between carbon markets, or the consequences of grade repetition at school.*

Simultaneously, the publication process has been strengthened to allow for a double peer review: all papers submitted, if accepted by the Board of editors, are reviewed by two referees, one internal to the OECD and the other external. This guarantees that the papers that are published are consistent with both the standards that the OECD applies to its own work for the member countries and the requirements of academic excellence. Of course, the views expressed in the papers published remain the responsibility of the sole authors and should not be attributed to the OECD.

Articles are published online as they become available and assembled in an annual volume released at the end of each year.

Editorial Board

Jean-Luc Schneider (Chairman), Deputy Director, OECD Economics Department, Paris, France

Sveinbjörn Blöndal, Head of Division, OECD Economics Department, Paris, France

Jorge Braga de Macedo, NOVA School of Business and Economics, Lisbon, NBER, Cambridge Mass. and CEPR, London; former president of the OECD Development Centre, Paris

Fabrizio Coricelli, Professor of Economics, Paris School of Economics, Université Paris 1 and Research Fellow, Centre for Economic Policy Research (CEPR), London, United Kingdom

Chiara Criscuolo, OECD Directorate for Science, Technology and Industry, Paris, France

Anna D'Addio, OECD Directorate for Employment, Labour and Social Affairs, Paris, France

James K. Hammitt, Harvard University (Center for Risk Analysis), United States and Toulouse School of Economics (LERNA-INRA), France

Bo Hansson, OECD Directorate for Education, Paris, France

Bo E. Honoré, Professor of Economics and Director of the Gregory C. Chow Econometric Research Program, Princeton University, US; Member of the Board of Trustees of the Danish National Research Foundation

Sébastien Jean, Centre d'études prospectives et d'informations internationales (*CEPII*), Paris, France. Policy associate and member of the Strategic Advisory Board, Leverhulme Centre for Research on Globalisation and Economic Policy (*GEP*), University of Nottingham, United Kingdom.

Nick Johnstone, Head of Division, OECD Directorate for Science, Technology and Industry, Paris, France

Francis Kramarz, Professeur École Polytechnique, Directeur, Centre de Recherche en Économie et Statistique (CREST) (Research Center for Economics and Statistics), Institut National de la Statistique et des Études Économiques (INSEE) (National Institute of Statistics and Economic Studies), Paris, France

Joaquim Oliveira Martins, Head of Division, OECD Public Governance and Territorial Development Directorate and Associate Professor, University Paris-Dauphine, Paris, France

David Turner, Head of Division, OECD Economics Department, Paris, France

Gert Wehinger, OECD Directorate for Financial and Enterprise Affairs, Paris France

Andreas Wörgötter, Head of Division, OECD Economics Department, Paris, France and Associate Professor (non-salaried position), Institute of Mathematical Methods in Economics, Vienna University of Technology

Lyn Urmston, OECD Economics Department, Paris, France

Table of contents

Reconciling fiscal consolidation with growth and equity
By Boris Cournède, Antoine Goujard *and* Álvaro Pina 7

Choosing the pace of fiscal consolidation
By Łukasz Rawdanowicz 91

Policy incentives for private innovation and maximising the returns
By Ben Westmore 121

Technological effects of intra-OECD trade in manufacturing: A panel data analysis over the period 1988-2008
By Javad Abedini 165

Demographic or labour market trends: What determines the distribution of household earnings in OECD countries?
By Wen-Hao Chen, Michael Förster *and* Ana Llena-Nozal. 179

Towards global carbon pricing: Direct and indirect linking of carbon markets
by Rob B. Dellink, Stéphanie Jamet, Jean Chateau *and* Romain Duval 209

Using a quasi-natural experiment to identify the effects of birth-related leave policies on subjective well-being in Europe
By Anna Cristina d'Addio, Simon Chapple, Andreas Hoherz *and* Bert Van Landeghem . 235

Grade repetition: A comparative study of academic and non-academic consequences
By Miyako Ikeda *and* Emma García. 269

OECD Journal: Economic Studies
Volume 2013
© OECD 2014

Reconciling fiscal consolidation with growth and equity

by
Boris Cournède, Antoine Goujard *and* Álvaro Pina*

Despite sustained efforts made in recent years to rein in budget deficits, a majority of OECD countries still face substantial public finance consolidation needs. While essential to avoid the disruption and large costs ultimately associated with unsustainable public finances, fiscal consolidation complicates the task of achieving other policy goals. In most cases, it weighs on demand in the short term. And, if too little attention is paid to the mix of instruments used to achieve consolidation, it can undermine long-term growth, exacerbate income inequality and slow the process of global rebalancing. It is therefore important for governments to adopt consolidation strategies that minimise these adverse side-effects. The analysis proposes consolidation strategies that take into account other policy goals as well as country-specific circumstances and preferences. To do so, increases in particular taxes and cuts in specific spending areas are assessed for their effects on short- and long-term growth, income distribution and external accounts. The results of detailed illustrative simulations indicate that a significant number of OECD countries may have to raise harmful taxes or cut valuable spending areas to deliver sufficient consolidation, underscoring the need for structural reforms to counteract these side-effects. The results are robust to an extensive range of sensitivity checks.

JEL classification: H62, H63, H68

Keywords: Fiscal consolidation, growth, equity, global imbalances, income distribution, structural reforms

* The authors are members of the OECD Economics Department. Álvaro Pina is also affiliated with ISEG (Lisboa School of Economics and Management, Universidade de Lisboa) and UECE (Research Unit on Complexity and Economics, Lisboa). This paper originated as a document prepared for a meeting of the Working Party No. 1 of the Economic Policy Committee (March 2013). Versions of the paper have also been presented at a Department of Economics seminar at ISEG, at the 9th BMRC-QASS Conference on Macro and Financial Economics (both in May 2013) and at a special Lisbon Council event in July 2013. The authors would like to thank Alain de Serres, who co-ordinated the project, the meeting participants as well as Pier-Carlo Padoan, Jørgen Elmeskov, Jean-Luc Schneider, Alain de Serres, Sebastian Barnes, Ray Barrell, Henrik Braconier, Orsetta Causa, Roger Farmer, Alberto Gonzalez Pandiella, Peter Hoeller, Isabell Koske, Stephen Matthews, Oliver Roehn and Eckhard Wurzel for helpful comments on earlier drafts, Caroline Abettan and Lyn Urmston for excellent editorial support, Agnès Cavaciuti for expert statistical assistance and they are indebted to Benjamin Carton and Andreas Wörgötter for valuable comments and suggestions. The views expressed in this paper are the authors' and are not necessarily shared by the OECD or its member countries. Corresponding author: boris.cournede@oecd.org.
The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

1. Introduction and main messages

Despite considerable progress in recent years, many OECD countries were still facing sizeable fiscal consolidation at the end of 2012 to bring back or keep public debt within manageable levels and avoid the cumulative costs and risks associated with the build-up of unsustainable public-finance positions. The origin of fiscal consolidation needs is complex and manifold, including *inter alia* initial conditions, the fiscal damage from financial crises, spending pressures from demographic developments and liquidation losses from public assets. Building on previous OECD and other work, the article presents a structured approach to choosing the instruments of fiscal consolidation strategies in ways that minimise adverse side-effects on growth and equity in the short and the long term as well as on external imbalances. As such, the focus of the paper is on the composition of consolidation. For in-depth recent studies about the size of consolidation needs and the optimal timing of consolidation, see for instance Sutherland et al. (2012) and Rawdanowicz (2012), respectively. The paper provides illustrative quantitative applications of the approach.

The main messages can be summarised as follows:

- In most OECD countries, compared with what had been achieved by end 2012, additional consolidation is needed in the short to medium term to put government debt on a trajectory toward more prudent levels (defined for simplicity as gross debt at 60% of GDP) and even more to keep debt stable in the very long run, i.e. in 2060.
- Consolidation instruments (increases in particular taxes and cuts in specific spending areas) can be ranked according to their effects on short- and long-term growth, income distribution and current accounts, with the rankings taking into account country circumstances.
- Based on these rankings, illustrative consolidation packages to optimise the side-effects of consolidation on other policy objectives can be drawn up for each country.
- The packages are based on using instruments sequentially, and within reasonable limits, starting from the most desirable and moving down the ranking until consolidation needs are satisfied.
- Based on this approach, half of OECD countries appear to be in a position to fulfil their short- to medium-term consolidation needs entirely through instruments that are well ranked (that is to say ranked in the top half). This suggests that, in these countries, well-designed consolidation packages can avoid severe adverse side effects on growth, equity and current-account imbalances.
- In the simulations, eight countries use some poorly-ranked instruments but achieve more than half their short- to medium-term consolidation through well-ranked instruments.
- Finally, in the illustrative simulations, three countries (Japan, United Kingdom and the United States) implement more than 50% of their short-to-medium-term consolidation

packages through instruments that have low rankings meaning that they are likely to involve substantial adverse side-effects.

- Despite the generally stronger consolidation requirements in the long term, 20 countries would manage to keep debt durably stable at 60% of GDP by relying only on well-ranked instruments.
- In the simulated very long-term packages, six countries use some poorly-ranked instruments but can nonetheless achieve more than 50% of their adjustment with well-ranked instruments.
- Finally, a few countries would have to implement most of their simulated very long-term consolidation packages relying on poorly-ranked instruments (that have more adverse effects on long-term growth and equity objectives).
- On average across countries, spending reductions would account for 41% of short- to medium-term and 65% of long-term consolidation packages, the rest being achieved through tax hikes. The difference mostly reflects the greater concern for demand effects in the short term.
- The proposed illustrative consolidation packages lead to some, but not much, convergence in spending and revenue structure across countries over time. This result can be interpreted as meaning that the proposed approach is largely respectful of the diversity of national preferences over spending and revenue structure.
- Extensive sensitivity analysis indicates that all the above results are largely robust to uncertainty about the assessments of the effects of instruments except the spending-tax split in simulated short-term consolidation plans. The simulation results are also very robust to changes to the method used to adapt instrument rankings to country circumstances. The assumptions made to define the maximum amount by which each consolidation instrument can be used have a direct influence on the degree to which countries with large adjustment needs have to rely on the most harmful categories of tax hikes or spending cuts.
- Trade-offs between consolidation and other policy objectives can be eased by exploiting the scope for efficiency gains through structural reforms.

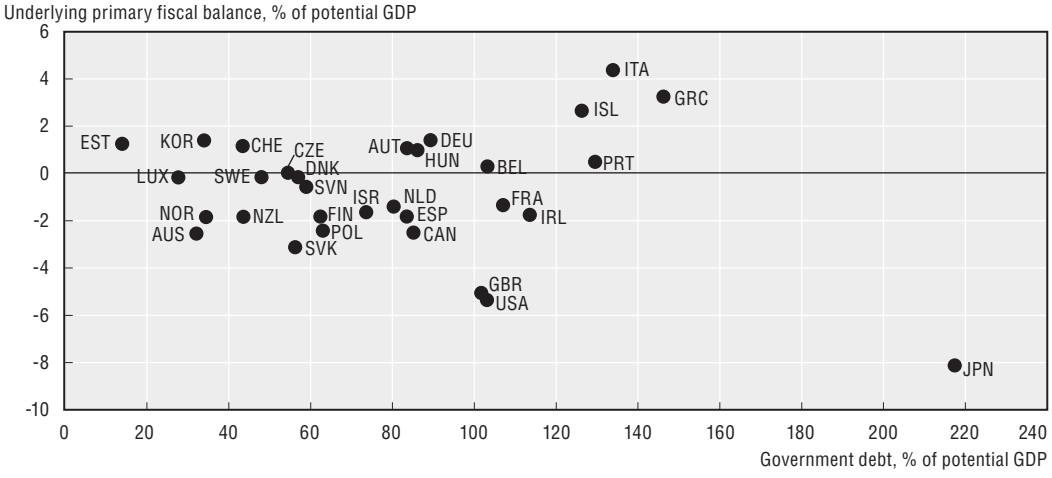
In a preliminary step providing inputs for the subsequent analysis of ways to minimise the side-effects of consolidation, the study first estimates fiscal consolidation needs in the short to medium term and the long term (Section 2). It then moves to its core subject and discusses the definition of growth, equity and current-account objectives before presenting the list of potential consolidation instruments, evaluating their effects along these three dimensions and proposing a generic illustrative hierarchy of instruments (Section 3). On that basis, Section 4 proposes a method for developing country-specific hierarchies of instruments taking into account country specificities. A file available online allows readers to build their own rankings of consolidation instruments by keying in their preferred weights on growth, equity and current-account objectives.[1] The study proceeds with an illustrative evaluation of how far down each country has to go on its list from more to less welcome instruments to meet its consolidation objectives without departing too much from its revealed preferences about government spending and revenue items and checks the robustness of the findings (Section 5). The results underscore the need for structural changes to be part of fiscal adjustment (Section 6). Section 7 then makes a few concluding remarks and offers suggestions for future research.

Appendix A reports country details of the main simulations as well as summary statistics from alternative simulations based on random draws. Appendix B provides supporting material on: the methodology used to estimate consolidation needs (summarised in Boxes 1 and 2), the dataset, the assessment of instrument impact, the analysis of debt behaviour during consolidation episodes, a variant without clustering analysis and another variant with increased room for manœuvre. Further supporting empirical evidence is provided by Barbiero and Cournède (2013) and Goujard (2013).

2. Estimated consolidation needs

The legacy of the financial crisis and earlier fiscal imbalances has burdened many OECD governments with high debt levels often accompanied by still significant structural deficits (Figure 1) which call for large consolidation efforts to reduce debt to more prudent levels. As a preliminary intermediate step to permit a quantitative analysis of the composition of consolidation strategies, this section presents estimates of consolidation needs for both the short to medium term and the long term. The present calculations are based on a gradual consolidation effort, embodied in smooth time paths for the structural primary balance (see Appendix A, Section 1 for more details). This approach ensures that the debt ratio is on a stable trajectory at the end of the consolidation horizon (2060). Second, in order to ensure that by 2060 the debt ratio not only stabilises but does so at the desired target level (60% as in Johansson et al., 2013), the present work differentiates short- from long-term consolidation needs, as explained in greater detail below. As developed in Box 1, this approach differs in purpose and methodology from the consolidation requirements reported in the *OECD Economic Outlook*, May 2013 (OECD, 2013a).

Figure 1. **Debt and underlying primary balances in 2012**



The short- to medium-term consolidation need is defined as the difference between a baseline and the peak of a trajectory for the underlying primary balance that brings gross general government debt to 60% of GDP by 2060. A uniform target of 60% of gross debt to GDP has been chosen for the sake of simplicity and because it represents an important reference in EU countries. However, in practice, national circumstances warrant different objectives. In particular, the presence or not of large amounts of financial assets on the government balance sheet is an important point to consider when setting targets for gross

Box 1. **Short- vs. long-term consolidation needs and average requirements**

The estimated consolidation needs presented here differ from the average consolidation requirements reported in OECD (2013a) as they serve different purposes and therefore use different assumptions. The present set of consolidation needs forms a basis for the subsequent quantitative analysis of detailed consolidation packages that minimise side effects. The focus is firstly on how far these packages need to go in the short to medium term to bring debt under control and secondly on what has to be done to keep debt stable in the very long term, that is to say in 2060 and beyond. This differs from the objective of the requirements reported in OECD (2013a) which was to show how much effort beyond that already built into the near-term projection is needed on average from 2015 to 2030. From these different purposes and perspectives result different methodological choices with the main differences summarised as follows:

- The reference point for comparisons is 2012 in the current study, so that needed changes in individual areas of tax and spending can be compared with the latest historical point (or estimate). The reference point in OECD (2013a) is fiscal projections to 2014 to provide an idea of how much remains to be done in aggregate after the expected consolidation to 2014.
- The present estimates refer to the peak effort needed in the short- to medium-term and in 2060 whereas the requirements reported in OECD (2013a) relate to the average effort over 2015-30. The former is needed for the present exercise as the point to assess how far, at the peak, instruments have to be used, and whether these instruments have to be maintained or can be partly reversed afterwards. To assess the size of aggregate consolidation efforts in an extended medium-term perspective as is the case in OECD (2013a), however, the average offers a more robust measure given that many different paths with many different peaks can be imagined for moving to debt stabilisation.
- In order to allow more realistic estimates of consolidation needs in the very long run (2060), the present estimated needs are calculated over a baseline where government expenditure on health and long-term care increases gradually over time as in de la Maisonneuve and Oliveira Martins (2013). The baseline for comparisons in OECD (2013a) does not incorporate such cost pressures which have a lesser impact when looking at average effort over 2015-30.
- For the sake of comparability of consolidation packages and in line with the long-term focus of the study, the present set of estimates assumes that all countries reach 60% gross debt-GDP ratios by 2060. In OECD (2013a), in line with the extended medium-term focus, the time horizon is 2030 but, to avoid too abrupt changes, some countries are allowed to reach their 60% target after 2030.

Despite the differences of purposes and method, the cross-country correlation between the present set of short- to medium-term consolidation needs and the requirements presented in OECD (2013a) is very strong with a coefficient of 96%.

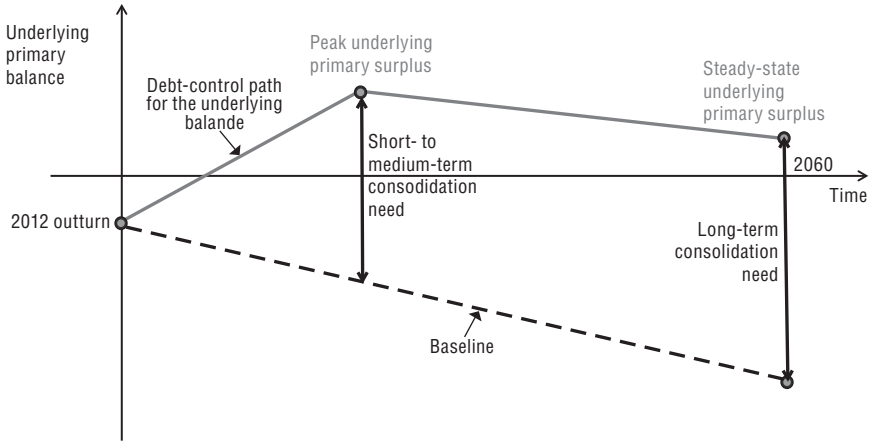
Source: OECD (2013a), *OECD Economic Outlook*, May 2013.

debt. A variant including drawing down financial assets to smooth the adjustment is included and discussed below. Another factor bearing on the level of debt that the government can carry without exacerbating risks of instability is private-sector indebtedness and associated risks of deleveraging. Recent OECD work suggests that the prospect of deleveraging is real in a number of OECD countries (Bouis et al., 2013).

Evidently, different consolidation paths can be taken to attain the 60% target, each leading to a different profile for the underlying primary balance (see Box 4.5 in OECD, 2013a). For the purpose of this exercise, and although some countries plan to adjust faster (see OECD, 2013a), the underlying primary balance is assumed to improve from its 2012 level at a rate of 1% of potential GDP each year for as long as necessary to put debt on a trajectory toward the target. This is a simplifying assumption to derive estimates on a uniform basis across countries because this study does not focus on the optimal timing of consolidation but on its composition. For a study about the optimal timing of consolidation from a welfare perspective, see for instance Rawdanowicz (2012).

After that initial period of consolidation, which varies considerably in length across countries, the underlying primary balance is assumed to converge very gradually to the 2060 level which stabilises debt at 60% of GDP (see Figure 1 and Section 1 in Appendix B). Initial improvement in the underlying primary balance at the fast annual pace of 1% (1½ per cent in Japan) helps to ensure that debt is put on a downward path in a not too distant future.[2] Interest rates and GDP growth, important drivers of debt dynamics, are assumed to follow the long-term baseline projections published OECD (2013a) of which the main assumptions are summarised in Box 2. The calculations are based on the effective interest rate paid by governments on the stock of debt, as projected in the OECD (2013a) long-term projections, so that the maturity structure of the debt stock is taken into account. Last historic point (2012) data for general government debt and underlying primary balances are also taken from the OECD (2013a). Figure 3 shows two concrete examples of baseline and debt-reducing trajectories.

Figure 2. **Defining short- to medium-term and long-term consolidation needs**



The long-term consolidation need compares the "debt-control" underlying primary balance with the baseline at the end of the projection period. The baseline corresponds to a policy scenario where sufficient reforms are introduced for public pension spending to remain constant relative to potential GDP and for government expenditure on health and long-term care to grow at a contained pace. Other tax and expenditure components are assumed to be unchanged from their 2012 levels relative to GDP except for cyclical effects associated with the projected closure of output gaps.

Box 2. **Main assumptions behind the OECD long-term growth projections**

Consolidation needs are calculated using the OECD long-term baseline growth projections to 2060 published in the *OECD Economic Outlook,* May 2013 (OECD, 2013a). The main features of this set of projections, which use the OECD long-term modelling framework presented in Johansson et al. (2013) are as follows:

- The backbone of these scenarios is a set of long-run projections for potential output based on physical and human capital and labour as production factors plus labour-augmenting technological progress. The latter is assumed to converge at a speed that depends on the starting point, with countries farther away from the technology frontier converging faster.
- The gap between actual and potential output is gradually eliminated from 2013, for most countries within four to five years, depending on the initial size of the output gap.
- Inflation is assumed to converge gradually to its target rate (2% in most countries).
- Policy-controlled short-term interest rates increase gradually as output gaps close before converging to neutral short-term rates following potential GDP growth rates. Long-term interest rates are a forward convolution of short-term rates plus a fixed-term premium, a fiscal-risk premium reflecting debt levels and a global-balancing premium. Effective interest rates are calculated as an average between short-term and long-term interest rates reflecting the maturity structure of government debt.

Source: Johansson et al. (2013).

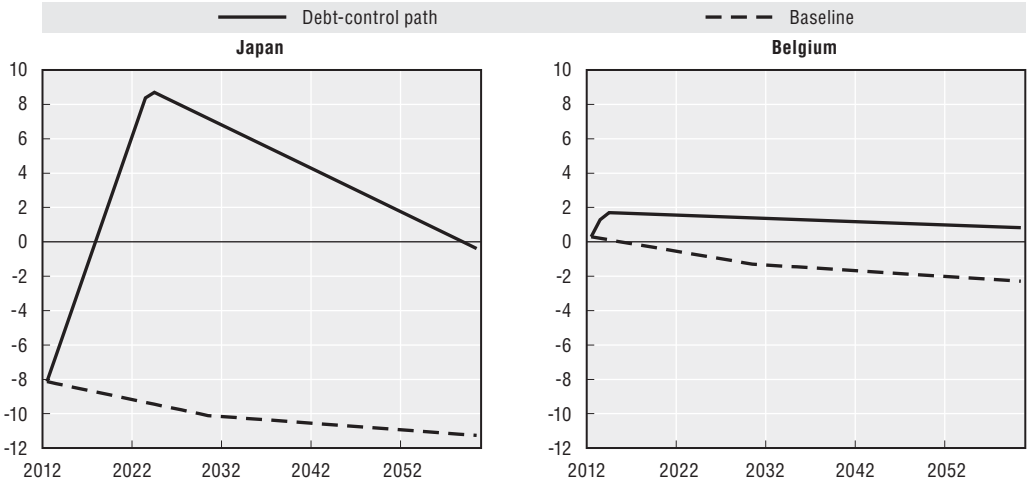
The baseline scenario therefore incorporates significant reform in the areas of pensions and health.

- Many countries expect large increases in government pension expenditure relative to GDP on current policy settings (Figure 4). The baseline scenario assumes that, in these countries, substantial reforms are implemented, including adjustments of the effective retirement age, so as to keep stable the ratio of public pension spending to GDP.
- Similarly, the continuation of past trends in public spending on health and long-term care would appear to result in large further increases, as apparent in the projected "cost-pressure" scenario presented in de la Maisonneuve and Oliveira Martins (2013) and plotted in Figure 5. Hence, the baseline incorporates (unspecified) measures to contain cost pressures in health and long-term care, which could include a more frequent re-evaluation of drug prices, centralised bargaining for drug purchases, more user choice of health providers and incentives to enhance prevention *inter alia* (Joumard et al., 2010). Nonetheless, even under the cost-containment assumption, health spending is projected to rise as a share of GDP, which explains the trend decline in the primary balance in the baseline paths shown in Figures 2 and 3.

Estimates based on the approach described above suggest that in Greece, Japan, Portugal, Spain, the United Kingdom and the United States, a short- to medium-term consolidation in excess of 5% of potential GDP is required to reduce debt to 60% of GDP by 2060 (Figure 6). This is the result of currently high debt (Greece, Ireland, and Portugal), a large initial underlying primary deficit (Spain), or their combination (Japan, United Kingdom and United States). In a first subgroup of these countries (Greece, Portugal, Spain), the estimated large consolidation needs are associated with substantial risk premia in long-term interest rates. In a second subset (Japan, the United Kingdom and

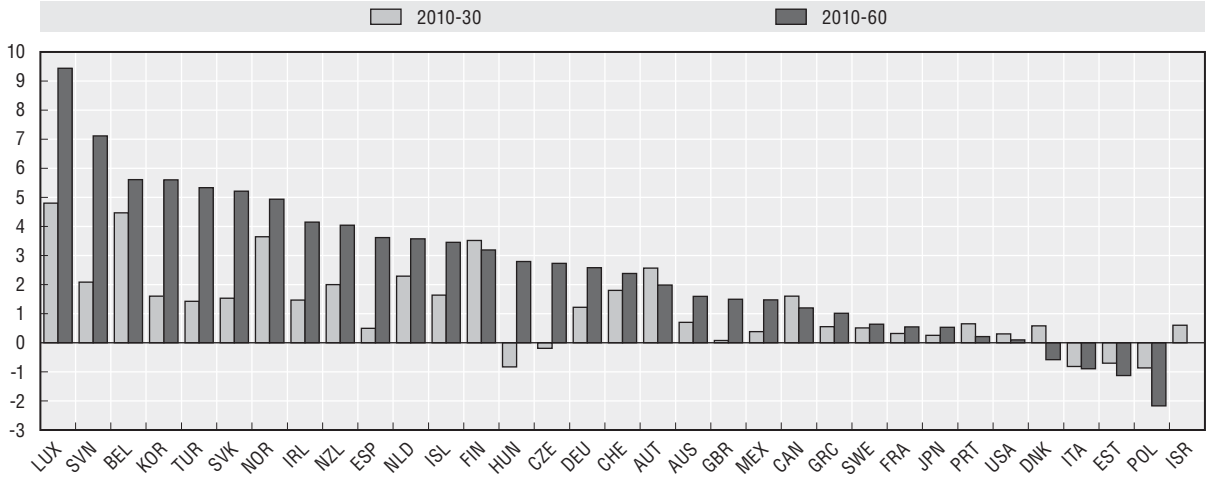
Figure 3. **Illustration of the budget consolidation profile compared with baseline in two countries**

Simulated underlying primary balance, per cent of potential GDP



Source: *OECD Economic Outlook*, May 2013 database and OECD calculations.

Figure 4. **Projected change in government pension expenditure on unchanged policies**

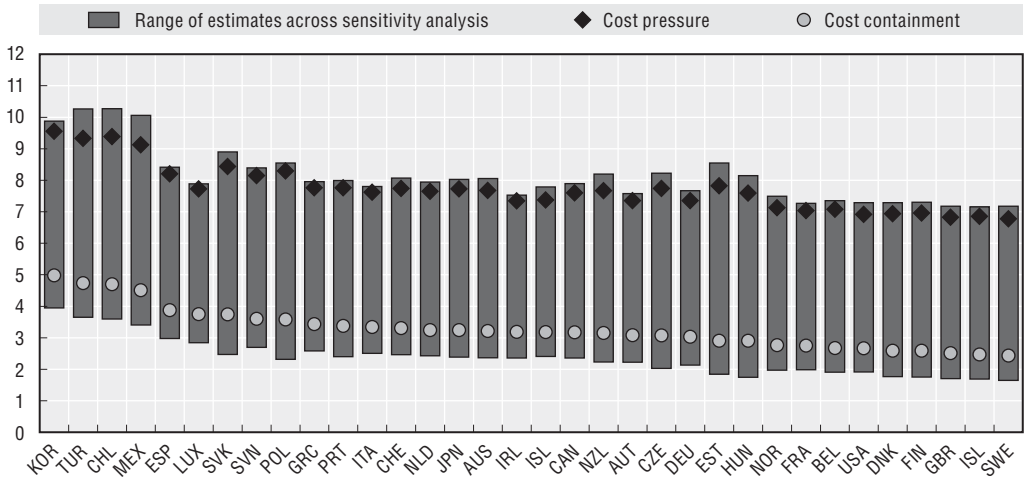


Source: EC (2012) for EU countries and Norway, OECD (2011a) for other countries except Japan where the estimate is taken from Merola and Sutherland (2012) and Israel where it has been estimated for 2030 based on projections to 2025 communicated by the Bank of Israel.

the United States), long-term interest rates are much lower, which can be related to the large-scale purchases of government bonds by central banks under their quantitative-easing policies. To bring debt to the same level, another group needs short- to medium-term consolidation by more than 3% of GDP – though less than 5% – because of high debt levels (France, Iceland) or a significant underlying primary deficit (Finland, Poland, Slovak Republic). Other countries, including in particular Italy and Germany, face little or no short- to medium-term structural consolidation needs, though high debt in the former makes this conclusion vulnerable to interest-rate changes. When needed, consolidation is in most cases relatively brief in the simulations: three out of four countries that require short- to medium-term consolidation complete it in four years or less. Many countries have made consolidation plans that go a long way toward meeting these consolidation needs (OECD, 2013a).

Figure 5. **Projected percentage point increase in total public health and long-term care spending, 2010-2060**

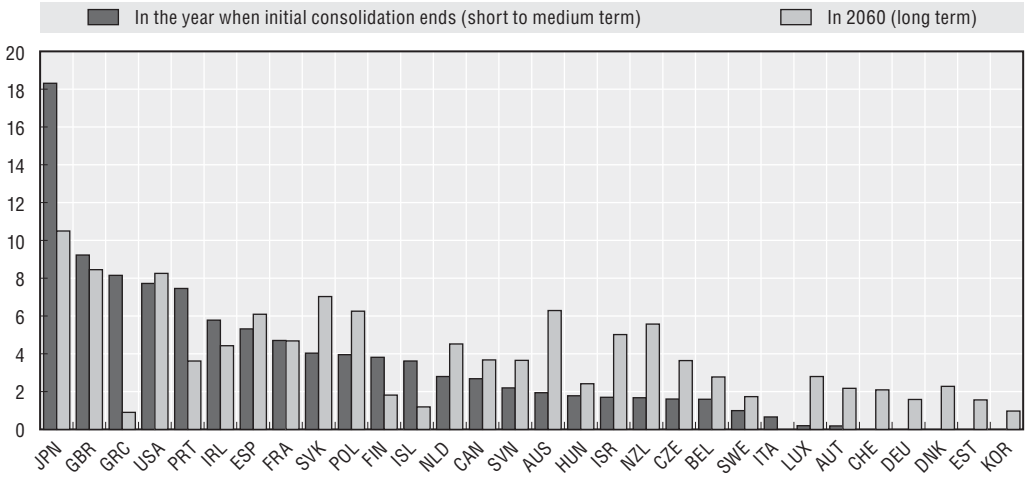
Range of estimates across sensitivity analyses



Notes: Countries are ranked by the increase of expenditures between 2010 and 2060 in the cost-containment scenario. The vertical bars correspond to the range of the alternative scenarios, including sensitivity analysis.
Source: de la Maisonneuve and Oliveira Martins (2013).

Figure 6. **Estimated consolidation needs**

Difference between debt-control and baseline underlying primary surplus, per cent of potential GDP



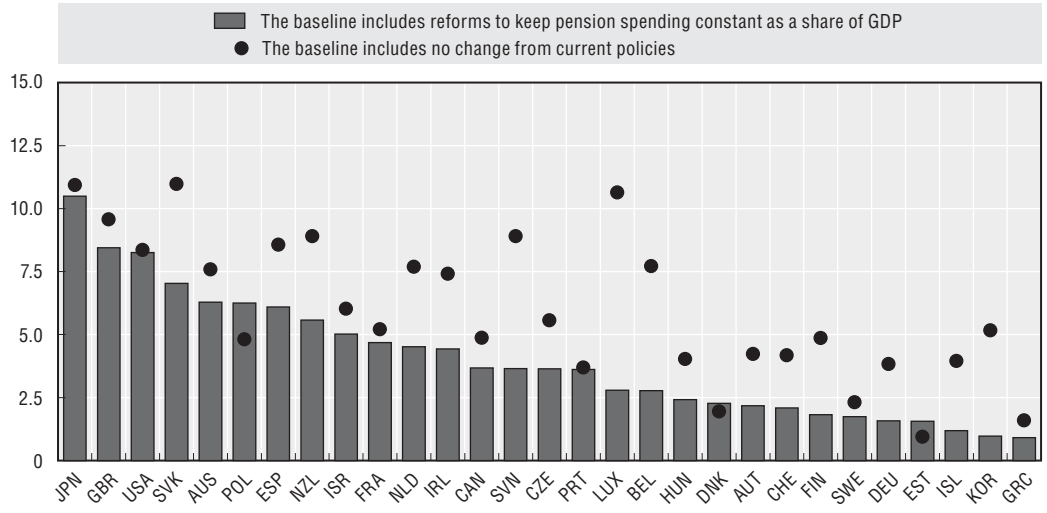
Source: OECD Economic Outlook, May 2013 long-term database and OECD calculations.

Consolidation needs are estimated to be greater in the long than the short term for the majority of countries. The difference is particularly large in countries where short-term needs are limited thanks to low initial debt levels. The high estimated level of long-term consolidation needs reflects large expected spending increases on health and long-term care. That said, since the cross-country variation in projected increases in government health spending is limited, it does not account for much of the differences in estimated long-term consolidation needs. The latter are primarily due to the starting point for the underlying primary surplus in 2012. Another significant source of differences is that the OECD long-term growth scenarios project interest rates rising well above nominal GDP growth rates by 2060, which leaves governments holding large amounts of financial assets

with substantial capital income to service their debt. This effect reduces the estimated long-term consolidation needs of Canada, Finland, Japan, Korea and Norway by 2½ per cent of GDP or more compared with a situation where these countries' governments had no financial assets. This set of long-term estimates is subject to particularly strong uncertainty.[3] If no pension reform was assumed in the baseline, that is to say if public pension expenditure was allowed to increase in line with unchanged policy projections, estimated long-term consolidation needs would be considerably larger in many countries (Figure 7). These very large differences underscore the critical need for pension reform in countries that have not yet adjusted their systems to ensure that government pension spending remains contained in the face of ageing. In addition to being key to fiscal sustainability, successful pension reform also brings important benefits in terms of greater labour supply (Duval, 2003) and intergenerational equity (Gonand, 2010).

Figure 7. **Long-term consolidation needs: Estimates with and without pension reform**

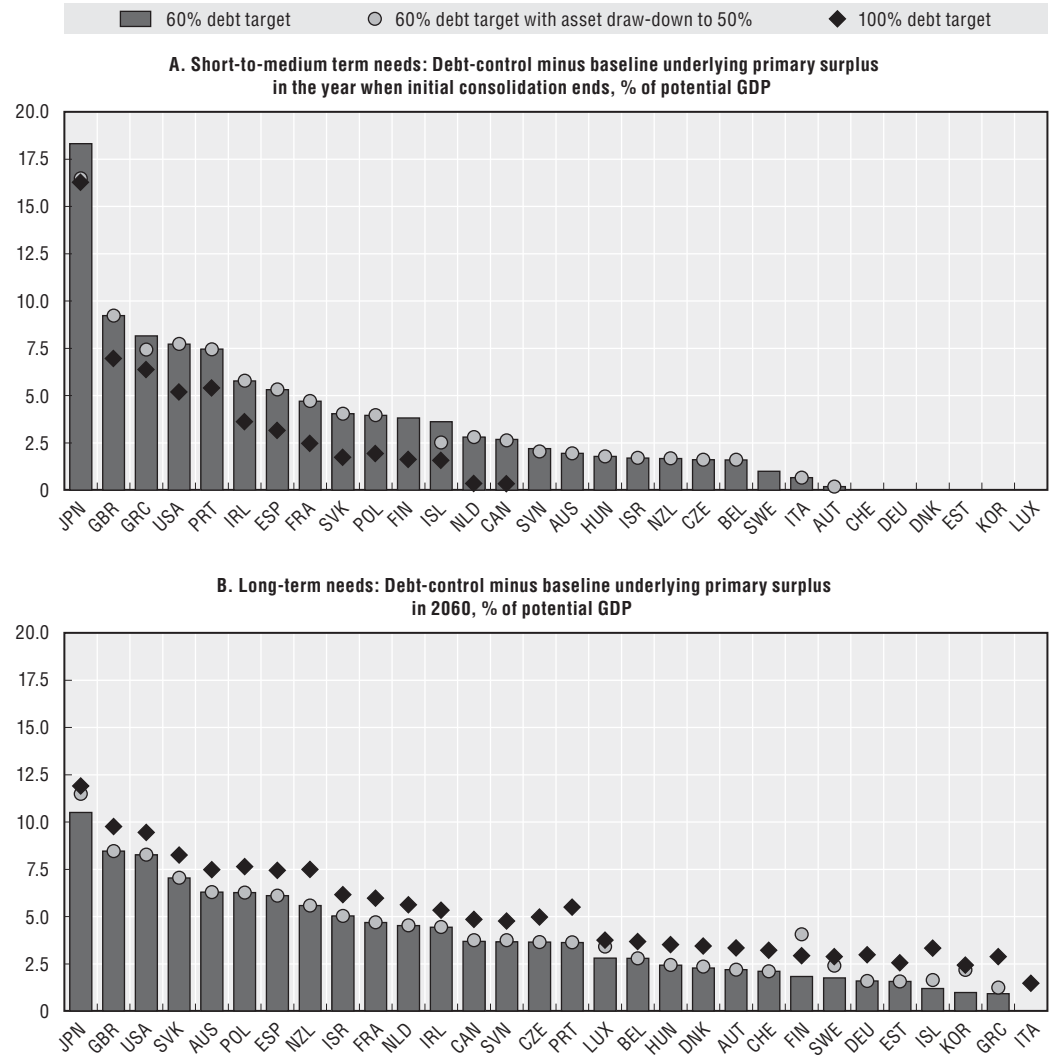
Difference between debt-control and baseline underlying primary surplus in 2060, per cent of potential GDP



Source: OECD Economic Outlook, May 2013 and OECD calculations.

The choice of a gross debt target can exaggerate consolidation needs for governments that have large sellable financial assets or substantial implicit assets for instance in the form of deferred tax on pension savings. A limited group of OECD governments (Estonia, Finland, Korea, Luxembourg, Norway, Sweden) report net positive financial asset positions. A significant number of OECD governments hold financial assets that are valued at more than half of their country's GDP (Canada, Denmark, Finland, Greece, Iceland, Japan, Korea, Luxembourg, Norway, Slovenia, Sweden).[4] Selling assets to meet consolidation needs, for instance by drawing them down to 50% of GDP in countries that currently hold more, eliminates estimated short- to medium-term consolidation needs almost entirely in Denmark and fully in Sweden (Figure 8, upper panel). This draw-down hypothesis also reduces consolidation needs significantly in Japan, where they diminish by nearly 2 percentage points but remain nonetheless elevated at 16½ per cent of GDP. In the long term, however, this draw-down hypothesis results in larger consolidation needs (Figure 8, lower panel) because asset depletion reduces the amount of capital income on government assets compared with the assumption of keeping them constant as a share of GDP. In practice, the ease with which

Figure 8. **Estimated consolidation needs**



Source: *OECD Economic Outlook*, May 2013 long-term database and OECD calculations.

financial assets can be liquidated varies across countries depending on their nature, on the extent to which they are earmarked to prefund budgetary commitments and on whether they are owned by the central or other levels of government (Rawdanowicz et al., 2011).

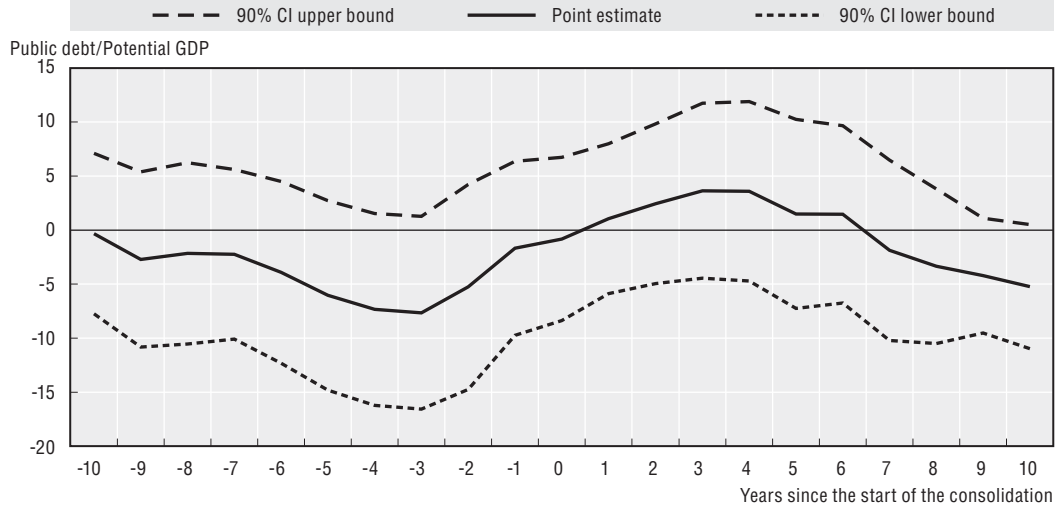
The chosen level of the debt target also has implications for consolidation needs. While there is no obvious optimal maximum ratio of public debt to GDP, the 60% value has been retained as the main reference point for the simulations because of its widespread use as a policy target within the OECD membership. Aiming at a higher 100% target would reduce estimated short- to medium-term consolidation needs by about 2% of GDP in most OECD countries (Figure 8A). Allowing greater indebtedness however comes at the cost of larger interest payments to keep the debt ratio stable, pushing up estimated long-term consolidation needs by about 1% of GDP in most OECD countries (Figure 8B). These considerations mean that the consolidation needs used in the rest of the paper, which correspond to the 60% of GDP gross debt target without asset draw-down, should be seen as illustrative as the choice of a debt target and the level of the target ought to be country specific.

Feedback from consolidation to activity could also influence consolidation needs in ways that are not fully reflected in the present set of estimates. In many countries, deeper consolidation would, through multiplier effects, reduce growth and at least temporarily create more adverse debt dynamics than assumed in the present simulations in the short to medium term. These effects could be magnified by the fact that most countries consolidate, implying that each country faces additional headwinds from external demand in its consolidation effort.[5] Afterwards, however, the return of output to potential would create more favourable growth and therefore debt dynamics than the one underpinning the present calculations. Simulations incorporating such effects by Rawdanowicz (2012) suggest that, even if multipliers are large, their effects on debt dynamics during and after the consolidation largely cancel out so that they have little effect on the estimated size of consolidation needs. One channel through which deep fiscal tightening can influence consolidation needs sizeably is if it generates hysteresis effects that depress potential output permanently, something which is assumed not to happen in the projections presented here. This consideration underscores the need to design consolidation strategies in ways that minimise the risk of generating hysteresis (see Section 4 below).

Feedback from consolidation to real interest rates could reduce consolidation needs. Consolidation strategies that are credibly seen as bringing back debt firmly within manageable levels are likely to lower risk premia and real interest rates (Turner and Spinelli, 2012). This favourable effect materialises in full only after the disinflationary consequences of any consolidation-induced contraction wear off. The historical experience is that, on average across large fiscal consolidation episodes, it takes three years after they start before the ratio of debt to potential GDP begins to fall and seven years before it becomes smaller than at the start (see Figure 9 and Blöchliger et al., 2012). Lower real interest rates improve debt dynamics directly, by fuelling demand, and also by

Figure 9. **Large fiscal consolidation and the government debt-to-potential-GDP ratio**

Deviation from cross-country and time-period averages, per cent



Note: The solid line shows the association between fiscal consolidation and the public debt to GDP ratio. Consolidations are defined as a 1.5% of GDP action-based consolidation effort: countries and period episodes are taken from Devries et al. (2011). Consecutive years of consolidation are dropped as in Alesina and Ardagna (2010). Driscoll-Kraay (1998) standard errors robust to heteroskedasticity, arbitrary spatial correlation and autocorrelation up to five years. Government debt and GDP data are taken from the *OECD Economic Outlook No. 92 Database*. See Appendix B, Section 4 for more details on the methodology.

boosting potential output, but it takes time for these effects to materialise. They could however be particularly strong, and materialise faster, in crisis countries where credible consolidation can carry them from a situation of high and rising indebtedness, elevated risk premia and low growth to a "good equilibrium" characterised by falling debt, lower risk premia and higher growth (Padoan et al., 2012).

3. The effects of consolidation instruments on other policy objectives

3.1. Other policy objectives

While the point of fiscal consolidation is to reduce debt, it cannot ignore other policy objectives. The present study looks at the extent to which fiscal consolidation can proceed while minimising adverse effects on short-term growth, preserving long-term prosperity, avoiding exacerbating income inequality in the short and long term and contributing to global rebalancing. In addition to being an objective in its own right, equity may influence the sustainability of fiscal adjustment programmes. Consolidation strategies perceived as inequitable are more likely to be reversed and to fail to reduce debt.

The distinction made here between short- and long-term effects does not relate to specific time spans but to adjustment processes. Short-term effects correspond to the direct impact of measures as they are implemented. Long-term effects describe their consequences when cyclical adjustment has run its course and behaviour has responded fully to the measures (meaning in particular that any general-equilibrium impacts have materialised).

3.2. Instruments

The instruments considered are policies that permanently affect government underlying primary spending and revenues. Government underlying primary spending is broken into ten categories, including four consumption items, three transfer items, subsidies, public investment (Table 1) and a residual item which is not considered as an instrument of consolidation. The expenditure breakdown broadly follows national accounts classifications with the difference that user charges are not netted out from government consumption. Instead, user charges are included among the eight consolidation instruments considered on the revenue side. Cutting tax expenditures, a potentially large and attractive source of revenue, is nevertheless not included as an instrument because of the lack of sufficiently reliable and internationally comparable data across countries. Section 6, however, discusses how reductions in tax expenditures can

Table 1. **Instruments of consolidation**

Expenditure cuts	Revenue increases
Public consumption: education	Personal income taxes
Public consumption: health	Social security contributions
Public consumption: other (except family)	Corporate income taxes
Cash transfers: pensions	Environmental taxes
Cash transfers: unemployment benefits	Consumption taxes (non-environmental)
Cash transfers: sickness and disability	Recurrent taxes on immovable property
Public consumption and cash transfers: family	Other property taxes
Subsidies	Sales of goods and services
Public investment	

Source: See Appendix B.

contribute to policy strategies that combine fiscal consolidation with structural reform. Appendix B, Section 2 provides details on the definition of the categories, on the sources used and on the methods employed to gather data from different sources in a way that adds up to government primary spending as recorded in national accounts.

3.3. The effects of instruments on objectives

An attempt is made at evaluating the effect of revenue increases and expenditure cuts on growth, equity and global rebalancing objectives. The effects of instruments on the current account are also evaluated because consolidation strategies should take into account co-ordinated efforts in multilateral settings such as the G20 to achieve balanced growth at the global level. For the purpose of this exercise, the instruments are assessed on their own without considering how their side-effects on long-term growth and equity could be minimised through structural reforms in the tax or spending area under consideration, other structural reforms or redistributive policies. The distinction between purely fiscal changes and structural reform is obviously not so clear cut in practice.[6] Still, it is useful insofar as it allows for an assessment of the side-effects that some consolidation instruments can imply for other policy objectives (this section) before discussing the benefits of joint policy strategies that combine consolidation with structural reform (Section 6).

The present assessment builds on previous work by the OECD and the wider literature complemented by new estimates presented in Barbiero and Cournède (2013). Table 2

Table 2. **Summary assessment of growth and equity effects of fiscal consolidation instruments**

	Growth		Equity		Current account[1]
	Short-term	Long-term	Short-term	Long-term	Short- to medium-term
Spending cuts					
Education	– –	– –	–	– –	+
Health services provided in kind	– –	–	–	–	++
Other government consumption (excluding family policy)	– –	+	–		+
Pensions		++			++
Sickness and disability payments	–	+	– –	–	++
Unemployment benefits	–	+	–		++
Family	–	–	– –	– –	+
Subsidies	–	++	+	+	+
Public investment	– –	– –			++
Revenue increases					
Personal income taxes	–	– –	+	+	+
Social security contributions	–	– –	–	–	
Corporate income taxes	–	– –	+	+	++
Environmental taxes	–	+[2]	–		+
Consumption taxes (other than environmental)	–	–	–		++
Recurrent taxes on immovable property	–				+
Other property taxes	–		++	+	+
Sales of goods and services	–	+	–	–	+

1. Current-account effects refer to a deficit country, and would switch sign in the case of a surplus country.
2. This + sign reflects positive welfare effects as the long-term impact on output narrowly defined as GDP may be ambiguous.

Source: See main text and Section 3 of Appendix B.

summarises this assessment which is based on the main points discussed immediately below while Appendix B, Section 3, provides additional supporting material on the growth and equity effects of consolidation instruments. Besides showing the estimated direction of the effect, some crude indications of the relative strength are also provided, based on empirical evidence.

3.3.1. *Long-term growth effects*

A number of fiscal consolidation instruments can enhance the long-term level of output by improving efficiency in the economy or the supply of resources. Evidence suggests that, in advanced economies, reducing the size of government increases long-term output although there is clearly no consensus on what constitutes the optimal size of the public sector even from a strict efficiency point of view (OECD, 2003; Cournède and Gonand, 2006; Barbiero and Cournède, 2013; Afonso and Tovar Jalles, 2013). This output-enhancing effect of reducing government spending is likely to be stronger in areas such as subsidies[7] where public expenditure frequently distorts the allocation of resources in the economy. Similarly, cuts in public spending that can prompt a positive response of labour utilisation, such as in pensions, are likely to have a particularly favourable effect on the long-term level of output per capita. Reductions in public spending on unemployment benefits can also boost employment and output per capita insofar as they do not bring unemployment benefits down to a level prompting inefficient employee-job matches that could curb productivity. Cuts in disability payments can boost labour utilisation (Hagemann, 2012) although this effect will arise only insofar as workers with significant residual capacity are receiving rehabilitation assistance.

Some revenue measures can also contribute positively to long-term output when they promote more efficient use or allocation of services or resources that were previously inadequately priced. To the extent that their current levels correspond to under-pricing, higher user charges reduce the waste of economic resources, thereby boosting productivity and output. Better pricing the use of environmental services through taxation can lead to welfare gains through improved environmental amenities that are not measured in GDP. While, as other forms of taxation, environmental taxes reduce labour supply and the accumulation of human-made capital, they also have long-term effects that go in the direction of boosting output compared with a baseline of wasteful use of environmental capital (de Serres et al., 2010). For instance, if no action is taken, climate change can involve large losses of physical and human capital as well as reduced productivity through more frequent and intense storms, rising sea levels, additional deaths from specific diseases (e.g. malaria) and deteriorating air quality (de Serres et al., 2010). Whether the net long-term output effect is positive is ambiguous conceptually and difficult to estimate empirically especially because of the very long lags involved.

In contrast, other consolidation instruments can reduce the productive potential of economies. At a general level, raising the tax burden tends to reduce factor supply and long-term output (OECD, 2003; Bouis et al., 2011). Evidence on the impact of the tax structure (Johansson et al., 2008; Bouis et al., 2011) indicates that taxes on mobile or adjustable production factors affect aggregate supply with particular severity. In the present classification of instruments, personal income taxes, social security contributions and corporate income taxes fall into this category. Other taxes such as value-added or consumption taxes have proven to exert still meaningful but less strong distortionary effects (Johansson et al., 2008).

Spending reductions can entail potentially large long-term losses in output when they cut into areas where governments provide particularly valuable public goods or growth-enhancing services that are insufficiently produced by market forces. Empirical evidence (OECD, 2003; Sutherland and Price, 2007) suggests that cuts in public investment or government spending on education broadly fall into this category. As developed in Section 6, cuts in government investment or education that respectively focus on low-externality projects or are accompanied by education reform can have more limited, or even favourable, growth effects. However, as mentioned earlier, the simple assessment summarised in Table 2 is concerned only with plain fiscal changes without structural reform, implying a lower provision of public goods and services. Cuts in health care can also reduce output per capita by reducing labour supply and productivity. When controlling for taxes, public health spending appears to have a positive, albeit moderate, effect on output per capita (Barbiero and Cournède, 2013).[8] Through its contribution to well-being, health spending is most likely to have additional positive welfare effects that are not measured in GDP.

Cuts in childcare can reduce output per capita primarily by depressing labour force participation (OECD, 2007). Reductions in family benefits have a more ambiguous effect on output per capita through three channels that work in opposite directions. Firstly, they can prompt greater labour-market participation, boosting output per capita. Secondly, such cuts can increase child poverty (Whiteford and Adema, 2007), hampering the formation of human capital and resulting in lower long-term output per capita. Thirdly, cuts in family benefits are likely to have a negative, albeit small, effect on fertility rates (OECD, 2011b).[9] Overall, the net effect of cuts in the aggregate of childcare and family benefits on long-term output per capita is most likely to be negative. Some consolidation instruments are likely to have neutral or very weak long-run effects on output. Such is the case of taxes with relatively low distortive effects, such as property taxes (Johansson et al., 2008).

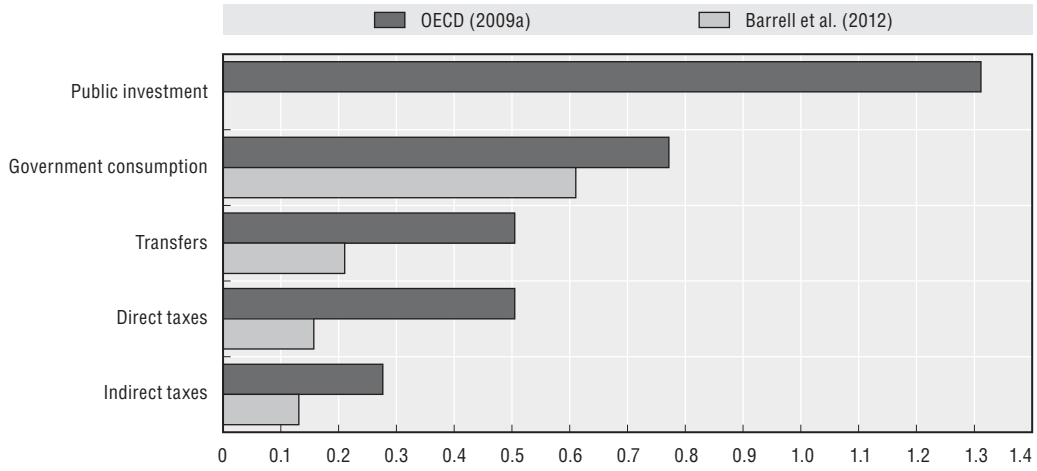
3.3.2. Short-term growth effects

Most fiscal consolidation instruments are harmful for growth in the short run, but there are differences among them and a few exceptions.[10] Although the vast literature on fiscal multipliers has not achieved consensus, international experience largely suggests that they are highest for public investment and government consumption and substantial but smaller for transfers and taxes (Figure 10; OECD, 2009a; Barrell et al., 2012). The main reason behind this difference is that changes in government investment and consumption affect activity directly while the effects of changes in taxes and transfers transit through the accounts of households and firms, offering greater possibilities for offset from saving behaviour. Consistent with this ranking, empirical evidence indicates that private-sector offsets from changes in government balances depend on their composition and are strongest for revenues, intermediate for spending and weakest for investment (Röhn, 2010).

The short-term output effects of instruments will depend on their design. In most cases, this design dependence does not preclude a broad assessment of their effect, but as far as cuts in pension spending are concerned, even the direction of the impact can change depending on how they are implemented. If cuts fall on current pensioners, they correspond to a reduction in transfers and are likely to affect output with a similar multiplier. In contrast, if pension spending is cut by raising the retirement age including for workers close to this age when the change is implemented, some positive demand effects

Figure 10. **Estimates of short-term fiscal multipliers for different consolidation instruments**

GDP contraction from a permanent 1 percentage-point increase in the underlying primary balance, per cent



Notes: The effects plotted in the chart are unweighted averages of country estimates reported in the quoted documents. The effect is averaged over the first and second years of consolidation for OECD (2009a) estimates and refers to the first year for Barrell et al. (2012) estimates. The simulations underlying Barrell et al. (2012) multipliers assume unchanged monetary policy in the year of the fiscal shock, but they incorporate the positive output effect of a fall in long-term interest rates resulting from the anticipation of a more accommodative monetary-policy path in the years following the shock. No multiplier estimate is available for public investment in Barrell et al. (2012).

are possible (Kerdrain et al., 2010) at the same time as supply expands, with an ambiguous net effect on the degree of economic slack.

In countries that are experiencing confidence crises because of their fiscal positions, the estimated multipliers reported above, which are calculated as historical averages, may not apply to their current circumstances. In fiscal-crisis countries, the absence of consolidation could translate into a massive loss of confidence triggering economic collapse. If it helps avoiding such extreme counterfactual scenarios, consolidation may be highly expansionary. There is also a possibility that, in such circumstances, different instruments may have different expansionary effects, notably by signalling the degree of determination of public authorities and thereby the likelihood that consolidation may be maintained. In particular, cuts in spending areas that raise serious political-economy challenges, such as subsidies, have been found to increase the probability of large consolidations to be successful (Molnar, 2012). There is however no consensus on the existence of these potential expansionary effects of consolidation, on their strength, on measuring when they may apply and how they may differ across instruments at a disaggregated level. For these reasons, these potential expansionary effects are not integrated in the assessment but should be seen as caveats regarding the extent to which the summary assessment presented in Table 2 applies to actual or potential crisis countries.

3.3.3. *Effects on equity*[11]

Many consolidation instruments work in the direction of aggravating income inequality (Table 2). Transfers in particular have strong redistributive power so that cuts in benefits are generally regressive, perhaps with the exception of public pensions where the equity effect is likely to be muted in countries where they are based on earned income and close to actuarial neutrality. Reducing the provision of public services likewise contributes

to increasing inequality in effective consumption (OECD, 2011d).[12] Also, a number of taxes fall more heavily on lower-income households, with the implication that increasing them would raise disposable income inequality.

Some fiscal consolidation instruments, on the other hand, can reduce income or wealth inequality. Such is particularly the case of hikes in inheritance and capital gains taxes, which the classification used in the present study includes among "other property taxes".[13] Increasing taxes that are typically designed to be progressive, such as personal income taxes, or concentrated on capital income, such as corporate income taxes, also goes in the direction of reducing disposable income inequality. The same holds for hikes in revenue instruments that are concentrated on capital income such as corporate income taxes (although some of their burden also falls on labour).

The equity implications of fiscal consolidation instruments can also evolve as behaviour responds to fiscal changes. Cuts in unemployment insurance payments, disability benefits or other social assistance programmes that are partly used as a way of withdrawing from the labour market can over time foster greater labour force participation. Since labour income tends to be greater than benefit payments, the supply response will work over time to reduce the regressive impact of cuts. On the tax side, environmental taxes, although they tend to be regressive in the short term, provide benefits that accrue in priority to low-income groups as those are more exposed to environmental degradation (Serret and Johnstone, 2006). Some of these effects, such as better health allowing greater labour supply, are reflected in higher measured income. Other often lagged effects such as improved wellbeing from better environmental conditions are not reflected in income distribution data. Consumption taxes, which are regressive in the short term because low-income households save a smaller share of their income than better-off ones, are neutral in a lifetime perspective taking into account the period when former savers spend what they previously accumulated. Finally, the redistributive benefits of some consolidation measures can wane over time as individuals put in place effective avoidance strategies as appears to be the case for inheritance taxes (Kopczuk, 2007).

3.3.4. Short- to medium-term effects on the current account

At a broad level fiscal consolidation works to push the current account towards a surplus over the short to medium term, but different instruments can have different effects depending on how they shape private saving and investment decisions. The impacts of individual consolidation instruments over and above the general macroeconomic effect are assessed based on the results reported in Kerdrain et al. (2010). Reductions in health care spending and in unemployment or disability benefits are likely to strengthen the current account through increased precautionary saving, whereas cutting pension benefits should lead to higher saving by the working-age population to smooth consumption over the life cycle. An increase in corporate taxation could improve the current account through lower investment (Schwellnus and Arnold, 2008; Vartia, 2008). Higher consumption taxes tend to penalise imports relative to exports, and thus may temporarily strengthen the current account, while the opposite holds for social security contributions.

3.4. A generic hierarchy of instruments

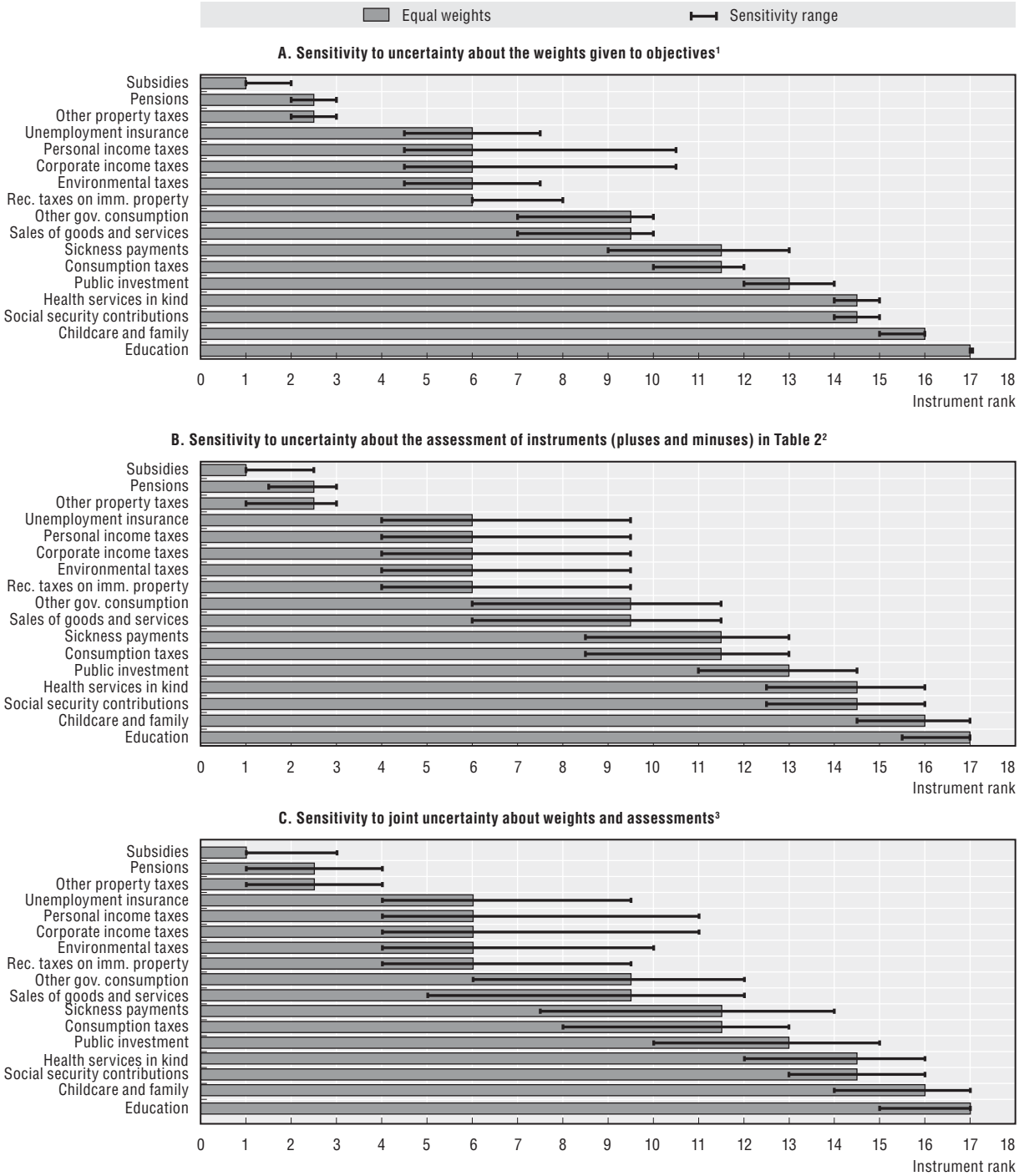
Based on the estimated impacts reported above, a generic hierarchy of consolidation instruments can be established (Figure 11). This is done simply by putting the same weight on each objective, assigning numerical values to the pluses and minuses and using the resulting scores to rank the instruments. The generic hierarchy puts no weight on the current account because the pursuit of global rebalancing operates in opposite ways depending on the sign of the imbalance and not at all in countries that have broadly balanced positions. Instead, current-account effects enter at a more country-specific level (see further below).

A long-term variant of the generic hierarchy can also be established for the purposes of looking solely at very long-term consolidation strategies by considering only long-term growth and equity effects, with equal weights. In this long-term variant, the instruments follow this ranking: 1. Subsidies; 2. Pensions; 3. Other government consumption, Unemployment benefits, Environmental taxes and Other property taxes; 7. Sickness and disability payments, Recurrent taxes on immovable property and Sales of goods and services; 10. Consumption, Personal income and Corporate income taxes; 13. Public Investment, Health services; 15. Family policy and Social security contributions; 17. Education.

Figure 11 also illustrates the sensitivity of instrument ranking to different weighting schemes and to uncertainty about the assessment of effects. When changing the weights attributed to objectives, a certain degree of sensitivity is indeed observed as instruments score differently across objectives, but the ranking of most instruments remains broadly stable in particular at both ends of the spectrum (Figure 11A). Similar robustness is observed when modifying the assessment of effects, even though the changes applied are strong, being equivalent to adding or withdrawing one plus or one minus sign in a full column of Table 2 (Figure 11B). Even combining these two sources of uncertainty leaves the ranking broadly stable, especially at both ends of the hierarchy (Figure 11C). Reductions in subsidies and in pension spending as well as increases in other property taxes come out robustly as preferred consolidation instruments. At the lower end, spending cuts in the areas of education, health care and family policy, as well as hikes in social-security contributions, appear as particularly unfavourable in terms of generating adverse side effects for growth and equity. In contrast, the middle part of the ranking is more fluid.

In addition to the arbitrary nature of the scoring and weighting scheme, considerable caveats surround the rankings above. They are based on an assessment of equity and growth effects of consolidation instruments which is drawn primarily from studies that estimate average effects in historical experience across countries. In practice, however, the growth and equity effects of instruments vary across countries: for instance, cutting investment in new roads in a country where highway density is already high should be less harmful to long-term growth than in a country with severe infrastructure gaps. Taking this cross-country variation into account is beyond the scope of this study, but it nonetheless goes beyond a pure one-size-fits-all approach. More specifically, the economic and social situation of countries in need of consolidation is taken into account by changing the weight of the different objectives, as is developed below. Also, the way in which the room for manœuvre is evaluated for each instrument takes into account whether or not the level of taxation or spending in this area is particularly high in the country under consideration.

Figure 11. **A possible generic hierarchy of consolidation instruments and its sensitivity to assumptions**



Note: The rankings are based on the assessment in Table 2. Scores of +1 and -1 are given to each + and - sign, respectively, each objective (except the current account) is given a weight, and the resulting indicator is used to rank instruments. The sensitivity range displays the 10th and 90th percentiles of the instrument rankings in random draws.

1. Randomly generated weights ranging each from 0.15 to 0.55 and summing to unity have been given to each objective in 10 000 draws. Weights have been restricted to be no smaller than 0.15 because each objective is considered important.
2. For deriving ranges, each individual instrument score along each objective shown in Table 2 is kept with a probability of 3/4 or increased by +1 with a probability of 1/8 or reduced by -1 with a probability of 1/8. A total of 10 000 random draws have been made.
3. Each individual instrument score based on the assessment in Table 2 is kept with a probability of 3/4 or increased by +1 with a probability of 1/8 or reduced by -1 with a probability of 1/8. Weights ranging each from 0.15 to 0.55 and summing to unity have been given to each objective. A total of 40 000 random draws have been made.

4. Adjusting instrument rankings for country circumstances over the short to medium term

The generic hierarchy is adapted to country-specific circumstances by adjusting the weights put on growth, equity and global rebalancing objectives. Summary indicators are defined for each of the growth, equity and current-account dimensions, and then used to compare country situations and form country groups. This makes it possible to derive a set of weights for each group and therefore a hierarchy of instruments for each group. While technically feasible, a country-specific ranking of instruments would give a false impression of accuracy with respect to country-specific instrument impacts and risk obscuring the substantial uncertainties and error margins of the exercise.

The group-specific rankings derived here will guide the choice of instruments for short- to medium-term consolidation efforts in the illustrative simulations. In the long run, however, a single hierarchy of instruments (presented in Section 3) is assumed to apply. As further addressed below, this is because some of the dimensions taken on board to form country groups lose relevance as the time horizon expands (e.g. short-run growth and current-account imbalances), while a solid basis is absent for giving differentiated weights to long-run growth impacts.

4.1. Characterising country circumstances

4.1.1. Short-run growth

This study attaches different weights to the short-run growth impacts of fiscal retrenchment depending on the degree of cyclical weakness faced by countries and their vulnerability to hysteresis.[14] A deeper negative output gap makes any short-run output losses from consolidation more painful, especially if fiscal multipliers of the Keynesian kind have become larger under such circumstances. Indeed, some recent studies find multipliers to be larger in recessions than expansions (Auerbach and Gorodnichenko, 2012; Baum et al., 2012), particularly in a context of financial crisis with monetary policy constrained by the zero nominal interest rate bound (IMF, 2010; Christiano et al., 2011; Corsetti et al., 2012). In turn, hysteresis effects could translate short-run slack into permanently lower levels of potential output through channels such as higher structural unemployment and a smaller capital stock (Bouis et al., 2012).

The degree of openness also has an influence on the magnitude of multipliers. The well-known inverse relationship between trade openness and multiplier size (OECD, 2009) could be invoked to give a lower weight to short-run growth impacts in more open economies. However, this would ignore the stronger negative spillover effects on partners' output that fiscal consolidation in those economies will tend to exert (Goujard, 2013). Consistent with ruling out a beggar-thy-neighbour approach to fiscal consolidation, the impact of openness on fiscal multipliers is therefore not taken into account.

The average of two variables, the output gap in 2012 and the 2007-12 percentage point change in the long-term unemployment rate, is used as a synthetic indicator of how different countries fare on the counts above. Long-term unemployment is used as a proxy of vulnerability to hysteresis, since it is a key variable in the transmission of short-run labour market slack to structural unemployment (Guichard and Rusticelli, 2010). It is taken in changes (and not levels) so as to capture impacts from the current crisis rather than pre-existing structural characteristics, which are better addressed through structural reforms in labour markets, as well as in product markets and tax and welfare systems.

4.1.2. *Long-term growth*

Assessing for which countries fiscal policy needs to be more supportive of long-run growth, with a concomitantly larger weight given to this objective, would be a hazardous task. Using weaker growth prospects as an argument for a larger weight runs into the difficulty that long-term growth projections are inevitably fraught with uncertainty and depend, to a significant degree, on policy assumptions in a wide range of areas, such as education, retirement age or product-market and trade regulations (Johansson et al., 2013). The long-term growth impacts of fiscal consolidation instruments are therefore deemed equally important for all countries.

4.1.3. *Income distribution*

The impacts of fiscal instruments on income distribution arguably gain increased prominence in more unequal countries. The links between inequality, growth and welfare are admittedly complex, and, to some extent, inequality differences across countries are rooted in social preferences, so that strong opposition to regressive changes might arise at comparatively low levels of inequality in strongly egalitarian societies. Still, beyond certain levels, inequality, and particularly poverty, may be bad for growth. Channels of transmission of inequality's detrimental effects include hampered investment in human capital, an area where inequalities can be self-perpetuating (Causa and Johansson, 2009; Hoeller et al., 2012). The Gini coefficient and the poverty rate (defined as income below 60% of the median) are combined into one indicator to summarise where countries stand as regards inequality. While the Gini coefficient encapsulates the whole income distribution, the poverty rate focuses on the lower tail. These two variables are computed after taxes and cash transfers, thus reflecting both the direct (on disposable incomes) and indirect (on market incomes) impacts of those fiscal tools on income distribution, though not the direct impact (on effective consumption) of in-kind transfers.

4.1.4. *Current account balance*

Addressing significant external imbalances is also a widely shared objective of economic policy (G20, 2009), which calls for taking account of the current-account impacts of different budget items when designing consolidation strategies. Imbalances carry risks for the individual countries concerned (the prospect of a hard landing for debtors, or growing credit risk for surplus countries), all the more so when they are particularly large, but also for the global economy (OECD, 2012). National positions are characterised on the basis of estimates of cyclically-adjusted current-account balances, which correct headline balances for the difference in output gaps between countries (Ollivaud and Schwellnus, 2013): a country facing a deeper downturn than its trading partners will temporarily tend to post a headline current account stronger than the adjusted one, as imports become more depressed than exports. The summary indicator used is the average of two variables: adjusted current-account balances in 2012 as percentages of both national and OECD GDP. The ratio of the cyclically-adjusted current-account balance to OECD GDP, which captures the absolute size of imbalances, serves as a proxy for their global implications which countries are assumed to internalise as part of the global rebalancing agenda.

4.2. *Hierarchies of instruments for groups of countries*

A cluster analysis has been performed to identify groups of countries that share similar characteristics regarding short-term growth, equity and external imbalances (Box 3).

Box 3. **Forming country groups and deriving group-specific weights for objectives**

Country groups capturing similarities along three dimensions (short-term growth, equity and the current account) are derived from a hierarchical cluster analysis based on the three summary indicators discussed in the text, with squared Euclidean distance to measure differences between groups, and a k-means algorithm used in a second stage to minimise the distance of each country to its cluster centre. The several variables used to characterise country circumstances as well as the ensuing summary indicators (repeated below for ease of reference) have all been normalised (i.e. set to zero mean and unit standard deviation) so that scale differences do not affect results:

- Short-term growth concerns have been summarised by the average of the output gap in 2012 and the 2007-12 percentage point change in the long-term unemployment rate.
- Equity concerns by the average of the Gini coefficient and the 60% threshold poverty rate (both applied to measures of household income after taxes and transfers) at the end of the 2000s.
- And current-account disequilibria by the average of cyclically-adjusted current-account balances in 2012 as a percentage of, respectively, national GDP and of OECD GDP.

The three co-ordinates of each cluster centre (one for each of the three summary indicators) have then been used to derive cluster-specific weights for short-term growth, equity (where the derived weight is applied to both the short and the long term) and the current account (as explained in the main text, the weight of the long-term growth objective is pre-set at a uniform 0.25 for all country groups).

These weights are derived in two steps. *First*, since weights cannot be negative, the cluster centre co-ordinates are transformed into values in the [0,1] interval while preserving their ordering (from smallest to largest). This is done for each summary indicator by fitting a logistic cumulative distribution function (F) to the co-ordinates (x) of the several cluster centres. The current-account summary indicator has been handled differently, both to treat in a symmetric way deficits and surpluses, and to focus on large rather than small imbalances: F has been fitted to the log of the indicator's absolute value, and the smaller ensuing $F(x)$ values have been reset to zero (in clusters comprising up to half of the total number of countries under analysis).

Second, the three $F(x)$ values obtained for each cluster have been renormalised, so that weights for each cluster add up to one. More precisely, for cluster c, the weights of the four dimensions, $d = 1,\ldots,4$ (counting short-term and long-term equity separately) are derived from the $F(x_{cd})$ values (again, using twice the $F(x)$ value for equity) so as to sum to 0.75 (one minus the weight of long-term growth):

$$w_{cd} = \frac{3}{4} \times \frac{F(x_{cd})}{\sum_{i=1}^{4} F(x_{ci})}$$

The ensuing weights (w) are displayed in Table 3. It should be borne in mind that weights compare the relative importance of the several dimensions considered *within* a cluster, but not *across* clusters. It is therefore possible that a cluster with pressing concerns on several fronts (like clusters No. 2 and 3) will display smaller weights for some objectives than a cluster where those same objectives, though less problematic in a cross-cluster comparison, are the sole concern which stands out (such as short-term growth in cluster No. 4).

Based on the summary indicators discussed above, five clusters have been identified:

1. The first cluster is formed by 11 geographically-dispersed countries (Australia, Canada, Estonia, Israel, Italy, Japan, Korea, New Zealand, Poland, Portugal and the United Kingdom), which mainly have in common above-average levels of inequality. Short-term growth risks are generally moderate (Italy and Portugal being exceptions) and current-account positions, though with considerable heterogeneity, do not include cases of extreme imbalances and are on average fairly close to balance.
2. The United States finds itself alone in the second cluster, as the sheer absolute size of its current account deficit places it at a considerable distance even from other deficit countries. Inequality is high and cyclical developments carry potentially substantial hysteresis risks although the materialisation of these risks would run counter to historical experience.
3. The third cluster comprises three euro area members from the geographical periphery (Greece, Ireland and Spain) sharing very high cyclical slack and hysteresis risks. Greece and Spain (but not Ireland) also display above-average inequality and large underlying external deficits.
4. A fourth cluster is formed by 11 European countries: Austria, Belgium, the Czech Republic, Denmark, Finland, France, Hungary, Iceland, Norway, Slovakia and Slovenia. It is the most egalitarian cluster. As in the first group of countries, current-account imbalances are on average small, though with significant intra-group heterogeneity,[15] and short-term growth risks are generally moderate.
5. The fifth and final cluster comprises five countries, Germany, Luxembourg, the Netherlands, Sweden and Switzerland, all with large current-account surpluses. Inequality levels are below-average and short-term growth vulnerability risks are among the lowest in the OECD.

For each of these clusters, specific weights are calculated for the short-term growth, equity and current-account objectives (Table 3). The weights depend on the degree to which each objective is relevant for the cluster as a whole in comparison with the other objectives (but do not compare the importance of each objective across different clusters of countries). For instance, short-run growth will attract a strong weight in groups of countries where cyclical weakness and hysteresis risks – whether very high in themselves (cluster 3) or only moderate (cluster 4) – are clearly a more important concern than equity or current-account issues. Similarly, the high weight attached to the current-account objective in cluster 5 stems from the contrast between large surpluses and mostly benign short-term growth and equity outlooks. As mentioned above, the same weight (25%) is given to long-term growth in all clusters. These cluster-specific weights are used to aggregate the pluses and minuses reported in Table 2 after converting them into numerical values, give scores to instruments and rank them.

Table 4 displays the ensuing cluster-specific instrument rankings. Rank variation across country groups is widest for instruments with the sharpest trade-offs between growth, equity and the current account. For instance, personal and corporate income taxes come out as good candidate instruments for cluster 1, where equity objectives carry a high weight, but much less so for countries such as those forming clusters 4 and 5 where relatively equal income distribution is assumed to lead to less emphasis on outcomes in this area. In the cluster 5, the ranking of corporate income tax hikes is brought down further by their likely detrimental effects on investment which complicate external surplus

Table 3. **Weights put on the growth, equity and current account dimensions across groups of countries**

Cluster		Growth		Equity		Current account
		Short-term	Long-term	Short-term	Long-term	Short-term
1	Australia, Canada, Estonia, Israel, Italy, Japan, Korea, New Zealand, Poland, Portugal, United Kingdom	0.13	0.25	0.30	0.30	0.01
2	United States	0.13	0.25	0.21	0.21	0.20
3	Greece, Ireland, Spain	0.29	0.25	0.18	0.18	0.10
4	Austria, Belgium, Czech Republic, Denmark, Finland, France, Hungary, Iceland, Norway, Slovak Republic, Slovenia	0.47	0.25	0.14	0.14	0.00
5	Germany, Luxembourg, Netherlands, Sweden, Switzerland	0.12	0.25	0.15	0.15	0.33

Table 4. **Possible hierarchies of consolidation instruments for groups of countries**

Instruments	Generic ranking	Cluster-specific ranking					Long-term ranking
		1	2	3	4	5	
Subsidies	1	1	1	2	2	1	1
Pensions	2-3	3	2	1	1	3	2
Other property taxes	2-3	2	3	3	3	2	3-6
Unemployment benefits	4-8	7	4	4	4	9	3-6
Personal income taxes	4-8	5	8	9	9-10	8	10-12
Corporate income taxes	4-8	4	5	7	9-10	12	10-12
Environmental taxes	4-8	8	6	5	4	4	3-6
Recurrent taxes on immovable property	4-8	6	7	6	6	5	7-9
Other government in kind consumption	9-10	9	9	11	11	6	3-6
Sales of goods and services	9-10	10	10	8	7	7	7-9
Sickness and disability payments	11-12	13	11	10	8	11	7-9
Consumption taxes (other than environmental)	11-12	11	12	12	12	13	10-12
Public investment	13	12	13	13	15	15	13-14
Health services provided in kind	14-15	14	14	14	16	16	13-14
Social security contributions	14-15	15	16	15	13	10	15-16
Family	16	16	15	16	14	14	15-16
Education	17	17	17	17	17	17	17

Note: The rankings are based on the assessments in Table 2 with scores of +1 and –1 given to each + and – sign, respectively, and weights taken from Table 3. The current-account scores of Table 2 switch sign for surplus clusters. The long-term ranking in the final column is based on equal weights given to impacts on long-term growth and equity. Cluster 1 regroups Australia, Canada, Estonia, Israel, Italy, Japan, Korea, New Zealand, Poland, Portugal and the United Kingdom. Cluster 2 includes only the United States. Cluster 3 comprises Greece, Ireland and Spain. Cluster 4 is formed by Austria, Belgium, the Czech Republic, Denmark, Finland, France, Hungary, Iceland, Norway, Slovakia and Slovenia. Cluster 5 is made up by Germany, Luxembourg, the Netherlands, Sweden and Switzerland.

reduction. In contrast, increases in social contributions move up a few ranks due to their effect that goes in the direction of reducing current-account surpluses via labour costs.

Short-term growth impacts do not vary considerably across instruments (as there are no positive effects) and are therefore a less powerful determinant of rank variation. Their influence is nonetheless visible, for instance in the lower position of other government consumption in clusters 3 and 4 where short-term contractionary effects carry more weight. Rank variation across country groups is smallest for the instruments that have similar impacts on virtually all objectives, such as education, subsidies or property taxes.

4.3. Robustness of country-specific rankings to the choice of a clustering technique

While there is something intrinsically mechanical to the derivation of weights from a mathematical clustering technique, the ensuing instrument hierarchies are reasonably robust to shocks to weights as suggested by the relative stability of the ranking plotted in Panel B of Figure 11. This robustness has been checked further in a variant exposed in Section 5 of Appendix B: countries have been divided in three groups solely on the basis of their current-account positions and equal weights have been given to each objective within each group. The instrument hierarchies that result from this alternative weighting scheme are very highly correlated with the clustering-based ones plotted in Table 4: the rank correlations between alternative and clustering-based hierarchies are greater than 93% in all clusters except cluster 4 where it is 89%.

5. How far down instrument rankings do countries need to go? Some illustrative simulations

In this section simulations are performed to investigate how far down instrument rankings countries may need to go in order to meet their consolidation needs. Countries are assumed to implement budget tightening according to the relevant instrument ranking, i.e. to start by adjusting the most beneficial (or least detrimental) instrument and only proceed down the list after exhausting the estimated room for manœuvre available in the preceding instrument. In practice, implementing this approach would raise political-economy challenges: the top ranking instruments tend to be either streams of spending accruing to politically powerful constituencies, such as subsidies or pensions, or forms of taxation where planned increases often meet with strong resistance, such as property taxes. Nevertheless, it may still provide a useful benchmark for considering a consolidation strategy.

The analysis is conducted separately for the short to medium term and for the long term, and requires three building blocks, themselves differentiated according to the respective time dimension: *i)* estimated consolidation needs for both horizons, as presented in Section 2; *ii)* a hierarchy of instruments, which is common to all countries in the long run simulation (as presented in Section 3) but varies across clusters in the short to medium term (Table 4 and Section 4); *iii)* estimates of the available margin for adjustment in each instrument, which is discussed next.

5.1. Room for manœuvre in instruments

Although it is an important building block when drawing up an illustrative consolidation plan, estimating the room for manœuvre for each policy instrument – or, put differently, the margin of feasible adjustment – is necessarily judgemental. As such, it can only be done in an indicative and approximate way that is to some degree arbitrary. In a cross-country setting, it is impossible to fully account for the economic circumstances, social preferences and institutions which, in each country, shape the relative size of budget items. At one extreme, it could be assumed that the current structure of budgets already equalises the marginal costs and benefits of adjusting the different instruments (whose growth and equity impacts vary across countries, as acknowledged above), and is therefore optimal. If so, consolidation should be pursued, at least at the beginning, through a proportional adjustment of budget items. At another extreme, the budget structure status quo, hard to change as it may be, could be viewed as the sub-optimal outcome of political and institutional distortions, the correction of which would require sweeping changes. For

instance, it could be the case that property taxes should be increased further even in countries where they are already high by international comparison.

This exercise attempts to strike a balance between the above considerations by assuming that there is some margin, albeit limited, to scale back expenditure items that are large relative to a significant number of other OECD countries and similarly to increase revenue streams that are relatively low. One reason for not pushing adjustment along each individual item too far is that the positive and negative assessments underpinning the rankings can be expected to be most reliable in standard situations. The effects may change if adjustment along one item takes a country to a more extreme situation. For instance, up to a point, reducing spending on unemployment benefits improves incentives to take up a new job and boosts long-term output through higher employment, but if cuts are pushed too far they can impair the quality of labour-market matches and harm output through lower productivity while also resulting in insufficient macroeconomic stabilisation. On the tax side, marginal rate increases from a high starting point are more distortive than from a low-rate baseline. At the same time, social preference and political feasibility considerations call for putting an upper bound on the amount of change to any spending cuts (tax hikes) in a given item, no matter how high (low) the departure point is.

In operational terms, simulations assume that room for manœuvre exists in a revenue instrument if a country is below the 66th percentile in the cross-country distribution of cyclically-adjusted receipts from this instrument as a share of potential GDP. In other terms, there is room to increase a given tax if one-third of OECD countries raise more income, relative to potential GDP, with this tax. Similarly, room for manœuvre on the spending side exists if a country is above the 33rd percentile in the cross-country distribution of cyclically-adjusted spending on this instrument as a share of potential GDP. The room for manœuvre is given in the general case by: *i)* the gap between the value in the country under consideration and the 66th or 33rd percentile; or by *ii)* the standard deviation of the cross-country distribution of the instrument at hand, whichever is smallest (Box 4). It turns out that this simulation design imposes only a moderate degree of convergence in budget structures across countries (Section 5.3).

A few additional adjustments have been made to improve comparability and minimise the mechanical nature of the simulations (described below and presented in technical form in Box 4):

- Spending on pensions, education and unemployment benefits as a share of potential GDP has been corrected for the number of potential beneficiaries, defined in terms of age cohorts or labour market status. For instance, this acknowledges that, all else being equal, a higher rate of structural unemployment implies a smaller room for manœuvre in reducing the unemployment benefits bill.
- Further to the above correction, a special adjustment is made to reduce the available room for cuts in pension spending to acknowledge that the baseline already incorporates significant effort. More specifically, the room for manœuvre is set at the standard deviation of the cross-country distribution minus the projected increase until 2060 in government pension expenditure on unchanged policies (Figure 4). This increase is taken as an estimate of the implicit effort in keeping baseline pension spending constant as a share of potential GDP. In the short- to medium-term, the room for manœuvre is set at a quarter of its long-run value, as the budget savings yielded by most measures

Box 4. **Simulation design: technical details on instrument adjustment**

This box sets out in more detail the rules governing the use of fiscal instruments in simulations. It is assumed that country *i* has some room for manœuvre (*RM*) in revenue item *T* when its value T_i (adjusted for the cycle and one-offs and taken as a percentage of potential GDP) stands below the cross-country 66th percentile for that item (T^{P66}). Room for manœuvre is then given by the minimum of the cross-country standard deviation (σ_T) and the gap to the 66th percentile ($T^{P66} - T_i$), except if the latter falls below a certain threshold, in which case a minimum scope for adjustment (set at a quarter of the standard deviation) is admitted. This tries to avoid cases of either implausibly large or pointlessly small instrument adjustment. Formally:

$$RM_i(T) = \begin{cases} \min[\sigma_T, \max(T^{P66} - T_i, \sigma_T/4)], T_i < T^{P66} \\ 0, T_i \geq T^{P66} \end{cases}$$

On the spending side, room for manœuvre exists when the country value (G_i) exceeds the 33rd percentile (G^{P33}) and is defined along similar lines:

$$RM_i(G) = \begin{cases} \min[\sigma_G, \max(G_i - G^{P33}, \sigma_G/4)], G_i > G^{P33} \\ 0, G_i \leq G^{P33} \end{cases}$$

Cross-country distributions refer to 2012 for the short- to medium-term simulations (see Appendix A for a presentation of the dataset) and to 2060 for the long-term simulations. Because in the baseline most fiscal variables are assumed to remain constant as a percentage of (potential) GDP, their cross-country distributions will be identical in both years, and so will countries' room for manœuvre. The only exception is health care, where long-term pressures are taken into account as described in Section 2, giving rise to larger spending relative to GDP in 2060 than in 2012.

Room for manœuvre in a total of five fiscal instruments, three on the spending side and two among revenues, is the object of additional adjustments, thus departing somewhat from the above formulas. To take account of country differences in demographic and labour-market circumstances, expenditure on pensions, education and unemployment benefits is adjusted for the number of potential beneficiaries, defined respectively as individuals aged 65 and over, aged 25 and below, and those unemployed when the economy is operating at potential. In each of the three cases, a scaled variable G_i/R_{Gi} is constructed, where R_{Gi} is the ratio of potential beneficiaries to total population; this scaled variable corresponds to spending per potential beneficiary relative to GDP per capita, and therefore, as regards pensions and unemployment benefits, bears some resemblance to a replacement rate. The cross-country distribution of G/R_G yields estimates of room for manœuvre – $\overline{RM}_i(G/R_G)$ – by application of the formula above in the case of education and unemployment benefits, whereas for pensions, in addition, a smaller upper bound on long-run spending reductions is set (σ minus the 2012-60 change – when positive – in government pension spending on unchanged policies, also scaled by R_{Gi}). Room for manœuvre scaled back to percentage points of potential GDP is then given by:

$$RM_i(G) = R_{Gi}.\overline{RM}_i(G/R_G)$$

As regards education and unemployment benefits, these adjustments are performed separately for 2012 and 2060 and therefore introduce some differentiation in cross-country distributions (and ensuing room for manœuvre) between the two simulation horizons. In the area of pensions, the room for manœuvre in the short to medium run is assumed to be four times smaller than in the long run, the latter being computed as described above.

Box 4. **Simulation design: technical details on instrument adjustment** *(cont.)*

Long-run population projections are taken from Eurostat and United Nations sources as in Johansson et al. (2013) while projected NAIRU and trend labour force come from the *OECD Economic Outlook*, May 2013 long-term database.

The room for manœuvre in personal income tax (PIT) and social security contributions (SSC) is assessed on the basis of the cross-country distribution of their sum according to the general formula above. The individual distributions of PIT and SSC are nonetheless taken into account and impose additional constraints on increases in these items, which cannot exceed one standard deviation of the respective individual instrument nor take place if that instrument already exceeds the 66th percentile of its cross-country distribution.

When instruments are tied in rankings (which is rare in the short- to medium-term simulations, but relatively common in the long run), spending items are used before revenue items, as spending-based fiscal adjustments are often found to be more durable (Alesina and Ardagna, 2012; Molnar, 2012). In case of ties between two spending instruments (or two revenue ones) both are used in proportion to the respective room for manœuvre.

Besides the 17 instruments considered, property income and residual items (both on revenue and expenditure sides) also stem from the budget decomposition performed (Section 3.2 and Appendix B). Residual items are not regarded as instruments, since their heterogeneous character makes it hard to assess their growth and equity impacts. However, they are assumed to take part in the overall consolidation effort by keeping a (broadly) constant share in total receipts and disbursements. More precisely, if after a given revenue (spending) instrument has been used the consolidation needs of the country at hand have not yet been fully met, before moving to the next instrument the residual revenue (spending) item is increased (decreased) so as to restore it to its baseline share in total underlying primary revenue (expenditure).

(e.g. raising the retirement age, or lowering the replacement rate for new retirees) will only accrue gradually over time.

- Leeway for raising personal income tax and social security contributions is assessed on the basis of the cross-country distribution of their sum, taking account of the wide heterogeneity across countries in the balance between those two revenue items which are often largely substitutes. For instance, a country that raises very low amounts of social contributions may nevertheless have little room for manœuvre along this instrument if it has very high personal income taxation, as is the case in Denmark.

5.2. *Meeting consolidation needs*

5.2.1. *Short- to medium-term consolidation needs*

Under the simulation design outlined above, almost all countries have the scope to meet their short- to medium-term consolidation needs within the constraints put on instrument use. The only exception is Japan where the constraints imposed by the chosen simulation design limit consolidation to 15% of GDP against an estimated need of 18½ per cent of GDP. This discrepancy implies that, in practice, the constraints imposed on instrument use would have to be eased.

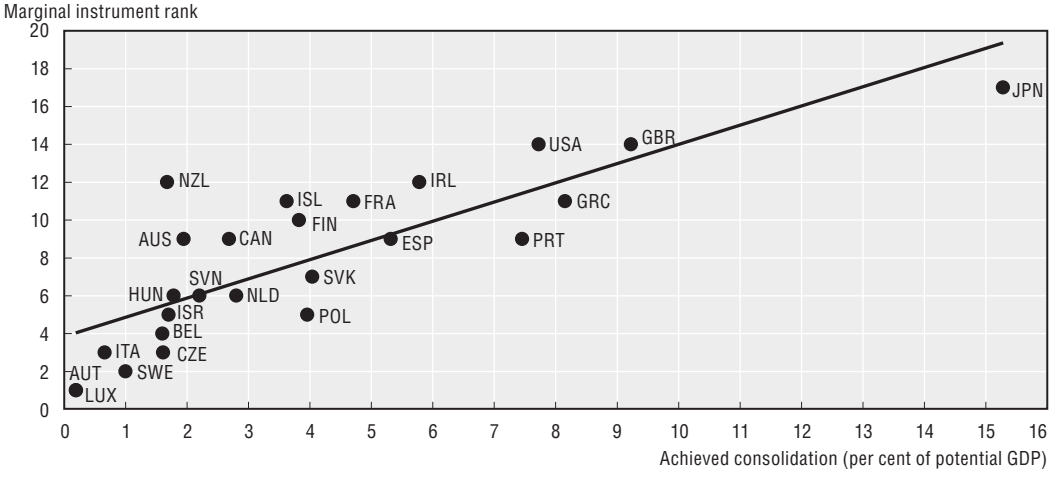
However, even when fully meeting consolidation needs, several countries are forced to resort to a significant degree to instruments which lie towards the bottom of their respective instrument hierarchies (Table 4), and thus generally have an overall detrimental

impact on objectives. On the basis of the marginal (i.e. worst) instrument used (Figure 12) as well as the full consolidation packages pursued (Tables A1 to A5, Appendix A), three groups of countries can be identified:

- Sixteen countries (Australia, Austria, Belgium, Canada, Czech Republic, Hungary, Israel, Italy, Luxembourg, Netherlands, Poland, Portugal, Slovak Republic, Slovenia, Spain and Sweden) only need to use instruments featuring in the top half (first nine places) of their respective cluster-specific rankings. All these countries have short- to medium-term consolidation needs which do not exceed 3 percentage points of potential GDP. Though the simulated adjustment is not without economic costs, these will be mainly of a Keynesian nature, while negative impacts on equity or on long-term growth will be absent or, at worst, limited.
- Six countries (Finland, France, Greece, Iceland, Ireland and New Zealand) use marginal instruments placed in the lower half of the respective cluster-specific hierarchies (ranked 10th or worse), but manage to implement consolidation packages where more than 50% of the adjustment comes from instruments in the upper half. While the use of detrimental instruments remains moderate, fiscal tightening will entail costs which go beyond short-run aggregate demand, raising concerns about impacts on equity and long-term growth.
- Three countries (Japan, United Kingdom and the United States) have to resort to marginal instruments ranked 14th or worst, with more than 50% of pursued consolidation packages consisting in the use of instruments placed in the lower half of rankings. Short- to medium-term consolidation therefore presents considerable challenges for these countries as it appears difficult to avoid potentially strong detrimental effects on both growth and equity.

Among the countries covered in this study, six do not need any short- to medium-term fiscal tightening (Denmark, Estonia, Germany, Korea, Norway and Switzerland) so no packages have been simulated for them.

Figure 12. **Marginal instrument rank and achieved short- to medium-term consolidation**



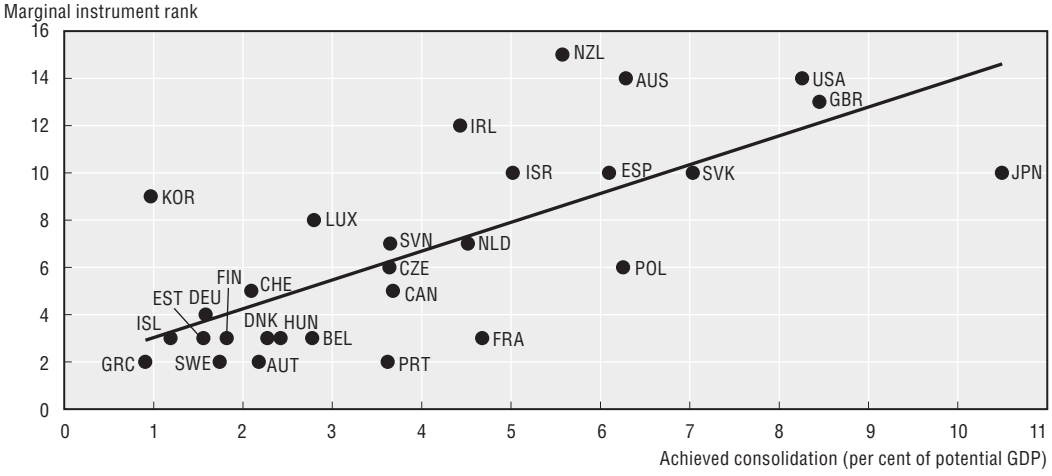
5.2.2. *Long-term needs*

Despite generally larger consolidation needs in the long run, all countries can meet them fully while complying with the constraints imposed by the simulation design. However, as with short- to medium-term consolidation packages, there is a risk of significant negative impacts on equity and long-term growth for some countries. As above, one can identify three groups of countries according to their marginal instrument (Figure 13) and full consolidation packages (Appendix Table A6):

- Twenty countries (Austria, Belgium, Canada, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Korea, Luxembourg, Netherlands, Poland, Portugal, Slovenia, Sweden and Switzerland) with low or moderate consolidation needs enjoy the favourable position of only having to use instruments in the upper half (top nine places) of the uniform long-run hierarchy, of which the overall impact on long-run growth and equity can be deemed mostly beneficial or fairly neutral.[16]
- Six countries (Ireland, Israel, Japan, Slovak Republic, Spain and the United Kingdom) resort to marginal instruments in the lower half of the ranking (10th to 17th places), which may entail more detrimental consequences for growth and equity objectives. However, these countries have consolidation packages where more than half (and in some cases virtually all) of the adjustment comes from better instruments (those in the upper half of the hierarchy).
- Three countries (Australia, New Zealand and the United States) with large long-term consolidation needs face the unpleasant prospect of both employing low-quality marginal instruments and letting poor instruments (those in the lower half of the ranking) account for more than half of the total fiscal adjustment. Therefore this group faces a substantial risk of overall negative impacts of consolidation on growth and equity.

With the assumptions outlined above, Italy and Norway have no estimated long-term consolidation needs and therefore no simulated packages at that horizon.

Figure 13. **Marginal instrument rank and achieved long-term consolidation**



Despite estimated consolidation needs being generally larger in the long than the short run, more countries rely fully on well-ranked instruments in their simulated long-term packages than in the short- to medium term ones. One reason is that other

government consumption, an area with substantial consolidation potential in many countries, is much better ranked in the long term when demand effects are no longer taken into account. Another reason is that the simulations are designed to offer more room for adjustment in public pension spending in the long than the short term, reflecting that expenditure savings in general accrue gradually in this area. Finally, more countries are estimated to face positive consolidation needs in the long than the short term.

At either simulation horizon, countries facing the unpleasant prospect of having to resort on a large scale to low-ranking instruments have two non-mutually exclusive options. The first, further discussed in Section 6, is to supplement the use of such instruments by structural changes that make them more growth- or equity-friendly. The second option is to use the best instruments more intensively than implied by the somewhat arbitrary constraints. The simulation design implies that countries such as Australia, New Zealand and the United States which start out with an above-average use of the least detrimental forms of taxation or below-average spending in the least effective areas tend to lack room for manœuvre in the best budget instruments. If the constraint that adjustment cannot take a country into the group of the ten OECD countries that tax most or spend least in the area of under consideration is relaxed by moving from ten- to five-country reference groups, then New Zealand and the United States achieve close to half of their simulated long-term consolidation with well-ranked instruments. With this relaxation of the constraint, Australia carries out three-quarters of its simulated long-term adjustment with well-ranked instruments.

In either simulation horizon, countries facing the unpleasant prospect of having to rely heavily on detrimental instruments have two non-mutually exclusive options. The first, further discussed in Section 6, is to supplement the quantitative use of such instruments by structural changes that make them more growth- or equity-friendly. The second option is to increase the room for manœuvre in the best instruments, by implementing larger spending cuts or tax hikes in those budget items already being used, or accepting to increase revenue items (cut expenditure items) which do not rank low (high) in international comparison. The simulation design implies that countries such as Australia, New Zealand and the United States which start out with an above-average use of the least detrimental forms of taxation or below-average spending in the least effective areas tend to lack room for manœuvre in the best budget instruments. Relaxing the constraints on instrument use makes it possible for these countries to achieve more of their consolidation with better instruments, as apparent in the alternative simulation summarised in Section 5.4 and reported in Section 6 of Appendix B.

5.3. *Consolidation packages and shifts in budget structure*

The sequential nature of instrument use in the simulations, based on hierarchies which have strong resemblances across groups of countries (in the short to medium term) or are even common to all countries (in the long run), results in some instruments featuring much more often than others in consolidation packages. As a consequence, revenue and expenditure structures evolve and undergo some convergence across countries. While this sub-section discusses general trends across countries, Appendix A, Tables A1 to A6 provide detailed information by country about the illustrative consolidation packages.

In the short to medium term, subsidies and other property taxes are the most widely used instruments, and together account, on average, across countries for about a third of

consolidation packages (Table 5). This large figure, however, should be kept in perspective, as it is heavily influenced by the fact that some countries with small short-term consolidation needs, such as Austria, Italy and Sweden, use subsidy reductions as the prime or even the sole adjustment tool. Spending reductions on unemployment benefits and pensions as well as increases in environmental taxes, corporate and personal income taxes, and recurring property taxes come next in frequency of use. Cuts in the areas of health, education and family policy are very rare in simulated packages, as are increases in

Table 5. **Summary indicators about consolidation packages**

Description	Memorandum items				Number of countries resorting to instrument	
	Baseline average (% of potential GDP)		Baseline standard deviation (pp potential GDP)			
	Short-term	Long-term	Short-term	Long-term	Short-term	Long-term
Subsidies	1.2	1.2	0.8	0.8	14 (13-15)	15 (14-15)
Pensions	8.1	8.1	3.3	3.3	11 (9-12)	12 (12-12)
Other property taxes	0.7	0.7	0.6	0.6	16 (12-16)	11 (8-14)
Unemployment benefits	1.1	1.1	0.9	0.9	11 (6-12)	13 (7-15)
Personal income taxes	8.6	8.6	3.3	3.3	9 (5-12)	7 (0-8)
Corporate income taxes	2.9	2.9	0.9	0.9	11 (4-14)	5 (1-6)
Environmental taxes	2.3	2.3	0.7	0.7	13 (7-15)	11 (10-14)
Recurrent taxes on immovable property	1.3	1.3	1	1	9 (3-12)	4 (3-8)
Other government in kind consumption (excluding family policy)	8.4	8.4	2.4	2.4	8 (4-12)	14 (11-17)
Sales of goods and services	2.8	2.8	1	1	7 (3-11)	7 (6-9)
Sickness and disability payments	2	2	0.6	0.6	4 (1-9)	7 (5-12)
Consumption taxes (other than environmental)	9	9	2.4	2.4	4 (3-10)	8 (3-9)
Public investment	2.6	2.6	1	1	4 (1-6)	4 (2-6)
Health services provided in kind	6.5	9.6	1.4	1.3	3 (2-4)	4 (3-7)
Social security contributions	11.2	11.2	5.4	5.4	1 (1-3)	0 (0-3)
Family	2.4	2.4	1.1	1.1	0 (0-2)	1 (0-3)
Education	5.3	5.3	1.1	1.1	1 (1-2)	0 (0-1)

Note: Instruments are ranked as in Figure 11. ST and LT denote, respectively, short to medium term and long term. All figures in the table refer to the 24 countries common to both simulation horizons. Average shares of instruments are computed across national consolidation packages (Tables A1 to A6 in Appendix A). Policy reversals (cases of stronger instrument use in ST than in LT) exclude cases solely due to a smaller room for manœuvre (i.e. in both ST and LT instrument use exhausts the available room for manœuvre, which is smaller in LT than in ST). Simulated interdecile intervals shown between brackets are computed using 200 random draws. Each individual assessment (i.e. each cell) in Table 2 is increased by one, decreased by one or retained with probabilities 1/8, 1/8 and 3/4, respectively.

Table 5. **Summary indicators about consolidation packages** *(cont.)*

Description	Average use among countries resorting to instrument (% of potential GDP)		Average share of instrument in total consolidation (%)		Policy reversals	
					Number of cases	Average reversal (pp of potential GDP)
	Short-term	Long-term	Short-term	Long-term	Short-term	Long-term
Subsidies	0.6 (0.6-0.6)	0.6 (0.6-0.6)	19.8 (15.6-23.9)	15.3 (13.4-15.3)	0 (0-2)	0 (0.0-0.3)
Pensions	0.5 (0.4-0.5)	1.7 (1.5-1.8)	8.2 (3.4-9.5)	22.1 (16.0-23.5)	0 (0-1)	0 (0.0-0.2)
Other property taxes	0.4 (0.4-0.5)	0.4 (0.4-0.5)	10.1 (6.4-12.4)	3.6 (2.0-6.7)	8 (3-9)	0.3 (0.2-0.4)
Unemployment benefits	0.6 (0.5-0.7)	0.5 (0.4-0.6)	7.7 (2.8-9.3)	6.3 (3.3-10.5)	3 (0-3)	0.4 (0.0-0.7)
Personal income taxes	1.9 (1.8-2.3)	1 (0.0-1.5)	11 (5.5-19.2)	3.5 (0.0-7.0)	6 (5-9)	2 (1.6-2.3)
Corporate income taxes	0.5 (0.3-0.5)	0.2 (0.0-0.5)	4.2 (1.0-7.6)	0.8 (0.0-1.9)	10 (3-11)	0.4 (0.3-0.5)
Environmental taxes	0.6 (0.4-0.6)	0.5 (0.5-0.6)	8.1 (2.6-10.4)	4 (3.0-6.9)	6 (2-7)	0.3 (0.1-0.5)
Recurrent taxes on immovable property	0.8 (0.6-0.8)	0.7 (0.4-0.8)	7.6 (1.2-11.1)	2.5 (0.8-4.7)	6 (2-9)	0.8 (0.5-0.8)
Other government in kind consumption (excluding family policy)	1 (1.0-1.5)	1.4 (1.3-1.5)	7.2 (3.6-18.8)	20.5 (13.2-27.3)	4 (1-5)	0.6 (0.7-1.3)
Sales of goods and services	0.6 (0.5-0.7)	0.7 (0.4-0.8)	3 (0.7-6.5)	3.5 (1.2-4.6)	2 (1-6)	0.3 (0.1-0.4)
Sickness and disability payments	0.4 (0.1-0.5)	0.5 (0.4-0.6)	1.4 (0.1-4.7)	2.8 (1.5-4.5)	2 (0-4)	0.4 (0.0-0.4)
Consumption taxes (other than environmental)	1.9 (1.4-2.1)	1.3 (1.0-1.8)	3.7 (3.0-11.6)	5.9 (2.2-10.2)	2 (2-6)	0.7 (0.5-1.6)
Public investment	0.5 (0.5-1.1)	0.5 (0.4-0.7)	1 (0.3-3.5)	1.4 (0.5-2.4)	3 (1-3)	0.5 (0.1-1.1)
Health services provided in kind	1.4 (1.2-1.5)	0.6 (0.4-1.0)	1.8 (1.1-2.3)	1.5 (0.7-3.9)	0 (0-0)	0 (0.0-0.0)
Social security contributions	0.9 (0.8-1.6)	0 (0.0-0.7)	0.2 (0.2-2.6)	0 (0.0-1.8)	1 (1-3)	0.9 (0.8-1.7)
Family	0 (0.0-1.0)	0.5 (0.0-0.7)	0 (0.0-1.2)	0.4 (0.0-1.6)	0 (0-2)	0 (0.0-0.8)
Education	0.3 (0.3-0.3)	0 (0.0-0.3)	0.1 (0.1-0.2)	0 (0.0-0.2)	0 (0-0)	0 (0.0-0.0)

social-security contributions, reflecting their negative side-effects across the growth and equity dimensions. This finding is very robust to uncertainty about the effects of instruments and their ranking (Table 5).

The simulated long-term consolidation packages exhibit some differences from their short- to medium-term counterparts for two main reasons:

- Firstly, instruments resulting in cuts to public expenditure move up the ranking in the long term as their larger Keynesian demand effects are no longer taken into account. Cuts in other government consumption as a result play a much more important role in long- than in short-term simulated packages.
- Secondly, more room for manœuvre is assumed to be available in the area of pensions (over and above the effort implicit in the baseline) in a 2060 perspective than over the

medium term. Consequently, pensions are used more intensively to meet consolidation needs in the long than the short term.

These two factors result in a number of policy reversals, that is to say cases where a given country makes a larger use of a given instrument in the short to medium term than in the long term. Such policy reversals mainly concern taxes, and in particular property and corporate income taxes (Table 5), which generally fall from the upper to the lower half of instrument hierarchies as the time horizon expands.

As a result from this shift in the use of consolidation tools, the average share of spending reductions across national consolidation packages rises from 41% in the short to medium term to 65% in the long term. At both simulation horizons, the share of spending is particularly high among countries with modest consolidation needs, which to a large extent can be fulfilled with instruments like subsidies or pensions, which occupy top places in most rankings. In contrast, countries with substantial consolidation gaps often need to use large tax items as well, leading to a more balanced revenue-expenditure split or even to revenue-side adjustment becoming predominant.

Revenue and expenditure structures undergo some change from 2012 (the starting year) to 2020 (the medium term) and 2060 (the long term), reflecting both baseline developments (in the form of health spending pressures) and consolidation patterns. On the expenditure side, the share of health care in total spending rises at the expense of virtually all other instruments (Table 6). In the long run, this increase prevents overall primary spending from falling relative to GDP. In line with the different intensities of instrument use, pensions and other government consumption fall the most relative to GDP (mainly after 2020), while family and education spending remains broadly unchanged. Revenue items broadly follow a hump-shaped profile where they rise relative to GDP to meet short- to medium-term needs before falling back to on average 1 percentage point above their 2012 ratios to GDP in the very long term.

Budget structures also display some convergence across countries, as a consequence of broadly similar instrument orderings, but not much primarily because of the constraints imposed on instrument use (Table 7). On both revenue and expenditure sides, the cross-country dispersion of the size of budget items displays a general downward trend, which, unsurprisingly, tends to be more marked in the case of widely used instruments, such as subsidies, other government consumption, environmental levies and consumption taxes. Even in these cases, however, dispersion is only reduced by a third in the long term, and by much less in most other instruments.

5.4. *Robustness of the simulated consolidation packages*

5.4.1. *Sensitivity to uncertainty about the assessment of instrument impacts*

Numerous checks have been performed to test the robustness of the findings to uncertainty about the assessments of the side-effects of consolidation instruments. A large number (200) of alternative scenarios have been simulated: in each of these, one in every four assessments in Table 2, which is the equivalent of a full column, is chosen randomly and modified by adding a plus or minus sign. For each random draw, cluster-specific and long-term rankings corresponding to the new assessment of impacts are calculated. Then, for each random draw, full consolidation packages are simulated for the short to medium term as well as the long term. The values between brackets in Tables 5

to 6 and Tables A1 to A6 show the 10th and 90th percentiles of the values spanned by the results in the simulations associated with the random draws.

Table 6. **Evolution of expenditure and revenue structures**

Averages across countries

	2012		2020		2060	
	% of potential GDP	Share of spending or revenue (%)	% of potential GDP	Share of spending or revenue (%)	% of potential GDP	Share of spending or revenue (%)
Expenditure						
Public investment	2.6	6.2	2.5 (2.5-2.6)	6.3 (6.1-6.4)	2.5 (2.4-2.6)	6 (5.8-6.1)
Education	5.3	12.5	5.2 (5.2-5.2)	12.9 (12.9-13.0)	5.3 (5.2-5.3)	12.4 (12.4-12.5)
Health services provided in kind	6.5	15.5	6.4 (6.4-6.4)	15.7 (15.6-15.9)	9.5 (9.3-9.5)	22.4 (22.0-22.5)
Other in kind consumption (excluding family policy)	8.4	19.9	8.1 (7.7-8.2)	19.9 (19.2-20.2)	7.6 (7.5-7.8)	18 (17.7-18.4)
Pensions	8.1	19.1	7.9 (7.8-7.9)	19.4 (19.3-19.5)	7.2 (7.2-7.3)	17.1 (17.0-17.4)
Sickness and disability payments	2	4.7	1.9 (1.8-2.0)	4.7 (4.5-4.8)	1.8 (1.8-1.9)	4.3 (4.2-4.4)
Unemployment benefits	1.1	2.6	0.8 (0.8-1.0)	2 (2.0-2.4)	0.9 (0.8-1.0)	2 (1.8-2.2)
Family	2.4	5.6	2.4 (2.3-2.4)	5.8 (5.7-5.8)	2.3 (2.3-2.4)	5.5 (5.3-5.6)
Subsidies	1.2	2.8	0.8 (0.8-0.9)	2 (2.0-2.1)	0.8 (0.8-0.8)	1.8 (1.8-1.9)
Residual spending items	4.7	11.1	4.5	11.2	4.5	10.5
Total of spending instruments	42.2	100	40.5 (40.3-40.7)	100	42.3 (42.1-42.5)	100
Revenue						
Personal income taxes	8.6	21.3	9.3 (9.1-9.5)	21.7 (21.4-22.4)	8.8 (8.6-9.1)	21.2 (20.6-21.7)
Social security contributions	11.2	27.8	11.2 (11.2-11.4)	26.3 (26.2-26.7)	11.2 (11.2-11.3)	26.7 (26.7-27.0)
Corporate income taxes	2.9	7.1	3.1 (2.9-3.2)	7.2 (6.9-7.4)	2.9 (2.9-3.0)	7 (6.9-7.1)
Environmental taxes	2.3	5.8	2.6 (2.5-2.7)	6.2 (5.8-6.3)	2.6 (2.5-2.7)	6.2 (6.1-6.4)
Consumption taxes (other than environmental)	9	22.4	9.3 (9.2-9.6)	21.9 (21.8-22.5)	9.4 (9.1-9.7)	22.5 (21.9-23.1)
Recurring taxes on immovable property	1.3	3.1	1.5 (1.3-1.6)	3.6 (3.2-3.8)	1.4 (1.3-1.5)	3.3 (3.1-3.5)
Other property taxes	0.7	1.7	1 (0.9-1.0)	2.3 (2.1-2.4)	0.9 (0.8-1.0)	2.1 (2.0-2.3)
Sales of goods and services	2.8	7	3 (2.9-3.0)	7 (6.8-7.2)	3 (2.9-3.1)	7.2 (6.9-7.3)
Residual revenue items	1.5	3.8	1.6	3.8	1.6	3.8
Total of revenue instruments	40.2	100	42.6 (42.3-42.8)	100	41.7 (41.5-41.9)	100

Note: The table reports the size of the 17 fiscal instruments in p.p. of potential GDP, averaged across the 24 countries common to both simulation horizons, in 2012, 2020 and 2060. Figures for 2020 and 2060 reflect baseline developments in health spending as well as the consolidation packages implemented by each country in the short to medium term and in the long term, respectively. For simplicity, the year 2020 is taken as the medium-term consolidation horizon, though the latter varies somewhat across countries. Simulated interdecile intervals shown between brackets are computed using 200 random draws. Each individual assessment (i.e. each cell) in Table 2 is increased by one, decreased by one or retained with probabilities 1/8, 1/8 and 3/4, respectively. Residual spending and revenue items are not reported, as their share in the respective total remains broadly constant by simulation design (Box 4).

Table 7. **Cross-country convergence in expenditure and revenue structures**

	Standard deviation (percentage point of potential GDP)		
	2012	2020	2060
Expenditure			
Public investment	1	0.9	0.9
Education	1.1	1.1	1.1
Health services provided in kind	1.4	1.3	1.2
Other government in kind consumption	2.4	2.3	1.9
Pensions	3.3	3.1	2.8
Sickness and disability payments	0.6	0.6	0.5
Unemployment benefits	0.9	0.6	0.6
Family policy	1.1	1.1	1
Subsidies	0.8	0.6	0.5
Revenue			
Personal income taxes	3.3	3	3.1
Social security contributions	5.4	5.4	5.4
Corporate income taxes	0.9	0.9	0.9
Environmental taxes	0.7	0.5	0.5
Consumption taxes	2.4	2	1.9
Recurrent taxes on immovable property	1	0.9	1
Other property taxes	0.6	0.5	0.5
Sales of goods and services	1	0.9	0.9

Note: The table reports variability indicators computed across the 24 countries common to both simulation horizons, in 2012, 2020 and 2060. Figures for 2020 and 2060 reflect baseline developments in health spending as well as the consolidation packages implemented by each country in the short to medium term and in the long term, respectively. The year 2020 is taken as a proxy for the medium-term consolidation horizon although the latter varies across countries (Appendix B, Section 1).

The conclusions from these 200 randomly-generated simulations can be summarised as follows:

- The degree to which countries have to use poorly-ranked instruments, or can avoid doing so, is robust to uncertainty about impact assessments, especially in countries with large consolidation needs (Tables A1 to A6). In particular, in the alternative scenarios, there are very few shifts from being able to achieve most of the consolidation with well-ranked instruments to being forced to rely heavily on badly-ranked instruments, neither is there much movement in the opposite direction.
- The average use of each instrument is quite stable across alternative scenarios for both very well and very poorly ranked instruments (Table 5). There is more variation for middle-ranked instruments.
- The finding that short- as well as long-term simulated consolidation packages very seldom involve cuts in the areas of health, education and family policy holds very strongly in the robustness checks (Table 5).
- While the split between spending and tax adjustment shows sensitivity to uncertainty, especially at the country level (Tables A1 to A5), the findings that long-term packages rely more on spending reductions than tax increases and that short-term adjustment give a larger role to tax increases are very robust (Table A6).
- Policy reversals show some sensitivity to uncertainty (Table 5). The reason is that policy reversals occur mostly for instruments that feature in the middle of the generic ranking, which is the most unstable part of the ranking.

- The finding that the simulated packages largely respect the diversity of revealed preferences about tax and spending structures by prompting only very limited cross-country convergence is very robust in the random simulations (Table 6).

5.4.2. *Sensitivity to the chosen clustering technique*

A variant of the framework for simulating short- to medium-term consolidation packages has been prepared to check the sensitivity of the results to the weights put on objectives as a result of the clustering techniques. In this alternative framework, the clustering technique is replaced by splitting covered OECD countries into three equal-sized groups depending on the sign and size of the current account: strongly positive, strongly negative and close to balance. Instead of deriving the weighting scheme from clustering analysis, in the variant framework all objectives receive the same weight. Section 5 of Appendix B describes the methodology used to prepare this variant in greater detail and reports the results in Tables B4 to B8.

In summary, the results are relatively close to the main set of short- to medium-term simulations and corroborate its main findings although they take country circumstances less well into account. In addition, 200 random simulations have been conducted around this variant to assess its sensitivity to uncertainty about the assessment of the effects of instruments. The results are the same for the variant as for the main framework: most results are quite robust to uncertainty about instrument assessment (Tables B4 to B8). The only notable exception is the tax-spending split especially in countries with small consolidation needs.

5.4.3. *Sensitivity to the assumptions made on the available room for manœuvre for each instrument*

Variants of the short- to medium-term and long-term simulations have been conducted to explore the influence of the constraints on instrument use. The constraint that a given tax can be raised until the country reaches the 66th percentile of the covered OECD countries (and a spending item cut until the 33rd percentile) has been relaxed by allowing instrument use until the 80th and 20th percentiles for tax and spending items, respectively. The constraint that each instrument cannot be used for an amount greater than one standard deviation of the cross-country distribution has been retained.

In the short- to medium-term as well as the long-term simulations, relaxing the constraint on instrument use in this manner makes it possible for countries to make much more of their adjustment with well-ranked instruments (see Tables B9 to B14 for detailed results). The difference is particularly apparent for countries that make most of their adjustment through poorly-ranked instruments in the main simulations. In the long-term packages, as mentioned above, Australia, New Zealand and the United States find themselves in this situation. When the constraint on instrument use is relaxed as presented here, almost three-quarters of the adjustment is achieved with well-ranked instruments in Australia, and about half in New Zealand and the United States.

At the other extreme, another possibility would be to constrain the room for manœuvre at the median of OECD countries. In other words, for a tax instrument, the adjustment would be allowed only as long as a country does not raise more revenue with it, as a share of GDP, than half the OECD countries covered in the study. For a spending instrument, the limit on the room for manœuvre would be to spend no less in this area than half the covered OECD countries. The asterisks appearing in Appendix Tables B1 to B6

indicate all cases where the adjustment along one instrument crosses the median. The large number of asterisks in these tables illustrate that crossing the median is common in the simulations. Consequently, constraining the adjustment to stop at the median would result in greater use of poorly-ranked instruments.

6. The case for combining structural reforms and fiscal adjustment

For the purpose of the simulations reported above, no consideration was given to the scope for achieving efficiency gains when raising taxes or reducing spending. Structural reforms, while desirable in their own right, can also ease the trade-offs between consolidation, equity and long-term growth objectives. In the most favourable cases, structural reform can even eliminate trade-offs with growth and equity goals as well as bring fiscal improvements (OECD, 2013b).[17] As a result, structural reforms have the potential to make fiscal consolidation more likely to succeed (Alesina and Ardagna, 2012; Mauro, 2011).

At a general level, structural reforms that improve efficiency in the delivery of public services can reduce the adverse growth impact of spending cuts in productive areas of government spending (Journard et al., 2010; St Aubyn et al., 2009; Sutherland et al., 2007). Similarly, the negative equity impact of spending cuts can be headed off by structural reforms that ensure a better targeting of public services and transfers and stimulate labour supply. On the tax side, the adverse growth impact of hikes can be reduced through the closing of loopholes and base broadening (including by curbing fraud and evasion) rather than via rate increases (OECD, 2010a and 2010b). Hence, an important way of improving the trade-off between raising more revenue and preserving growth-friendly incentives is to cut back tax expenditures, though this recommendation should not be made across the board as some tax expenditures work to preserve productive potential, alleviate poverty or both. Such is the case of tax credits for low-income earners, which tackle poverty traps created by other parts of the tax and transfer system and help boost the employment of low-skill workers. Well-designed corporate income tax credits for research and development activities represent another example as they can provide remuneration for the growth-enhancing externalities from R&D (Jaumotte and Pain, 2005; Johansson et al., 2008, Westmore, 2013).

Even though substantial consolidation often requires deep changes in taxation and spending parameters, these changes do not in themselves amount to effective structural reform. International experience suggests that in fast-paced consolidation episodes the need for quick spending reductions or tax increases means that the changes in tax and spending parameters introduced for consolidation purposes often tend to be at odds with efficiency objectives. In particular, when the main objective is to achieve fiscal results quickly, it appears very difficult to introduce effective structural reforms in areas such as pension and health care where long lead times are required. Furthermore, changes in tax and spending items that are presented to the public as needed essentially for fiscal adjustment can meet strong, sometimes successful, pressure to reverse them once consolidation is achieved. These lessons from experience indicate that effective structural reform does not necessarily ensue from fiscal consolidation efforts. Policy strategies that explicitly combine fiscal and structural efforts in a common framework are thus desirable to take full advantage of their complementarities.

7. Concluding remarks

The present study proposed one structured way of thinking about the detailed instruments with which fiscal-consolidation plans can be implemented to reduce adverse side effects on other policy objectives. With a number of assumptions, it draws up illustrative quantitative simulations aimed at giving indications about the feasibility of such strategies, in particular to assess to what extent focussing on more favourable instruments requires deviations from existing tax and spending structures. These simulations are found to be broadly robust to uncertainty about the assessment of the effects of consolidation instruments on other policy objectives. Despite their limitations, the simulations indicate that prioritisation offers most countries distinct possibilities to adjust fiscal balances without much harm to other economic objectives. However, they also highlight that a few OECD countries find themselves in a more challenging situation where it may be appropriate to use better instruments more aggressively than assumed in the main simulations so as to avoid having to resort to more harmful fiscal variables.

While the simulated packages can provide suggestions about ways to pursue growth- and equity-friendly consolidation, they are not meant to substitute for country-by-country analysis of the best design of consolidation. The actual design of fiscal consolidation has to take into account the ways in which consolidation instruments interact with each other and with a country's particular characteristics. The study of interactions between instruments of consolidation appears to offer an interesting avenue for further research. Another potentially valuable extension of the present work would be to take into account feedbacks from the composition of consolidation on long-term growth potential and therefore public-debt dynamics.

Notes

1. This can be downloaded from: *www.oecd.org/eco/public-finance/Simulation-ranking-web.xlsx*.
2. This initial improvement at a fast pace, which generates a peak in the trajectory for the underlying primary balance, is needed in most but not all countries. Countries with a better starting fiscal position do not need such a peak. Nevertheless, the time path for the underlying primary balance always exhibits a kink (often, but not always, a peak), which provides the point where short- to medium-term consolidation needs are calculated.
3. Small deviations from baseline assumptions can accumulate to large differences in 2060 because: 1) demography, 2) the impact of pension reforms, 3) the impact of health care reforms and 4) long-term care cost projections can all change.
4. These assessments use end-2012 data or estimates for general government financial assets and liabilities on a national accounts basis as reported in the May 2013 *OECD Economic Outlook Database*.
5. Empirical evidence suggests that, when trading partners consolidate, growth is reduced, even after controlling for growth in trading-partner countries (Goujard, 2013). This finding that consolidation has cross-border effects over and above direct growth spillovers is consistent with the view that consolidation tends to rebalance production toward servicing external demand (IMF, 2010).
6. On the spending side, for instance, cuts in education spending achieved through reduced service provision can be described as pure budgetary measures whereas efficiency gains that can maintain a similar level of service for lower costs represents structural reform. On the revenue side, one example where the distinction is clear is indirect taxation where an increase in the standard VAT rate can be seen as a pure fiscal change while measures such as reducing the reliance on reduced rates and exemptions are part of structural tax reform. One example where the distinction is difficult to make is unemployment insurance where almost any form of reduction in benefits will amount to a change in structural policy settings.
7. Some categories of subsidies, however, can work in the direction of growth potential. In particular, government subsidies can encourage business research and development activities where the

social rate of return exceeds the private rate of return because of cross-company spillovers (Jaumotte and Pain, 2005).

8. Although part of the empirical literature finds a negative effect of public health spending on GDP per capita, this appears to be related to the output cost of the associated taxes which the present study considers separately (see for instance Box 6.1 in OECD, 2011c).

9. While for about two decades a reduction in the fertility rate works in the direction of lowering the dependency ratio and boosting output per capita, in the following five to six decades it contributes to an increase in the dependency ratio with negative implications for output per capita.

10. In theory, fiscal policy changes could, under specific conditions, be expected to have no major impact on aggregate demand if households anticipate the future tax benefits from debt reduction and adjust their saving in a way that just offset the rise in public savings. In practice, given that such a Ricardian equivalence effect depends on many conditions, including the access to complete financial markets, which typically fail to hold, a significant short-term impact from consolidation on aggregate spending is generally observed (Sutherland et al., 2010)

11. The assessment of the effect of instruments on income inequality draws largely on OECD (2012) and Rawdanowicz et al. (2013).

12. The study however incorporates no assessment of the impact of public investment on inequality. At a conceptual level, the effect is ambiguous. By providing the basis for public capital services that are consumed without relation to income, public investment should promote equality in effective consumption. On the other hand, inasmuch as public capital is complementary to private capital and boosts returns on capital, it could work in the direction of exacerbating income inequality because of the concentration of control over private capital. While there is evidence in favour of net equality-enhancing effects of public investment in developing countries, there are no comparable findings for OECD countries.

13. No positive or negative assessment is included for real estate taxes because of a lack of clear evidence. In most OECD countries, lower-income households pay a higher share of their income in recurring property taxes than higher income taxes, so that on this count recurring property taxes might be described as regressive. However, this situation largely reflects larger home ownership among retirees, implying that recurring property taxation is not necessarily regressive in a dynamic perspective, and may even be progressive if adjusting income fully for the market value of owner-occupied housing services.

14. Besides affecting the choice of fiscal instruments, short-run growth impacts also have important implications for the optimal pace and timing of consolidation, an issue from which this paper largely abstracts and for which the reader is referred for instance to Rawdanowicz (2012).

15. As is well known, Norway has a huge current account surplus (17.4% of 2012 GDP in cyclically-adjusted terms). However, unlike other surplus countries, this is largely due to the exploitation of finite natural resources (fossil fuels). As the Norwegian external position reflects exceptional circumstances, it has not been taken into account when forming clusters, and Norway has been treated as if it had a balanced current account.

16. The top nine instruments have either *i*) beneficial impacts on both long-term growth and long-term equity (Table 2), as is the case of subsidy reduction, *ii*) impacts which are beneficial on one objective and fairly neutral on the other, as it happens with reductions in other government consumption, or *iii*) opposite impacts on long-term growth and long-term equity which can somehow be regarded as compensating each other, reductions in sickness and disability payments being an example. Implicit in this "compensation" argument is the use of +1 and –1 scores for each + and – sign in Table 2, which is admittedly a simplifying assumption, rather than an attempt to calibrate a social-welfare function.

17. The baseline scenario against which consolidation needs are estimated already incorporates significant structural reform (see Johansson et al., 2013). Correspondingly, far-reaching structural reforms going beyond the assumptions in Johansson et al. (2013) would be required to reduce consolidation needs compared with the estimates presented previously.

References

Afonso, A. and J.T. Jalles (2013), "Economic Performance, Government Size and Institutional Quality", in *Fiscal Policy and Growth*, Banca d'Italia, Rome.

Alesina, A. and S. Ardagna (2010), "Large Changes in Fiscal Policy: Taxes *vs*. Spending", in *Tax Policy and the Economy,* Vol. 24, The University of Chicago Press.

Alesina, R. and S. Ardagna (2012), "The Design of Fiscal Adjustments", *NBER Working Paper Series*, No. 18423.

Auerbach, A.J. and Y. Gorodnichenko (2012), "Measuring the Output Responses to Fiscal Policy", *American Economic Journal: Economic Policy* 2012, 4(2), pp. 1-27.

Barbiero, O. and B. Cournède (2013), "New Econometric Estimates of Long-Term Growth Effects of Different Areas of Public Spending", *OECD Economics Department Working Papers,* No. 1100.

Barrell, R., D. Holland and I. Hurst (2012), "Fiscal Consolidation Part 2: Fiscal Multipliers and Fiscal Consolidation", *OECD Economics Department Working Papers*, No. 933.

Baum, A., M. Poplawski-Ribeiro and A. Weber (2012), "Fiscal Multipliers and the State of the Economy", *IMF Working Paper*, No. 12/286, IMF, Washington, DC.

Blöchliger, H. (2012), "Fiscal Consolidation across Government Levels", paper presented to the joint IMF-OECD Fiscal Workshop on Instruments of Fiscal Consolidation, Paris, 11-12 November, available at *www.oecd.org/eco/public-finance/jointimf-oecdfiscalworkshop.htm*.

Blöchliger, H., D.H. Song and D. Sutherland (2012), "Fiscal Consolidation Part 4: Case Studies of Large Fiscal Consolidation Episodes," *OECD Economics Department Working Papers*, No. 935.

Blöchliger, H. (2013), "Decentralisation and Growth Part 1: Decentralisation, Economic Activity and Educational Performance", *OECD Economics Department Working Papers*, forthcoming.

Blöchliger, H. and B. Égert (2013), "Decentralisation and Growth Part 2: Decentralisation and its Impact on Economic Activity, Productivity and Investment", *OECD Economics Department Working Papers*, forthcoming.

Bouis, R., B. Cournède and A.K. Christensen (2012), "Implications of Output Gap Uncertainty in Times of Crisis", *OECD Economics Department Working Papers,* No. 977.

Bouis, R., A.K. Christensen and B. Cournède (2013), "Deleveraging: Challenges, Progress and Policies", *OECD Economics Department Working Papers*, No. 1077.

Bouis, R., R. Duval and F. Murtin (2011), "The Policy and Institutional Drivers of Economic Growth Across OECD and Non-OECD Economies", *OECD Economic Department Working Papers*, No. 843.

Bouis, R., O. Causa, L. Demmou, R. Duval and A. Zdzienicka (2012), "The Short-Term Effects of Structural Reforms: An Empirical Analysis", *OECD Economics Department Working Papers*, No. 949.

Causa, O. and C. Chapuis (2009), "Equity in Student Achievement Across OECD Countries: An Investigation of the Role of Policies", *OECD Economics Department Working Papers*, No. 708.

Causa, O. and Å. Johansson (2009), "Intergenerational Social Mobility", *OECD Economics Department Working Papers*, No. 707.

Christiano, L., M. Eichenbaum, and S. Rebelo (2011), "When is the Government Spending Multiplier Large?", *Journal of Political Economy*, Vol. 119, No. 1 (February), pp. 78-121.

Corsetti, G., A. Meier and G.J. Müller (2012), "What Determines Government Spending Multipliers?", *Economic Policy* (October), pp. 521-565.

Cournède, B., A. Goujard and Á. Pina (2013), "How to Achieve Growth- and Equity-Friendly Fiscal Consolidation? A Proposed Methodology for Instrument Choice with an Illustrative Application to OECD Countries", *OECD Economics Department Working Papers*, No. 1088.

Cournède, B. and F. Gonand (2006), "Restoring Fiscal Sustainability in the Euro Area: Raise Taxes or Curb Spending?", *OECD Economics Department Working Papers*, No. 520.

Crain, W.M. and J.B. O'Roark (2004), "The Impact of Performance-Based Budgeting on State Fiscal Performance", *Economics of Governance*, Vol. 5.

Curristine, T., Z. Lonti and I. Joumard (2007), "Improving Public Sector Efficiency: Challenges and Opportunities," *OECD Journal on Budgeting*, Vol. 7, No. 1, OECD Publishing.

de la Maisonneuve, C. and J. Oliveira Martins (2013), "Public Spending on Health and Long-Term Care: A New Set of Projections", *OECD Economics Department Working Papers*, forthcoming.

DeLong, J.B. and L.H. Summers (2012), "Fiscal Policy in a Depressed Economy", *Brookings Papers on Economic Activity*, Spring 2012.

de Serres, A., F. Murtin and G. Nicoletti (2010), "A Framework for Assessing Green Growth Policies", *OECD Economics Department Working Papers*, No. 774.

Devries, P., J. Guajardo, D. Leigh and A. Pescatori (2011), "A New Action-Based Dataset on Fiscal Consolidation", *IMF Working Papers*, No. 11/128, Washington, DC.

European Commission (2012), "Fiscal Sustainability Report 2012", *European Economy*, Vol. 8/2012.

Goujard, A. (2013), "Cross-Country Spillovers from Fiscal Consolidation", *OECD Economics Department Working Papers*, No. 1099.

Guichard, S. and E. Rusticelli (2010), "Assessing the Impact of the Financial Crisis on Structural Unemployment in OECD Countries", *OECD Economics Department Working Papers*, No. 767.

Hagemann, R. (2012), "Fiscal Consolidation: Part 6. What Are the Best Policy Instruments for Fiscal Consolidation?", *OECD Economics Department Working Papers*, No. 937.

Hoeller, P. et al. (2012), "Less Income Inequality and More Growth – Are They Compatible? Part 1. Mapping Income Inequality across the OECD", *OECD Economics Department Working Papers*, No. 924.

IMF (2010), "Will it Hurt? Macroeconomic Effects of Fiscal Consolidation", *IMF World Economic Outlook*, Chapter 3, October, Washington, DC.

Jaumotte, F. and N. Pain (2005), "Innovation in the Business Sector", *OECD Economics Department Working Papers*, No. 459.

Johansson, Å., C. Heady, J. Arnold, B. Brys and L. Vartia (2008), "Taxation and Economic Growth", *OECD Economics Department Working Papers*, No. 620.

Johansson, Å., D. Turner, Y. Guillemette, F. Murtin, C. de la Maisonneuve, P. Bagnoli, G. Bousquet and F. Spinelli (2013), "Long-Term Growth Scenarios", *OECD Economics Department Working Papers*, No. 1000.

Joumard, I., P. Hoeller, C. André and C. Nicq (2010), *Health Care Systems: Efficiency and Policy Settings*, OECD Publishing.

Kerdrain, C., I. Koske and I. Wanner (2010), "The Impact of Structural Policies on Saving, Investment and Current Accounts", *OECD Economics Department Working Papers*, No. 815.

Kopczuk, W. (2007), "Bequest and Tax Planning: Evidence from Estate Tax Returns", *Quarterly Journal of Economics*, MIT Press, Vol. 122.

Larch, M. and A. Turini (2008), "Received Wisdom and Beyond: Lessons from Fiscal Consolidations in the EU", *European Economy, Economic Papers*, No. 320.

Mauro, P. (ed.) (2011), *Chipping Away at Public Debt*, Wiley.

Molnar, M. (2012), "Fiscal Consolidation Part 5: What Factors Determine the Success of Consolidation Efforts?", *OECD Economics Department Working Papers*, No. 936.

OECD (2009), "The Effectiveness and Scope of Fiscal Stimulus", *OECD Economic Outlook: Interim Report*, March, OECD Publishing.

OECD (2003), *The Sources of Growth in OECD Countries,* OECD Publishing.

OECD (2007a), *Babies and Bosses: Reconciling Work and Family Life*, OECD Publishing.

OECD (2007b), *Performance Budgeting in OECD Countries*, OECD Publishing.

OECD (2009a), *Interim OECD Economic Outlook*, OECD Publishing.

OECD (2009b), *The Political Economy of Reform: Lessons from Pensions, Product Markets and Labour Markets in Ten OECD Countries*, OECD Publishing.

OECD (2009c), *Government at a Glance 2009*, OECD Publishing.

OECD (2010a), *Choosing a Broad Base-Low Rate Approach to Taxation,* OECD Publishing.

OECD (2010b), *OECD Economic Outlook,* No. 88, November, OECD Publishing.

OECD (2010c), *Tax Expenditures in OECD Countries,* OECD Publishing.

OECD (2011a), *Pensions At a Glance: Retirement Income Systems in OECD and G20 Countries*, OECD Publishing.

OECD (2011b), *Doing Better for Families*, OECD Publishing.

OECD (2011c), *Going for Growth: Economic Policy Reforms,* OECD Publishing.

OECD (2011d), *Divided We Stand: Why Inequality Keeps Rising,* OECD Publishing.

OECD (2013a), *OECD Economic Outlook*, No. 93, May, OECD Publishing.

OECD (2013b), *Economic Policy Reforms 2013: Going for Growth*, OECD, Publishing.

Ollivaud, P. and C. Schwellnus (2013), "International Imbalances: Recent Developments and Medium-Term Outlook", *OECD Economics Department Working Papers*, forthcoming.

Padoan, P.C., U. Sila and P. van den Noord (2012), "Avoiding Debt Traps: Fiscal Consolidation, Financial Backstops and Structural Reforms", *OECD Journal: Economic Studies*, OECD Publishing.

Rawdanowicz, L., E. Wurzel and P. Ollivaud (2011), "Current Issues in Managing Government Debt and Assets", *OECD Economics Department Working Papers,* No. 923.

Rawdanowicz, L. (2012), "Choosing the Pace of Fiscal Consolidation", *OECD Economics Department Working Papers*, No. 992.

Rawdanowicz, L., E. Wurzel and A.K. Christensen (2013), "The Equity Implications of Fiscal Consolidation", *OECD Economics Department Working Papers,* No. 1013.

Röhn, O. (2010), "New Evidence on the Private Saving Offset and Ricardian Equivalence", *OECD Economics Department Working Papers*, No. 762.

Schwellnus, C. and J.M. Arnold (2008), "Do Corporate Taxes Reduce Productivity and Investment at the Firm Level? Cross-Country Evidence From the Amadeus Dataset", *OECD Economics Department Working Papers*, No. 641.

Serret, Y. and N. Johnstone (2006) (ed.), *The Distributional Effects of Environmental Policy*, OECD Publishing.

St. Aubyn, M., Á. Pina, F. Garcia and J. Pais (2009), "Study on the Efficiency and Effectiveness of Public Spending on Tertiary Education", *Economic Papers*, No. 390, European Commission, Directorate-General for Economic and Financial Affairs.

Sutherland, D. and R. Price (2007), "Linkages Between Performance and Institutions in the Primary and Secondary Education Sector", *OECD Economics Department Working Papers*, No. 558.

Sutherland, D., P. Hoeller, B. Égert and O. Roehn (2010), "Counter-Cyclical Economic Policy", *OECD Economics Department Working Papers*, No. 760.

Sutherland, D., P. Hoeller and R. Merola (2012), "Fiscal Consolidation: Part 1. How Much is Needed to Reduce Debt to a Prudent Level?", *OECD Economics Department Working Papers,* No. 932.

Turner, D. and F. Spinelli (2012) "Explaining the Interest-Rate-Growth Differential Underlying Government Debt Dynamics", *OECD Economics Department Working Papers*, No. 919.

Vartia, L. (2008), "How do Taxes Affect Investment and Productivity? An Industry-Level Analysis of OECD Countries", *OECD Economics Department Working Papers* No. 656.

Westmore, B. (2013), "R&D, Patenting and Productivity: The Role of Public Policy", *OECD Economics Department Working Paper* (forthcoming).

Whiteford, P. and W. Adema (2007), "What Works Best in Reducing Child Poverty: A Benefit or Work Strategy?", *OECD Social, Employment and Migration Working Papers*, No. 51.

APPENDIX A

Detailed composition of consolidation packages

The present section provides detailed quantitative information about the illustrative consolidation packages presented in Section 5 of the main text. Tables A1 to A5 provide results about the illustrative short- to medium-term consolidation packages of countries with one table per cluster. Table A6 details the illustrative long-term consolidation packages for all countries covered. The instruments used are as described in Section 2 of Appendix B. The categories "used spending residual" and "used revenue residual" refer to the part of the adjustment that is achieved through residual items of primary expenditure and receipts which are not considered as instruments of consolidation as they have no direct economic interpretation. However, there is no reason to assume that they remain constant as a share of potential GDP when other budgetary items adjust, so the assumption is made that they remain fixed as shares of total primary spending or revenues (whichever is relevant).

Table A1. **Instrument use and achieved short- to medium-term consolidation vs. needs in cluster No. 1**

Percentages of potential GDP except if otherwise mentioned

Description	AUS	CAN	GBR	ISR	ITA	JPN	NZL	POL	PRT
Subsidies	0.6*	0.3	0	0	0.3	0	0	0	0
	(0.6-0.6)	(0.3-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0	0.5*	0.2	0.2	0	0.4*	0.7*	0.7*	0.4*
	(0.0-0.0)	(0.5-0.5)	(0.2-0.2)	(0.2-0.2)	(0.0-0.0)	(0.4-0.4)	(0.7-0.7)	(0.7-0.7)	(0.3-0.3)
Pensions	0	0	0.3	0.2	0.3	0.8	0	0.9	0.9
	(0.0-0.0)	(0.0-0.0)	(0.3-0.3)	(0.0-0.2)	(0.0-0.7)	(0.8-0.8)	(0.0-0.0)	(0.0-0.9)	(0.9-0.9)
Corporate income taxes	0	0.2	0.4*	0	0	0	0	0.9*	0.2
	(0.0-0.0)	(0.0-0.2)	(0.4-0.4)	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.9)	(0.0-0.2)
Personal income taxes	0	0	0.5	1.3	0	4.6*	0	1.3	4.1*
	(0.0-0.0)	(0.0-0.0)	(0.5-0.5)	(0.0-1.5)	(0.0-0.0)	(4.6-4.6)	(0.0-0.0)	(0.0-2.7)	(3.9-4.1)
Recurrent taxes on immovable property	0	0	0	0	0	0	0	0	0.8*
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.0-0.8)
Unemployment benefits	0	0.2*	0	0	0	0.3	0	0	0.3
	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.3-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)
Environmental taxes	0.7*	0.7	0.2*	0	0	0.7	0.7	0	0.2
	(0.0-0.7)	(0.0-0.7)	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)	(0.7-0.7)	(0.0-0.7)	(0.0-0.7)	(0.0-0.2)
Other government in kind consumption	0.6	0.6	2.2*	0	0	0	0	0	0.2
	(0.0-1.0)	(0.0-1.8)	(2.2-2.2)	(0.0-1.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-1.1)	(0.0-0.6)
Sales of goods and services	0	0	0.7*	0	0	1.0*	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.7-0.7)	(0.0-0.0)	(0.0-0.0)	(1.0-1.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)
Consumption taxes (other than environmental)	0	0	1.4*	0	0	2.5	0	0	0
	(0.0-0.6)	(0.0-0.8)	(0.0-1.4)	(0.0-0.0)	(0.0-0.0)	(2.5-2.5)	(0.0-0.0)	(0.0-0.0)	(0.0-0.6)
Public investment	0	0	0.3	0	0	1.1	0.2	0	0
	(0.0-0.6)	(0.0-0.6)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)	(1.1-1.1)	(0.0-0.9)	(0.0-0.0)	(0.0-0.0)
Sickness and disability payments	0	0	0.7*	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.4)	(0.0-0.0)	(0.0-0.2)
Health services provided in kind	0	0	1.5*	0	0	1.5*	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-1.5)	(0.0-0.0)	(0.0-0.0)	(1.5-1.5)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0	0.9*	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-3.0)	(0.0-0.0)	(0.0-0.0)	(0.9-0.9)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.4)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0	0.3	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)	(0.3-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0.1	0.1	0.9	0	0	0.6	0	0.1	0.1
Used revenue residual	0	0	0.2	0	0	0.5	0.1	0.1	0.4
Share spending efforts	65	45	62	14	100	30	9	25	20
	(35-97)	(23-81)	(35-64)	(0-89)	(53-100)	(30-31)	(9-56)	(0-53)	(17-25)
Achieved consolidation	1.9	2.7	9.2	1.7	0.7	15.3	1.7	4	7.5
Consolidation needs	1.9	2.7	9.2	1.7	0.7	18.3	1.7	4	7.5
Share top 9 instruments	100	100	45	100	100	48	90	100	100
	(70-100)	(78-100)	(27-61)	(100-100)	(100-100)	(47-66)	(88-100)	(100-100)	(96-100)
Instruments crossing the median	2	2	7	0	0	5	1	2	3
	(1-2)	(1-3)	(5-7)	(0-1)	(0-1)	(5-5)	(1-2)	(1-2)	(3-4)

Note: Simulated interdecile intervals in parentheses are computed using 200 random draws. Each individual assessment (i.e. each cell) in Table 2 is increased by one, decreased by one or retained with probabilities 1/8, 1/8 and 3/4, respectively.

* Denotes that the proposed adjustment takes the country from above to below the median for spending reductions or from below to above the median for tax increases. Simulated interdecile intervals are not reported for these items.

Table A2. **Instrument use and achieved short- to medium-term consolidation vs. needs in cluster No. 2**

Percentages of potential GDP except if otherwise mentioned

Description	United States
Subsidies	0 (0.0-0.0)
Pensions	0.5 (0.5-0.5)
Other property taxes	0.7* (0.7-0.7)
Unemployment benefits	0.2 (0.2-0.2)
Corporate income taxes	0.2 (0.2-0.2)
Environmental taxes	0.7 (0.7-0.7)
Recurrent taxes on immovable property	0 (0.0-0.0)
Personal income taxes	1 (1.0-1.0)
Other government in kind consumption (excluding family policy)	0 (0.0-0.0)
Sales of goods and services	0 (0.0-0.0)
Sickness and disability payments	0 (0.0-0.0)
Consumption taxes (other than environmental)	2.5 (2.5-2.5)
Public investment	0.3 (0.0-0.3)
Health services provided in kind	1.3* (1.3-1.5)
Family	0 (0.0-0.0)
Social security contributions	0 (0.0-0.6)
Education	0 (0.0-0.0)
Used spending residual	0.2
Used revenue residual	0.2
Share spending efforts	31 (31-37)
Achieved consolidation	7.7
Consolidation needs	7.7
Share top 9 instruments	45 (31-74)
Instruments crossing the median	2 (2-2)

Note: See notes to Table A1.

Table A3. **Instrument use and achieved short- to medium-term consolidation vs. needs in cluster No. 3**

Percentages of potential GDP except if otherwise mentioned

Description	ESP	GRC	IRL
Pensions	0	0.6	0
	(0.0-0.0)	(0.6-0.6)	(0.0-0.0)
Subsidies	0.2	0	0
	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Unemployment benefits	1.6	0	1.1
	(0.2-1.6)	(0.0-0.0)	(1.1-1.1)
Environmental taxes	0.7	0.2	0.3*
	(0.0-0.7)	(0.0-0.2)	(0.3-0.3)
Recurrent taxes on immovable property	0.4	1.0*	1.0*
	(0.0-0.4)	(0.7-1.0)	(0.5-1.0)
Corporate income taxes	0.9*	0.2	0.8*
	(0.0-0.9)	(0.0-0.2)	(0.0-0.8)
Sales of goods and services	1	0.3	0.3*
	(0.0-1.0)	(0.0-0.3)	(0.0-0.3)
Personal income taxes	0.3	3.4*	0.3
	(0.0-2.5)	(2.3-3.4)	(0.0-0.3)
Sickness and disability payments	0	0	0.3*
	(0.0-0.7)	(0.0-0.0)	(0.0-0.3)
Other government in kind consumption (excluding family policy)	0	1.9*	0
	(0.0-0.6)	(1.1-2.3)	(0.0-0.0)
Consumption taxes (other than environmental)	0	0	1.4
	(0.0-2.5)	(0.0-0.7)	(1.4-2.5)
Public investment	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0.1	0.1	0.2
Used revenue residual	0.1	0.4	0.2
Share spending efforts	36	33	27
	(18-58)	(23-41)	(22-27)
Achieved consolidation	5.3	8.2	5.8
Consolidation needs	5.3	8.2	5.8
Share top 9 instruments	100	75	69
	(100-100)	(36-100)	(52-100)
Instruments crossing the median	1	3	5
	(1-2)	(2-3)	(4-5)

Note: See notes to Table A1.

Table A4. **Instrument use and achieved short- to medium-term consolidation vs. needs in cluster No. 4**

Percentages of potential GDP except if otherwise mentioned

Description	AUT	BEL	CZE	FIN	FRA	HUN	ISL	SVK	SVN
Pensions	0.2	0	0.1	0	0.6	0.2	0	0	0
	(0.0-0.2)	(0.0-0.0)	(0.1-0.1)	(0.0-0.0)	(0.6-0.6)	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Subsidies	0	0.8	0.8	0.7*	0.7*	0.7*	0.8*	0.5*	0.5*
	(0.0-0.2)	(0.8-0.8)	(0.7-0.8)	(0.7-0.7)	(0.7-0.7)	(0.7-0.7)	(0.8-0.8)	(0.5-0.5)	(0.5-0.5)
Other property taxes	0	0	0.6*	0.4*	0	0.2	0.3	0.7*	0.7*
	(0.0-0.0)	(0.0-0.0)	(0.0-0.6)	(0.4-0.4)	(0.0-0.0)	(0.0-0.2)	(0.3-0.3)	(0.6-0.7)	(0.2-0.7)
Unemployment benefits	0	0.7	0	0.9	0.8*	0.2	0.6	0	0
	(0.0-0.0)	(0.0-0.7)	(0.0-0.5)	(0.0-0.9)	(0.6-0.8)	(0.0-0.2)	(0.0-0.6)	(0.0-0.0)	(0.0-0.0)
Environmental taxes	0	0	0	0	0.7*	0	0.5*	0.7	0
	(0.0-0.0)	(0.0-0.6)	(0.0-0.0)	(0.0-0.0)	(0.4-0.7)	(0.0-0.0)	(0.0-0.5)	(0.0-0.7)	(0.0-0.0)
Recurrent taxes on immovable property	0	0	0	0.8*	0	0.4*	0	1.0*	0.9*
	(0.0-0.0)	(0.0-0.3)	(0.0-0.6)	(0.0-0.8)	(0.0-0.0)	(0.0-0.9)	(0.0-0.0)	(0.0-1.0)	(0.0-0.9)
Sales of goods and services	0	0	0	0	0	0	0.3	1	0
	(0.0-0.0)	(0.0-0.7)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.0-0.3)	(0.0-1.0)	(0.0-0.0)
Sickness and disability payments	0	0	0	0.7	0	0	0.2	0	0
	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.0-0.4)	(0.0-0.2)	(0.0-0.0)	(0.0-0.2)
Personal income taxes	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.4)	(0.0-0.0)	(0.0-2.0)	(0.0-0.9)
Corporate income taxes	0	0	0	0.1	0.5*	0	0.5*	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.9)	(0.0-0.5)	(0.0-0.4)	(0.0-0.5)	(0.0-0.4)	(0.0-0.9)
Other government in kind consumption (excluding family policy)	0	0	0	0	1.2*	0	0.2	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-1.7)	(0.0-1.8)	(0.0-0.0)	(0.0-1.6)	(0.0-0.0)	(0.0-0.3)
Consumption taxes (other than environmental)	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-2.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-1.1)	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0	0.1	0.1	0.2	0.2	0.1	0.2	0.1	0.1
Used revenue residual	0	0	0	0	0	0	0	0.1	0.1
Share spending efforts	100	100	64	65	73	67	56	13	23
	(100-100)	(56-100)	(60-100)	(45-88)	(73-89)	(51-90)	(50-83)	(13-13)	(23-55)
Achieved consolidation	0.2	1.6	1.6	3.8	4.7	1.8	3.6	4	2.2
Consolidation needs	0.2	1.6	1.6	3.8	4.7	1.8	3.6	4	2.2
Share top 9 instruments	100	100	100	97	63	100	79	100	100
	(100-100)	(100-100)	(100-100)	(79-100)	(63-100)	(100-100)	(78-100)	(100-100)	(100-100)
Instruments crossing the median	0	0	1	3	5	2	3	3	3
	(0-0)	(0-1)	(1-2)	(2-4)	(3-5)	(1-2)	(2-3)	(2-4)	(2-3)

Note: See notes to Table A1.

Table A5. **Instrument use and achieved short- to medium-term consolidation vs. needs in cluster No. 5**

Percentages of potential GDP except if otherwise mentioned

Description	LUX	NLD	SWE
Subsidies	0.2	0.6*	0.7*
	(0.2-0.2)	(0.6-0.6)	(0.3-0.8)
Other property taxes	0	0.2	0.2
	(0.0-0.0)	(0.0-0.2)	(0.0-0.7)
Pensions	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)
Environmental taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Recurrent taxes on immovable property	0	0.8*	0
	(0.0-0.0)	(0.0-0.8)	(0.0-0.0)
Other government in kind consumption (excluding family policy)	0	1.1	0
	(0.0-0.0)	(0.0-1.9)	(0.0-0.0)
Sales of goods and services	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Personal income taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Unemployment benefits	0	0	0
	(0.0-0.0)	(0.0-0.4)	(0.0-0.0)
Social security contributions	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Corporate income taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Sickness and disability payments	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Consumption taxes (other than environmental)	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0	0.1	0.1
Used revenue residual	0	0	0
Share spending efforts	100	65	85
	(100-100)	(65-92)	(30-100)
Achieved consolidation	0.2	2.8	1
Consolidation needs	0.2	2.8	1
Share top 9 instruments	100	100	100
	(100-100)	(100-100)	(100-100)
Instruments crossing the median	0	2	1
	(0-0)	(1-2)	(1-2)

Note: See note to Table A1.

Table A6. **Instrument use and achieved long-term consolidation vs. needs**

Percentages of potential GDP except if otherwise mentioned

Description	JPN	GBR	USA	SVK	AUS	POL	ESP	NZL
Subsidies	0	0	0	0.5*	0.6*	0	0.2	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.5-0.5)	(0.6-0.6)	(0.0-0.0)	(0.2-0.2)	(0.0-0.0)
Pensions	3.2	1	1.9	0	0	3.7	0	3.2
	(3.2-3.2)	(1.0-1.0)	(1.9-1.9)	(0.0-0.0)	(0.0-0.0)	(3.7-3.7)	(0.0-0.0)	(3.2-3.2)
Other in kind consumption (excluding family policy)	0	2.2*	0	0	1	1.1	0.6	0
	(0.0-0.0)	(2.2-2.2)	(0.0-0.0)	(0.0-0.0)	(1.0-1.0)	(0.6-1.1)	(0.6-0.6)	(0.0-0.0)
Unemployment benefits	0.3	0	0.2	0	0	0	1.4	0.3
	(0.3-0.3)	(0.0-0.0)	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(1.4-1.4)	(0.3-0.3)
Environmental taxes	0.7	0.2*	0.7	0.7	0.7*	0.6	0.7	0.7
	(0.7-0.7)	(0.2-0.2)	(0.7-0.7)	(0.7-0.7)	(0.7-0.7)	(0.0-0.7)	(0.7-0.7)	(0.7-0.7)
Other property taxes	0.4*	0.2	0.7*	0.7*	0	0.6*	0	0.4*
	(0.4-0.4)	(0.2-0.2)	(0.7-0.7)	(0.7-0.7)	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.4-0.4)
Sickness and disability payments	0	0.7*	0	0	0	0	0.7*	0
	(0.0-0.0)	(0.7-0.7)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.4)	(0.4-0.7)	(0.0-0.0)
Recurrent taxes on immovable property	0	0	0	1.0*	0	0	0.4	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.4-1.0)	(0.0-0.0)	(0.0-0.1)	(0.0-0.4)	(0.0-0.0)
Sales of goods and services	1.0*	0.7*	0	1	0	0	1	1.0*
	(0.6-1.0)	(0.7-0.7)	(0.0-0.0)	(0.5-1.0)	(0.0-0.0)	(0.0-0.5)	(0.2-1.0)	(0.6-1.0)
Personal income taxes	2.8	0.5	1	1.8	0	0	0.4	2.8
	(0.0-4.4)	(0.0-0.5)	(0.0-1.0)	(0.0-3.3)	(0.0-0.0)	(0.0-0.0)	(0.0-1.5)	(0.0-4.4)
Corporate income taxes	0	0.4*	0.2	0.2*	0	0	0.1	0
	(0.0-0.0)	(0.0-0.4)	(0.0-0.2)	(0.0-0.4)	(0.0-0.0)	(0.0-0.0)	(0.0-0.9)	(0.0-0.0)
Consumption taxes (other than environmental)	1.5	1.4*	2.5	0.9	2.5	0	0.4	1.5
	(0.0-2.5)	(0.0-1.4)	(1.1-2.5)	(0.0-2.3)	(2.0-2.5)	(0.0-0.0)	(0.0-1.6)	(0.0-2.5)
Public investment	0	0.1	0.1	0	0.8	0	0	0
	(0.0-1.1)	(0.0-0.3)	(0.0-0.3)	(0.0-0.3)	(0.0-1.1)	(0.0-0.0)	(0.0-0.0)	(0.0-1.1)
Health services provided in kind	0	0.2	0.3	0	0.4	0	0	0
	(0.0-1.3)	(0.0-1.2)	(0.0-1.3)	(0.0-0.3)	(0.0-0.5)	(0.0-0.0)	(0.0-0.9)	(0.0-1.3)
Family	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.8)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	(0.4-0.7)	(0.8-1.1)	(0.5-0.8)	(0.1-0.2)	(0.2-0.3)	(0.2-0.3)	(0.1-0.1)	(0.4-0.7)
Used revenue residual	(0.1-0.3)	(0.1-0.2)	(0.1-0.3)	(0.2-0.4)	(0.0-0.0)	(0.0-0.1)	(0.0-0.1)	(0.1-0.3)
Share spending efforts	37	59	36	8	49	80	49	37
	(37-54)	(59-76)	(35-54)	(7-13)	(49-56)	(76-88)	(40-64)	(37-54)
Achieved consolidation	10.5	8.5	8.3	7	6.3	6.3	6.1	10.5
Consolidation needs	10.5	8.5	8.3	7	6.3	6.3	6.1	10.5
Share top 9 instruments	57	66	46	58	39	100	85	57
	(48-94)	(63-82)	(46-79)	(49-94)	(39-80)	(100-100)	(82-100)	(48-94)
Instruments crossing the median	2	6	1	4	2	1	1	2
	(1-3)	(6-7)	(1-2)	(3-5)	(2-3)	(0-2)	(1-2)	(1-3)

Note: See note to Table A1.

Table A6. **Instrument use and achieved long-term consolidation vs. needs** *(cont.)*

Percentage of potential GDP except if otherwise mentioned

Description	ISR	FRA	NLD	IRL	CAN	SVN	CZE	PRT
Subsidies	0	0.7*	0.6*	0	0.3	0.5*	0.8	0
	(0.0-0.0)	(0.7-0.7)	(0.6-0.6)	(0.0-0.0)	(0.3-0.3)	(0.5-0.5)	(0.8-0.8)	(0.0-0.0)
Pensions	0.8	2.2	0	0	0	0	0.4	3.4
	(0.8-0.8)	(1.8-2.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.4-0.4)	(2.8-3.4)
Other in kind consumption (excluding family policy)	2.3	1.1*	2.3	0	2.3*	2.0*	1.4*	0
	(2.3-2.3)	(0.0-1.7)	(2.3-2.3)	(0.0-0.0)	(1.8-2.3)	(2.0-2.0)	(0.9-1.4)	(0.0-0.5)
Unemployment benefits	0	0.4	0.5	0.9	0.2*	0	0.4*	0
	(0.0-0.0)	(0.0-0.7)	(0.5-0.5)	(0.9-0.9)	(0.0-0.2)	(0.0-0.0)	(0.1-0.4)	(0.0-0.3)
Environmental taxes	0	0	0	0.3*	0.3	0	0	0
	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.3-0.3)	(0.0-0.7)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0.2	0	0.2	0	0.2*	0.7*	0.3*	0
	(0.2-0.2)	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.5)	(0.2-0.7)	(0.0-0.6)	(0.0-0.0)
Sickness and disability payments	0.5*	0	0.6	0.3*	0	0.2*	0	0
	(0.1-0.5)	(0.0-0.0)	(0.0-0.7)	(0.3-0.3)	(0.0-0.0)	(0.0-0.2)	(0.0-0.3)	(0.0-0.0)
Recurrent taxes on immovable property	0	0	0	1.0*	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.6)	(0.5-1.0)	(0.0-0.0)	(0.0-0.5)	(0.0-0.2)	(0.0-0.0)
Sales of goods and services	0.5*	0	0	0.3*	0	0	0	0
	(0.1-0.5)	(0.0-0.0)	(0.0-0.0)	(0.1-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)
Personal income taxes	0.2	0	0	0.1	0	0	0	0
	(0.0-0.8)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Corporate income taxes	0	0	0	0.3	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.8)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Consumption taxes (other than environmental)	0.1	0	0	1	0	0	0	0
	(0.0-0.7)	(0.0-0.0)	(0.0-0.1)	(0.3-2.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	(0.3-0.4)	(0.2-0.4)	(0.3-0.3)	(0.1-0.1)	(0.1-0.3)	(0.3-0.3)	(0.2-0.3)	(0.0-0.2)
Used revenue residual	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)	(0.1-0.2)	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)
Share spending efforts	80	100	95	29	87	80	92	100
	(69-80)	(84-100)	(81-97)	(29-29)	(66-87)	(74-83)	(82-100)	(100-100)
Achieved consolidation	5	4.7	4.5	4.4	3.7	3.6	3.6	3.6
Consolidation needs	5	4.7	4.5	4.4	3.7	3.6	3.6	3.6
Share top 9 instruments	93	100	100	65	100	100	100	100
	(83-100)	(100-100)	(100-100)	(50-98)	(100-100)	(100-100)	(100-100)	(100-100)
Instruments crossing the median	2	2	1	4	3	4	3	0
	(1-2)	(2-3)	(1-2)	(3-5)	(2-3)	(3-4)	(2-4)	(0-0)

Note: See note to Table A1.

Table A6. **Instrument use and achieved long-term consolidation vs. needs** *(cont.)*

Percentage of potential GDP except if otherwise mentioned

Description	LUX	BEL	HUN	DNK	AUT	CHE	FIN	SWE
Subsidies	0.8*	0.8	0.7*	0.8	0.8	0.8	0.7*	0.7*
	(0.8-0.8)	(0.8-0.8)	(0.7-0.7)	(0.8-0.8)	(0.8-0.8)	(0.8-0.8)	(0.7-0.7)	(0.0-0.8)
Pensions	0	0	0.6	0.7	1.2	0	0	0.9*
	(0.0-0.0)	(0.0-0.0)	(0.3-0.6)	(0.2-0.7)	(0.8-1.3)	(0.0-0.0)	(0.0-0.0)	(0.3-1.5)
Other in kind consumption (excluding family policy)	0	1.3*	0.9	0.4	0	0.6	0.8	0
	(0.0-0.0)	(0.3-1.9)	(0.4-1.3)	(0.0-1.0)	(0.0-0.0)	(0.1-0.6)	(0.0-1.1)	(0.0-0.5)
Unemployment benefits	0.5*	0.6	0.1	0.1	0	0.2	0.3	0
	(0.5-0.5)	(0.0-0.9)	(0.0-0.2)	(0.0-0.6)	(0.0-0.3)	(0.0-0.2)	(0.0-0.9)	(0.0-0.0)
Environmental taxes	0.2	0	0	0	0	0.3	0	0
	(0.1-0.2)	(0.0-0.6)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.2)	(0.0-0.4)	(0.0-0.0)	(0.0-0.0)	(0.0-0.4)	(0.0-0.0)
Sickness and disability payments	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.2)	(0.0-0.1)	(0.0-0.1)	(0.0-0.0)	(0.0-0.3)	(0.0-0.2)	(0.0-0.0)
Recurrent taxes on immovable property	0.5	0	0	0	0	0	0	0
	(0.0-1.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)
Sales of goods and services	0.5	0	0	0	0	0	0	0
	(0.0-1.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Personal income taxes	0	0	0	0	0	0	0	0
	(0.0-0.4)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Corporate income taxes	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Consumption taxes (other than environmental)	0	0	0	0	0	0	0	0
	(0.0-0.5)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	(0.2-0.3)	(0.1-0.1)	(0.1-0.2)	(0.1-0.2)	(0.1-0.1)	(0.1-0.2)	(0.1-0.1)	(0.0-0.1)
Used revenue residual	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Share spending efforts	58	100	100	100	100	86	100	100
	(57-59)	(78-100)	(92-100)	(83-100)	(100-100)	(65-100)	(75-100)	(100-100)
Achieved consolidation	2.8	2.8	2.4	2.3	2.2	2.1	1.8	1.7
Consolidation needs	2.8	2.8	2.4	2.3	2.2	2.1	1.8	1.7
Share top 9 instruments	100	100	100	100	100	100	100	100
	(89-100)	(100-100)	(100-100)	(100-100)	(100-100)	(100-100)	(100-100)	(100-100)
Instruments crossing the median	2	1	1	0	0	0	1	2
	(2-3)	(1-2)	(1-2)	(0-1)	(0-1)	(0-1)	(1-2)	(1-2)

Note: See note to Table A1.

Table A6. **Instrument use and achieved long-term consolidation vs. needs** *(cont.)*

Percentage of potential GDP except if otherwise mentioned

Description	DEU	EST	ISL	KOR	GRC
Subsidies	0.3	0.3	0.8*	0	0
	(0.3-0.3)	(0.3-0.3)	(0.7-0.8)	(0.0-0.0)	(0.0-0.0)
Pensions	0.8	0	0	0	0.9
	(0.6-0.8)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.5-0.9)
Other in kind consumption (excluding family policy)	0	1.3*	0.2	0	0
	(0.0-0.0)	(0.4-1.3)	(0.0-0.3)	(0.0-0.0)	(0.0-0.4)
Unemployment benefits	0.4	0	0.1	0.1	0
	(0.0-0.6)	(0.0-0.0)	(0.0-0.3)	(0.1-0.1)	(0.0-0.0)
Environmental taxes	0	0	0	0	0
	(0.0-0.4)	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0	0	0	0	0
	(0.0-0.4)	(0.0-0.7)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)
Sickness and disability payments	0	0	0	0	0
	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Recurrent taxes on immovable property	0	0	0	0.3	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.6)	(0.0-0.0)
Sales of goods and services	0	0	0	0.5	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.9)	(0.0-0.0)
Personal income taxes	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.6)	(0.0-0.0)
Corporate income taxes	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Consumption taxes (other than environmental)	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.5)	(0.0-0.0)
Public investment	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	(0.1-0.1)	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)
Used revenue residual	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Share spending efforts	100	100	100	12	100
	(76-100)	(53-100)	(76-100)	(12-12)	(100-100)
Achieved consolidation	1.6	1.6	1.2	1	0.9
Consolidation needs	1.6	1.6	1.2	1	0.9
Share top 9 instruments	100	100	100	100	100
	(100-100)	(100-100)	(100-100)	(74-100)	(100-100)
Instruments crossing the Median	0	1	1	0	0
	(0-1)	(1-1)	(1-1)	(0-0)	(0-0)

Note: See note to Table A1.

APPENDIX B

Supporting material

B1. Estimating fiscal consolidation needs: Additional technical information

The present section provides additional background information on the calculation of fiscal consolidation needs that are presented in Section 2 of the main text. The present set of consolidation needs estimates is aimed at indicating a way in which the fiscal balance can evolve so as to reduce (or contain) public debt to a manageable level (set at 60% of GDP) and keep it there at the end of the projection period (2060). A target of 60% for gross general government debt relative to GDP has been chosen for reasons of consistency with the OECD long-term growth scenarios (OECD, 2012a; Johansson et al., 2013; OECD, 2013). In addition to reaching this 60% target, a second objective of the consolidation path is to keep debt stable at this 60% of GDP by 2060. The debt target is specified in gross terms, while government financial assets are assumed to remain constant as a ratio to GDP.

The debt to GDP ratio d_i^t in country *i* at the end of year *t* results primarily from the past debt ratio d_i^{t-1}, the real interest rate r_i^t paid on government debt, the real growth rate g_i^t, and the primary balance (as a share of GDP) p_i^t. Another driver of the debt ratio is the income generated by government financial assets (at the real rate $\tilde{r}_i^t$) net of the reinvestment needed to keep the asset-GDP ratio a_i constant. Consequently, the debt ratio evolves according to the following accumulation equation:

$$d_i^t = d_i^{t-1}(1 + r_i^t - g_i^t) - p_i^t + a_i(\tilde{r}_i^t - g_i^t) \quad \text{(E1)}$$

The primary balance ratio to GDP can itself be written as a function of the underlying primary balance as a ratio to potential GDP $\underline{p}_i^t$, the output gap as a ratio to potential GDP o_i^t, a semi-elasticity ε_i and the amount of one-off factors f_i^t:

$$p_i^t = \underline{p}_i^t + \varepsilon_i o_i^t + f_i^t \quad \text{(E2)}$$

In the present study, the fiscal consolidation path is defined as a trajectory for $\underline{p}_i^t$ that starts from the 2012 estimate in the *OECD Economic Outlook*, May 2013 and then takes projected values from 2013 to 2060. Using equation (E2), the fiscal consolidation path $\underline{p}_i^t$ is combined with the projections for $\varepsilon_i o_i^t + f_i^t$ in the *OECD Economic Outlook Long-Term Database*, May 2013 to obtain projections of the primary deficit ratio p_i^t. from 2013 to 2060. This path for the primary deficit ratio p_i^t is then fed into equation (E1) to calculate the debt trajectory starting from the 2012 estimate of d_i^t and the long-term projections for $r_i^t - g_i^t$ in the *OECD Economic Outlook*, May 2013. The rates r_i^t and $\tilde{r}_i^t$ are the effective interest rates paid and received by the government on its liabilities and assets, respectively, as projected in the *OECD Economic Outlook Long-Term Database*, May 2013, deflated by the rate of change

of the GDP deflator taken from the same source.[1] The effective interest rate r_i^t takes into account the maturity structure of the stock of government debt.

The consolidation path $\left(\underline{p}_i^t\right)_{t=2013}^{2060}$ must satisfy the desired objectives of debt reaching 60% of GDP by 2060 (O1) and being stable (O2):

$$d_i^{2060} = 60\% \quad \text{(O1)}$$

$$\frac{\partial}{\partial t} d_i^t \Big|_{t=2060} = 0 \quad \text{(O2)}$$

Using equation (E1) and then (O2) and observing that cyclical effects and one-offs have vanished by 2060, it follows that the stability objective (O2) is satisfied if and only if:

$$\underline{p}_i^{2060} = p_i^{2060} = d_i^{2060}\left(r_i^{2060} - g_i^{2060}\right) + a_i\left(\tilde{r}_i^t - g_i^t\right) \quad \text{(E3)}$$

The stability condition (O2) therefore pins down the end point for the consolidation path. Nevertheless, there are an infinite number of consolidation paths $\left(\underline{p}_i^t\right)_{t=2013}^{2060}$ that satisfy the first objective (O1) of a 60% debt-to-GDP ratio by 2060.

The present study chooses a particular functional form for the consolidation path. It assumes that the underlying primary balance moves by 1% of GDP each year until it reaches a certain value $\underline{p}_i^*$ and from then on evolves in equal yearly steps until its end value. In symbols instead of words, this functional form is defined as follows: $f\left(\underline{p}_i^*, \underline{p}_i^{2060}, t\right)$

$$= \underline{p}_i^{2012} + (t - 2012)\,\text{sign}\left(\underline{p}_i^* - \underline{p}_i^{2012}\right) \quad \text{if} \quad t \le t_i^* \quad \text{where} \quad t_i^* = 2012 + \text{abs}\left|\underline{p}_i^* - \underline{p}_i^{2012}\right| \quad \text{(D1)}$$

$$= \underline{p}_i^* + \frac{(t - t_i^*)}{(2060 - t_i^*)}\left(\underline{p}_i^{2060} - \underline{p}_i^*\right) \text{ if } t > t_i^* \quad \text{(D2)}$$

With this choice of functional form,[2] the problem reduces to finding the value for $\underline{p}_i^*$ such that, with $\underline{p}_i^t = f\left(\underline{p}_i^*, \underline{p}_i^{2060}, t\right)$, the debt-accumulation equation (E1) yields $d_i^{2060} = 60\%$.

The study covers all the OECD membership except Chile, Mexico and Turkey where the necessary national accounts data and projections for the general government sector are not available. As mentioned above, all the data and projections for the variables mentioned in this section are taken from the OECD long-term baseline published in the *OECD Economic Outlook*, May 2013.

The consolidation paths for all countries covered are plotted in Appendix 3 of Cournède, Goujard and Pina (2013). Most countries follow the hump shape depicted in Figures 2 and 3 of the main text $\left(\underline{p}_i^* > \underline{p}_i^{2060} > \underline{p}_i^{2012}\right)$: they first have to increase their primary balance to get rid of legacy debt before they can allow it to diminish to reach the lower level that is consistent with debt being stable as a ratio to GDP at 60%. A number of countries (Australia, Czech Republic, Finland, New Zealand, Poland, Slovak Republic, Slovenia), however, do not exhibit this pattern as they start from low debt so that the short- to medium-term level of the surplus required to make sure that debt remains below 60% until 2060 is lower than the surplus required in 2060 to keep the debt ratio stable $\left(\underline{p}_i^{2012} \le \underline{p}_i^* < \underline{p}_i^{2060}\right)$. Finally, in a group of countries (Denmark, Estonia, Korea, Luxembourg, Norway, Sweden, Switzerland) the uniform objective of a 60% debt-to-GDP ratio results mechanically in a profile of fiscal relaxation to reach the target from the bottom. Such a profile does not appear advisable in practice as the 60% target is set relatively high, leaving little room for shock absorption, and because fiscal relaxation often proves very difficult to reverse in practice. These countries are considered as having no short- to medium-term consolidation need, and the corresponding discussion does not apply to them.

Consolidation needs are estimated by comparing these "debt-control" trajectories for the underlying primary surplus that stabilise debt at 60% by 2060 with a baseline $\underline{\check{p}}_i^t$. As described in the main text, the baseline for the underlying primary surplus starts from the 2012 estimated value in the *OECD Economic Outlook*, May 2013 $\underline{\check{p}}_i^{2012} = \underline{p}_i^{2012}$. The baseline incorporates the assumption that pension reforms, including adjustments to the retirement age in line with longevity, keep public pension spending stable as a share of potential GDP. The baseline also incorporates some reform of health and long-term care so that public spending in this area increases in line with the projections h_i^t in the "cost-containment" scenario presented in de la Maisonneuve and Oliveira Martins (2013). In other areas, fiscal policy parameters are assumed to be unchanged in the sense that corresponding spending and revenue items stay constant as a proportion of potential GDP. As a result, the baseline underlying primary balance evolves as:

$$\underline{\check{p}}_i^t = \underline{p}_i^{2012} - \left(h_i^t - h_i^{2012}\right).$$

The short- to medium-term consolidation need c_i^{sm} is defined as the difference between the underlying primary surplus in the year t_i^* when the value $\underline{p}_i^*$ is reached and the baseline in the same year: $c_i^{sm} = \underline{p}_i^* - \underline{\check{p}}_i^{t_i^*}$. Given the choice of increasing the underlying primary balance by 1% of potential GDP per year, there is a very direct link between the duration and size of consolidation. The short- to medium-term part of the simulated consolidation is completed by 2017 in the vast majority of covered countries, the exceptions being Greece, Ireland, Japan, Portugal, the United Kingdom and the United States.

The long-term consolidation need c_i^l is defined as the difference between the underlying primary surplus that stabilises a 60% debt-GDP ratio and its baseline level in 2060: $c_i^l = \underline{p}_i^{2060} - \underline{\check{p}}_i^{2060}$. With a common target debt-ratio, differences in the long-term objective for the primary surplus only reflect differences in assumptions about real growth and interest rates by 2060 in the OECD long-term scenarios, which are small. The variation in estimated long-term consolidation needs (Figure 3 of the main text), therefore, primarily reflect differences in the baseline. The cross-country variation in projected increases in public health and long-term care spending is relatively limited in the cost-containment scenario. As a result, the main source of variation in estimated long-term consolidation needs comes from the starting point for the underlying primary surplus.

B2. The dataset of fiscal consolidation instruments

B.2.1. Defining the instruments

The revenue and expenditure sides of budgets are decomposed into sets of variables which may potentially be the object of consolidation measures, and are hence referred to as the instruments of consolidation. Their definition balances concerns of economic interpretability, parsimony and data availability, while ensuring a consistent accounting framework where no budget items are missed or overlap, and hence the various categories of spending and revenue sum to total spending and revenue. A total of 17 consolidation instruments are identified, plus residual and property income items (Tables B1 and B2). Revenue instruments span the main categories of taxes plus a non-tax item, sales of goods and services (mainly user charges). Spending instruments articulate the economic and functional classifications of expenditure, where the former focuses on the economic nature of transactions (such as wages, cash transfers or public investment) and the latter on their purpose (like education, health or defence). Residual spending and revenue

categories are heterogeneous and hence hard to characterise: though they are not taken as instruments, they are assumed to take part in the overall consolidation effort by keeping a constant share to underlying primary disbursements and receipts. Property income (paid or received) follows from the trajectories of interest rates and asset stocks.

Consolidation instruments are hence defined in a quantitative, internationally-comparable way, and their use to achieve budget targets does not map into specific policy measures. For instance, resorting to an increase in personal income taxes could be done either through higher marginal rates or by curbing tax expenditures. Likewise, spending on health care could be brought down either by reducing the levels of service provision or by achieving efficiency gains. Structural reforms of taxation and public spending, of which tackling tax expenditures and spending inefficiencies are prime examples, are strongly advocated, especially when large consolidation requirements force governments to use instruments with growth- or equity-damaging impacts. However, severe data limitations, as well as the need to preserve a consistent accounting framework, preclude the identification of tax expenditures and efficiency gains in specific areas as autonomous instruments. Though the scope for base broadening can be quantified for some taxes, such as VAT (through the VAT revenue ratio), efforts at international comparisons of tax expenditures across various tax bases have been subject to important limitations and covered only a subset of OECD members (OECD, 2010a). Potential savings from efficiency gains have been estimated in only a limited number of areas (e.g. Sutherland et al., 2007, for primary and secondary education; OECD, 2010b, for health care) and are also subject to important qualifications. For these reasons, the study considers the quantitative changes in the instruments of fiscal consolidation (spending cuts and revenue increases) separately from the structural measures that can accompany or contribute to them.

A dataset containing estimates in 2012 of all consolidation instruments in 31 OECD countries has been constructed, as described in more detail below (severe data limitations have prevented the inclusion of Chile, Mexico and Turkey). Using values in 2012 ensures consistency with the estimates of consolidation requirements and incorporates the effects of the various consolidation measures already implemented by member countries.[3] Further, instruments are defined in underlying terms (i.e. net of cyclical effects and of one-offs), as a percentage of potential GDP. This avoids distortions in the cross-country comparison of fiscal variables. For instance, the use of nominal values as a percentage of actual GDP could have resulted in overestimating the structural importance of spending items in countries with big negative output gaps.

B.2.2. Constructing the fiscal dataset

Underlying instrument values in 2012 have been estimated in three main steps. First, actual values in the most recent year available have been collected. Second, values have been updated to 2012 on the basis of the dataset in the *OECD Economic Outlook*, May 2013 (OECD, 2013). Third, instruments have been adjusted for the cycle and for one-offs, also in a way consistent with OECD (2013).

In the first step, data is collected or estimated for the most recent year available in both the main statistical sources, national accounts and the *OECD Revenue Statistics*. This year (call it year Y) varies across countries, though it is often 2010 or 2011 (Table B3). However, on the spending side, though these sources give the totals for wages, intermediate consumption and transfers in kind and for social transfers other than in kind, they do not provide the functional breakdown defining the first seven instruments, as

COFOG data is insufficiently disaggregated or simply not available. Alternative sources are then used (Tables B1, B2 and B3) and, when they refer to a year before year Y, values are extrapolated by assuming a constant share of the instrument in question in the respective national accounts total. For instance, *Education at a Glance* data for 2009 is extrapolated to 2010 (assuming this is year Y) by preserving the 2009 share of education spending relative to total wages, intermediate consumption and transfers in kind.

In the second step, the change from year Y to 2012 in the fiscal variables contained in OECD (2013) is allocated to the several instruments of consolidation (plus residual and property income items), ensuring that total revenue and total expenditure in the dataset of instruments equal those totals in OECD (2013). Tables B1 and B2 set the correspondence used. Since the dataset of instruments is more disaggregated than the OECD (2013) fiscal block, it is often assumed that instruments keep a constant share relative to certain variables. For instance, while the change in the OECD (2013) variable TKTRG is fully allocated to the expenditure residual item, the change in TOCR is split between sales of goods and services and the residual item by assuming that the year Y ratio between sales and the residual (net of TKTRG) is maintained. On the spending side, it is worth underlining that unemployment benefits are assumed to remain a constant proportion of cash transfers in cyclically-adjusted rather than nominal terms.

In the third step, instruments are adjusted for the cycle and for one-offs, drawing again on OECD (2013) fiscal variables (Tables B1 and B2) and, as above, often making an assumption of constant proportions. Adjusted instruments are in all cases taken as a percentage of potential, rather than actual, GDP.

B.2.3. Country-specific issues

For each country, Table B3 documents the date of the latest available information from several sources. Further, Swiss institutional features have given rise to a number of adjustments. In Switzerland, most health expenditure recorded as public in the health accounts is not classified as general government expenditure in national accounts, since the system is largely managed by private non-profit organisations (OECD, 2011a). This would imply an artificially low value for other government consumption (item 3 in Table B2), since the latter is obtained by difference. To correct this problem, the difference in public spending on health between health accounts and national accounts is added to other government consumption. On the revenue side, a similar adjustment is made for social contributions to ensure budget-balance neutrality and improve international comparability.

B.3. Additional considerations on the growth and equity effects of consolidation instruments

This section gives detail, additional to that in the main text, on the underpinning of the assessment summarised by the plus and minus signs in Table 2 of the main text. The assessment proposed in Table 2 focuses on the effects of consolidation instruments themselves and does not incorporate the possible response of joint reform of structural or redistributive policies. In principle, all the long-term-growth-damaging instruments listed below can be accompanied by structural reforms that could mitigate, offset, or even more than offset, their harmful growth implications. Similarly, all regressive instruments can be supplemented by redistributive provisions tailored to preserve, or equalise, the income distribution. For many instruments the same behavioural responses to their

implementation shape long-term effects along both the growth and equity dimensions which are therefore discussed together in the present section.

B.3.1. Government consumption

B.3.1.1. Health care

Government spending on health care reduces inequality in effective consumption (Hoeller et al., 2012). The favourable welfare effects of providing health services to people who cannot afford them otherwise are likely to be particularly large.

With good health contributing to human capital preservation, health spending can be expected to contribute positively to the productive potential of an economy (Joumard et al., 2010). Investment in health has indeed been found to boost economic output in OECD countries (Gyimah-Brempong and Wilson, 2004). On the plausible assumption that private funding cannot fully substitute for public expenditure, government spending on health care can be expected to boost potential output, a positive link that seems present in the data for OECD countries (see Barbiero and Cournède, 2013).

B.3.1.2. Education

Empirical estimates suggest that the distributional effect of government spending is progressive for primary and secondary education and regressive for tertiary education (OECD, 2011). Given that governments spend more on the primary and secondary levels of education, an across-the-board cut in education spending can be considered as regressive. Government spending on education generally tends to increase intergenerational social mobility as well (Causa and Johansson, 2010), meaning that the inequitable consequences of cuts in this area are likely to grow over time.

B.3.1.3. Other government consumption

Cuts in other government consumption that include reductions in government employment are likely to increase income inequality (Fournier and Koske, 2012). More broadly, reducing the government wage bill can be expected to increase income inequality in most countries (Bova et al., 2012). Cuts in government-provided services will also have the effect of increasing inequality in effective consumption.

B.3.2. Government investment

OECD work on the sources of economic growth found a positive long-term relationship between private and public capital accumulation, suggesting complementarity between the two and a positive effect of government investment on long-term output (see Table 2.7 in OECD, 2003a; and Sutherland et al., 2009). On US data, Leduc and Wilson (forthcoming) identify strong, positive long-term GDP effects of road infrastructure grants. Evidence from 583 studies surveyed by Bom and Ligthart (2013) indicates that public capital has positive short- and long-run effects on output. With a depreciation rate of 5% and a real interest rate of 3%, the marginal cost of public capital (8%) is lower than the estimated marginal rate of return (10 to 24%). This situation suggests underinvestment in public capital, particularly for core public capital where the marginal rate of return may be close to 24%.

Evidence on the effect of public investment on inequality in OECD countries is too limited for inclusion in the assessment summarised in Table 2 of the main text. Estimates by Calderon and Serven (2004) show that infrastructure development, a key area of

government investment, promotes income equality. This study has however been conducted with a database consisting primarily of developing countries, which raises a question as to whether its finding applies to advanced economies. Furthermore, a model-based study by Chatterjee and Turnovsky (2012) suggests a conceptual channel whereby government investment could increase wealth inequality: because of complementarities between public and private capital, a greater public capital stock raises the value of private capital, the holding of which is concentrated at the top end of the wealth distribution. Finally, the short-term impact of government investment on the income distribution is likely to be positive given the large role of low-skilled labour in construction work, but there is little empirical evidence to document this effect.

B.3.3. Transfers

Transfers equalise income more strongly than taxes (Rawdanowicz et al., 2012; Joumard et al., 2012; Bastagli et al., 2012). On the other hand, aggregate higher social protection expenditure, especially "passive" income support, much of which is achieved through transfers, has been found to be associated with lower economic growth (Arjona et al., 2002). Some categories, such as active labour market programmes (ALMP), are linked with higher growth (Arjona et al., 2002), but most of ALMP spending is recorded as government consumption or subsidies rather than transfers.

B.3.3.1. Pensions

The distributional impact of reducing public pension spending can change direction depending on the way in which it is implemented. Reducing spending by increasing the average retirement age will have the effect that people who would otherwise have retired and drawn a sometimes low pension will keep drawing a higher salary. The effect is likely to be a more equal distribution of income. On the other hand, as public pension systems are progressive, although to varying degrees across OECD countries, reducing replacement rates would tend to be regressive on impact. The effect will diminish over time if labour supply increases in response.

B.3.3.2. Unemployment benefits

Reducing unemployment benefits increases inequality, even in insurance-based systems because less paid workers are more likely to be laid off. In the long term, however, reducing the level of unemployment benefits translates into higher employment rates, implying little if any change in steady-state income distribution (Rawdanowicz et al., 2012).

B.3.3.3. Sickness and disability benefits

Reducing sickness and disability benefits would mechanically increase inequality as replacement rate tend to be higher for low-income earners (OECD, 2010). However, in the long term, poorly-targeted disability benefits can lock up individuals with residual working capacity in poverty traps with negative growth effects and potentially negative effects on equity. In that regard, well-targeted reductions in government spending on disability benefits are likely to have regressive effects that diminish over time (Rawdanowicz et al., 2012). Unfocused budgetary cuts will have lasting regressive effects since part of them fall on people who cannot supply labour to compensate.

B.3.4. Subsidies

In the present taxonomy of instruments, following the national accounts framework, subsidies are defined as current payments to businesses based on their production levels. Unless they correct market failures, such production subsidies have the potential to distort resource allocation and restrict competition (OECD, 2001) and have been long identified as reducing productive potential (Ford and Suyker, 1990). Agricultural and fossil-fuel subsidies in particular can, in addition, encourage inefficient use of natural resources (OECD, 2003b), eroding the resource base for long-term growth and harming those most exposed to environmental degradation, which are typically concentrated at the bottom of the income distribution (Serret and Johnstone, 2006).

B.3.5. Taxes

An increase in distortionary taxes (labour and capital income taxes, social security contributions, property taxes) of 1% of GDP is estimated to reduce the level of GDP by 3.2% after ten years according to a recent empirical study (Gemmell, Kneller and Sanz, 2011).

B.3.5.1. Tax expenditures

The present study does not include tax expenditures in its quantitative analysis of possible consolidation strategies for lack of comparable data across the OECD membership. Tax expenditures have the potential of distorting resource allocation and hurting growth. To give just one example, the combination of mortgage interest relief with no taxation of owner-occupied rental services favours investment in housing over other, typically more productive, forms of capital. With the notable exception of tax relief or credit for low-wage workers, many tax deductions have higher value for high-income earners and are therefore regressive (Rawdanowicz et al., 2012).

B.3.5.2. Consumption taxes

Consumption taxes are typically regressive because lower-income households spend a higher share of their income (Rawdanowicz et al., 2012). Fiscal authorities often try to limit the regressive effects of indirect taxes by applying reduced rates on bare necessities, but these reductions have large fiscal costs as they are enjoyed not only by low-income households but by all consumers.

B.3.5.3. Personal income tax and social security contributions

Progressive income taxation reduces growth and equalises after-tax income immediately (Joumard et al., 2012) and permanently (Echevarria, 2012). In many countries, however, a wide array of tax breaks, which have higher value for higher-income households, reduces the redistributive power of personal income taxation. Social security contributions, in contrast, tend to be regressive as they are often concentrated on labour income, which makes a smaller share of the income of higher-income households.

The direction of the growth effects of progressive income taxation, compared with flatter tax schedules, is ambiguous in theory. Progressive income taxation reduces the return on work more for more productive workers. Depending on whether they work more to maintain their income (the income effect) or less because of the higher after-tax relative value of leisure (the substitution effect), progressive taxation will boost or reduce output. Another channel is that, in the presence of financial constraints, progressive tax schedules leave lower-income workers with more after-tax income from which they can build savings

to start their own business. In the first instance, this channel means that progressive taxation should be growth-enhancing. However, general-equilibrium effects push up the wages of low-productivity workers under flat tax schedules, which can partly compensate, or more than offset, the previous direct effect depending on the characteristics of the progressive and flat tax schemes being compared (Bohacek and Zubricki, 2012). Overall, the question is empirical, and the evidence is largely conclusive that progressive taxation reduces long-term growth (Johansson et al., 2008). Social security contribution schemes are typically far less progressive than income tax (OECD, 2012b).

B.3.5.4. Corporate income tax

Corporate income taxes are one way of taxing capital income, directly and also by limiting the incentive for owners to shelter capital earnings from personal income taxation by keeping them in corporations (Johansson et al., 2008). As capital control is concentrated among higher-income households, corporate income taxation works in the direction of equalising income.

A large body of empirical evidence indicates that corporate income taxes are highly distortive to long-term output (see for instance Johansson et al., 2008). Because it is calculated on earnings after interest payments, corporate income taxation creates a strong bias in favour of debt over equity financing, contributing to excess leverage and instability (Sutherland et al., 2012; De Mooij, 2012). Higher corporate income taxes have been found to deter FDI (Becker et al., 2012).

B.3.5.5. Recurring taxes on immovable property

In a number of OECD countries lower-income households pay a larger share of their earnings in recurring real estate taxes than higher-income households, so that recurring property taxes can be seen as regressive in the static income distribution. However, this situation is to some extent (but not fully) related to larger property holdings among pensioners, implying that property taxation is not necessarily regressive in the perspective of lifetime income or wealth distribution (Rawdanowicz et al., 2012). Furthermore, some OECD countries have introduced exemptions or rebates for low-income households to alleviate, or reverse, the regressive effects of recurring real estate taxes (Joumard et al., 2012). Against this background, increases in recurring taxes on immovable property cannot be described as being generally progressive or regressive.

B.3.5.6. Other property taxes

Inheritance, gift and wealth taxes, which make up much of this category, are highly redistributive because tax schedules are generally strongly progressive and wealth is more concentrated than income (Joumard et al., 2012). The progressive effects of increases, however, are likely to diminish over time due to the mobility of many forms of capital and the greater incentive to put in place tax avoidance strategies for more wealthy households.

B.3.5.7. Environmental taxes

Environmental taxes can improve welfare by increasing the marginal cost of activities that harm the environment. This welfare benefit is not measured in GDP which for the most part does not record environmental services. One of the effects of environmental taxes, however, is to support long-term output to the extent that they discourage the excessive use of environmental resources that are needed for future production. However,

another of their effects, shared with other taxes, is to reduce labour supply and physical capital accumulation. Where applying to internationally-mobile activities, environmental taxation will entail smaller gross effects on GDP and bring larger environmental benefits if it is co-ordinated across countries.

Lower-income households, who spend more of their income than better-off households on basic resources, therefore tend to pay a greater share of their income in environmental taxes. *Prima facie*, environmental taxes may therefore be considered as regressive. However, lower-income households are more exposed to environmental degradation than the rest of the population. As a result, when the benefits of higher environmental taxation materialise, the net welfare effect of environmental taxation may not necessarily be regressive.

B.3.6. User charges

Charging more for the use of public services can boost growth by avoiding their inefficient use. The distributional effects of increases in user charges are likely to be regressive except in cases (such as tertiary education) where public services are used more heavily by higher-income groups.

B.4. Methodological information on estimating debt behaviour during consolidation episodes

This section provides detailed information about the methodology used to document the behaviour of debt-GDP ratios during consolidation episodes shown in Figure 6 of the main text. Episodes of fiscal consolidation are defined using the IMF Action-Based Dataset (Devries et al., 2011). This definition of fiscal adjustments and discretionary policy choices of fiscal authorities is, in principle, less affected by countries' macroeconomic conditions than measures of the fiscal stance based on cyclical adjustment (Guajardo et al., 2011). A period of fiscal adjustment is defined as a year in which the action-based budget balance improves by at least 1½ per cent of GDP. As a robustness check, two thresholds of 1.75 and 2% of GDP were also used, but the results were very similar to the ones reported in Figure 6. This definition rules out fiscal adjustments which are small annually but prolonged over several years. Furthermore, it selects both one-year consolidation episodes and multi-period adjustments as in Alesina and Ardagna (2010).

Figure 6 in the main text presents difference-in-differences estimates of the association between the public debt ratio d and consolidation episodes, from 10 years before to 10 years after the onset of a consolidation. These specifications control for country and year fixed effects according to the following specification:

$$d_{ict} = \sum_{k=-10}^{+10} \beta_k 1_{t=c+k} + \delta_i + \delta_t + \varepsilon_{ict}$$

where c stands for the starting year of a consolidation episode, I the country and t the year under consideration. The period from 15 to 11 years before the onset of a consolidation episode is omitted, meaning that all estimates are relative to these earlier years of data. Each country-year observation has the same weight in the estimation. On the solid lines, each point corresponds to the estimate of β_k, the difference between the level of public debt in pre- or post-consolidation year k relative to the rest of the sample. Dashed lines represent the 90% confidence interval adjusted for heteroskedasticity, spatial autocorrelation and within-country autocorrelation as in Driscoll and Kraay (1998) and Auerbach and Gorodnichenko (2012).

B.5. Alternative simulations based on simple country groups instead of clustering

A variant has been prepared to test the sensitivity of clustering-based instrument rankings and consolidation packages. In this variant, instead of deriving weights from clustering analysis, countries are divided in three groups of identical sizes based on their current-account positions, and each objective receives the same weight within each group. Current account-positions are evaluated on the basis of cyclically adjusted current account balances as shares of both national and OECD GDP. The country groups and weights are as follows:

1. A first group gathers 10 countries with the largest current-account deficits: Australia, France, Greece, Iceland, New Zealand, Poland, Portugal, Spain, the United Kingdom and the United States. Weights of 20% each are given to each objective (short- and long-term growth and equity, current account).
2. A second group gathers 11 countries with the largest current-account surpluses: Austria, Denmark, Germany, Ireland, Japan, Korea, Luxembourg, Netherlands, the Slovak Republic, Sweden and Switzerland. Weights of 20% each are given to each objective (short- and long-term growth and equity, current account). The signs of the assessments of current-account effects shown in Table 3 of the main text are reversed to take into account the surplus position of these countries.
3. The third group gathers the remaining 10 countries: Belgium, Canada, the Czech Republic, Estonia, Finland, Hungary, Israel, Italy, Norway and Slovenia. Weights of 25% each are given to each objective (short- and long-term growth and equity).

The rankings that result from this alternative weighting scheme are very highly correlated with the ones obtained with the clustering technique reported in the main text. Rank correlations between clustering-based and alternative rankings are 96% for Cluster 1, 98% for Cluster 2, 93% for Cluster 3, 89% for Cluster 4 and 95% for Cluster 5.

As a result, the consolidation packages are also very similar as is apparent from comparing Tables B4 to B8 which refer to the alternative simulations to Tables A1 to A5 which refer to the main ones. To facilitate comparison with the main Tables A1 to A5, the Tables B4 to B8 for the variant are shown with the same grouping of countries by cluster although uniform weights for the three above-listed groups have been used. Overall, the simulated consolidation packages in the variant are very close to the clustering-based ones. Noticeable differences appear only in Clusters 3 and 4 where clustering-based weights differ substantially from 25%. For instance, the instrument mix of the Slovak Republic and Spain changes with greater use of personal income taxes. The variant is similarly robust to shocks to the assessment of effects as the main clustering-based results. The interdecile variations shown between brackets in Tables B4 to B8 remain generally narrow.

B.6. Alternative simulations allowing additional room for manoeuvre

An important assumption in the simulation design relates to the amount up to which each instrument can be used. As presented in greater detail in Box 4, which also spells out the adjustments made to deal with special cases, two constraints are imposed in the main simulations:

1. A tax can be increased up until proceeds from this tax reach the 66th percentile of the cross-country distribution of receipts from this kind of tax, as a share of potential GDP.

Similarly, spending in one area can be reduced until the country reaches the 33rd percentile of the cross-country distribution of spending in this area, as a share of potential GDP.

2. Adjustment along one instrument is capped at one standard deviation of the cross-country distribution of tax or spending in the area under consideration as a share of potential GDP.

In the present set of alternative simulations, the first constraint is relaxed by allowing movement until reaching the 80th percentile for taxes and the 20th percentile for spending. The second constraint is retained.

Tables B9 to B13 present country details of the short- to medium-term packages simulated under the relaxed first constraint on instrument use. Compared with the main results reported in Tables A1 to A5, these alternative simulations show that allowing more room for manoeuvre makes it possible for countries to achieve much more of their adjustment with well-ranked instruments. The same finding holds for the simulated long-term packages reported in Table B14 for the variant with additional room for manoeuvre and Table A6 for the main scenario. These findings are obtained even though the relaxation of the constraint is relatively moderate in the sense that the one-standard-deviation constraint is maintained.

Table B1. **Revenue categories: Definitions and sources**

Item No.	Description	Definitions and sources[1] (codes in brackets)	Updating to 2012: EO93 variables	Adjusting for the cycle and for one-offs: EO93 variables
1	Personal income taxes	RS (1100 + 1300)	Δ TYH (part)	TYHA-TYH (part)
2	Social security contributions	RS (2000 + 3000) + NA (D612r)	Δ SSRG (100%)	SSRGA-SSRG (100%)
3	Corporate income taxes	RS (1200)	Δ TYB (part)	TYBA-TYB (part)
4	Environmental taxes	OECD/EEA Database on instruments used for environmental policy	Δ TIND (part)	TINDA-TIND (part)
5	Other consumption taxes	RS (5000 + 6000) – item 4	Δ TIND (part)	TINDA-TIND (part)
6	Recurrent taxes on immovable property	RS (4100 + 4600)	Δ TYH (part), Δ TYB (part), Δ TIND (part)[2]	TINDA-TIND (part), TYHA-TYH (part), TYBA-TYB (part)[2]
7	Other property taxes	RS (4200 to 4500)	Δ TYH (part), Δ TYB (part), Δ TIND (part)[2]	TINDA-TIND (part), TYHA-TYH (part), TYBA-TYB (part)[2]
8	Sales of goods and services	NA (P11 + P12 + P131)	Δ TOCR (part)	
9	Other (non-property) receipts	NA (TR) – sum of items 1-8 and 10	Δ TKTRG (100%), Δ TOCR (part)	TKTRGU-TKTRG-TRGOE (100%)
10	Property income received	NA (D4r)	Δ YPERG (100%)	

1. RS = *OECD Revenue Statistics*; NA = *National Accounts*.
2. Items 6 and 7 are each assumed to consist in equal parts of direct (TY) and indirect (TIND) taxes, and, within TY, TYH and TYB are assumed to observe the proportions of items 1 and 3.

Note: EO93 stands for the *OECD Economic Outlook*, No. 93 (OECD, 2013).

Table B2. **Expenditure categories: definitions and sources**

Item No.	Description	Definition: National Accounts codes	Data sources and proxies used[1]	Updating to 2012: EO93 variables	Adjusting for the cycle and one-offs: EO93 variables
1	Education (wages, intermediate consumption and social transfers in kind)	D1p + P2 + D6311p + 63121p + 63131p within COFOG 09	OECD Education at a Glance[2]	Δ (CGAA-CFKG + revenue item 8) (part)	
2	Health (wages, intermediate consumption and social transfers in kind)	D1p + P2 + D6311p + 63121p + 63131p within COFOG 07	Health Accounts (SHA)[3]	Δ (CGAA-CFKG + revenue item 8) (part)	
3	Wages, intermediate consumption and transfers in kind – other purposes (except family and children)	D1p + P2 + D6311p + 63121p + 63131p within remaining COFOG divisions (except 10.4, family and children)	NA (D1p + P2 + D6311p + 63121p + 63131p) – item 7 (benefits in kind) – items 1 and 2	Δ (CGAA-CFKG + revenue item 8) (part)	
4	Old-age and survivor pensions (social transfers other than in kind)	D62p within COFOG 10.2 + 10.3	SOCX Database, old age and survivors (cash benefits)	Δ SSPG (part)	
5	Sickness and disability (social transfers other than in kind)	D62p within COFOG 10.1	SOCX Database, incapacity related (cash benefits)	Δ SSPG (part)	
6	Unemployment benefits (social transfers other than in kind)	D62p within COFOG 10.5	SOCX Database, unemployment (cash benefits)	Δ SSPG (part)	YPGA – YPG (100%)
7	Family and children (wages, intermediate consumption and social transfers in cash and in kind)	D1p + P2 + D62p + D6311p + 63121p + 63131p within COFOG 10.4	SOCX Database, family (cash benefits + benefits in kind)	Δ (CGAA-CFKG + revenue item 8) (part), Δ SSPG (part)	
8	Subsidies	D3p	NA (D3p)	Δ (YPOTG net of revenue item 8) (part)	
9	Public investment	P51	NA (P51)	Δ IGAA (100%)	
10	Other primary expenditure	D62p (remainder) + D29p + D5p + D7p + P52 + P53 + K2 + D9p	NA (TE, total expenditure) – items 1-9 and 11	Δ TKPG (100%), Δ (YPOTG net of revenue item 8) (part), Δ SSPG (part)	TKPGU – TKPG – TPGOE (100%)
11	Property income paid	D4p	NA (D4p)	Δ YPEPG (100%)	

1. NA = National Accounts.
2. General government expenditures (all levels, all educational programmes, and all types of transactions).
3. General government total expenditure (HC.1-HC.9; HC.R.1) for the UK, and general government total current expenditure (HC.1-HC.9, individual and collective health care) for all other countries.

Table B3. **Country data sources**

	National accounts	Revenue statistics	Health accounts	Social Expenditure Database	Education at a Glance	Environmental taxes
Australia	2009	2009	2009	2009	2009	2009
Austria	2011	2011	2010	2009	2009	2011
Belgium	2011	2011	2010	2009	2009	2011
Canada	2010	2010	2010	2009	2009	2010
Czech Republic	2011	2011	2010	2009	2009	2011
Denmark	2011	2011	2010	2009	2009	2011
Estonia	2011	2011	2010	2009	2009	2011
Finland	2011	2011	2010	2009	2009	2011
France	2011	2011	2010	2009	2009	2011
Germany	2011	2011	2010	2009	2009	2011
Greece	2010	2010	2007	2009	2005	2010
Hungary	2011	2011	2010	2009	2009	2011
Iceland	2011	2011	2010	2009	2009	2011
Ireland	2011	2011	2010	2009	2009	2011
Israel	2011	2011	2009	2009	2009	2010
Italy	2011	2011	2011	2009	2009	2011
Japan	2010	2010	2009	2009	2009	2010
Korea	2010	2010	2010	2009	2009	2010
Luxembourg	2011	2011	2009	2009	2002	2011
Netherlands	2010	2010	2010	2009	2009	2010
New Zealand	2010	2010	2010	2009	2009	2010
Norway	2011	2011	2011	2009	2009	2011
Poland	2010	2010	2010	2009	2009	2010
Portugal	2010	2010	2010	2009	2009	2010
Slovak Republic	2011	2011	2010	2009	2009	2011
Slovenia	2011	2011	2010	2009	2009	2010
Spain	2011	2011	2010	2009	2009	2011
Sweden	2011	2011	2010	2009	2009	2011
Switzerland	2011	2011	2010	2008	2009	2011
United Kingdom	2011	2011	2010	2009	2009	2011
United States	2011	2011	2010	2009	2009	2010
Median	2011	2011	2010	2009	2009	2011
Min.	2009	2009	2007	2008	2002	2009
Max.	2011	2011	2011	2009	2009	2011

Source: OECD National Accounts and Revenue Statistics, *OECD Health Accounts*, *OECD Social Expenditure Database*, *OECD Education at a Glance*, and the *OECD/European Environment Agency Database* on instruments used for environmental policy and natural resources management.

Table B4. **Instrument use and short- to medium-term consolidation with fixed weights in cluster No. 1**

Percentage of potential GDP except if otherwise mentioned

Description	AUS	CAN	GBR	ISR	ITA	JPN	NZL	POL	PRT
Subsidies	0.6*	0.3	0	0	0.3	0	0	0	0
	(0.6-0.6)	(0.3-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0	0.5*	0.2	0.2	0	0.4*	0.7*	0.7*	0.3*
	(0.0-0.0)	(0.5-0.5)	(0.2-0.2)	(0.2-0.2)	(0.0-0.0)	(0.3-0.3)	(0.7-0.7)	(0.7-0.7)	(0.3-0.3)
Pensions	0	0	0.3	0.2	0.3	0.8	0	0.9	0.9
	(0.0-0.0)	(0.0-0.0)	(0.3-0.3)	(0.2-0.2)	(0.3-0.7)	(0.8-0.8)	(0.0-0.0)	(0.9-0.9)	(0.9-0.9)
Corporate income taxes	0	0.2	0.5*	0	0	0	0	0.9*	0.3
	(0.0-0.0)	(0.0-0.2)	(0.5-0.5)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.9)	(0.0-0.3)
Personal income taxes	0	0	0.5	1.3	0	4.6*	0	1.1	4.1*
	(0.0-0.0)	(0.0-0.0)	(0.5-0.5)	(0.0-1.3)	(0.0-0.0)	(4.6-4.6)	(0.0-0.0)	(0.0-2.0)	(3.2-4.1)
Recurrent taxes on immovable property	0	0	0	0	0	0	0	0.1	0.8*
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)	(0.3-0.8)
Unemployment benefits	0	0.2*	0	0	0	0.2	0	0	0.3
	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)
Environmental taxes	0.7*	0.7	0.2*	0	0	0.7	0.7	0.2	0.2
	(0.0-0.7)	(0.0-0.7)	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)	(0.7-0.7)	(0.0-0.7)	(0.0-0.7)	(0.0-0.2)
Other government in kind consumption	0.6	0.6	2.2*	0	0	0	0	0	0.2
	(0.0-1.0)	(0.0-1.8)	(2.2-2.2)	(0.0-1.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-1.1)	(0.0-0.6)
Sales of goods and services	0	0	0.7*	0	0	1.0*	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.6-0.7)	(0.0-0.1)	(0.0-0.0)	(1.0-1.0)	(0.0-0.0)	(0.0-0.4)	(0.0-0.3)
Consumption taxes (other than environmental)	0	0	1.3*	0	0	2.5	0	0	0
	(0.0-1.3)	(0.0-1.0)	(1.3-1.3)	(0.0-0.0)	(0.0-0.0)	(2.5-2.5)	(0.0-0.0)	(0.0-0.5)	(0.0-0.6)
Public investment	0	0	0.3	0	0	1.1	0	0	0
	(0.0-0.6)	(0.0-0.1)	(0.2-0.3)	(0.0-0.0)	(0.0-0.0)	(1.1-1.1)	(0.0-0.6)	(0.0-0.1)	(0.0-0.0)
Sickness and disability payments	0	0	0.7*	0	0	0	0.2	0	0
	(0.0-0.0)	(0.0-0.0)	(0.7-0.7)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)	(0.0-0.7)	(0.0-0.4)	(0.0-0.2)
Health services provided in kind	0	0	1.5*	0	0	1.5*	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.3-1.5)	(0.0-0.0)	(0.0-0.0)	(1.5-1.5)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0	0.9*	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.0-0.0)	(0.9-0.9)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.8)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0	0	0.3	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.3-0.3)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0.1	0.1	0.8	0	0	0.6	0	0.1	0.1
Used revenue residual	0	0	0.2	0	0	0.6	0.1	0.1	0.4
Share spending efforts	64	45	62	13	100	30	9	25	19
	(35-97)	(23-82)	(62-65)	(13-89)	(100-100)	(29-30)	(9-57)	(25-54)	(17-28)
Achieved consolidation	1.9	2.7	9.2	1.7	0.7	15	1.7	4	7.5
Consolidation needs	1.9	2.7	9.2	1.7	0.7	18.3	1.7	4	7.5
Share top 9 instruments	100	100	46	100	100	48	90	100	100
	(70-100)	(77-100)	(28-62)	(100-100)	(100-100)	(43-64)	(89-100)	(100-100)	(51-100)
Instruments crossing the median	2	2	7	0	0	5	1	2	3
	(1-2)	(2-3)	(6-7)	(0-1)	(0-0)	(5-5)	(1-1)	(1-3)	(3-4)

Note: Simulated interdecile intervals in parentheses computed using 200 random draws. Each dimension of the assessments in Table 2 is increased by 1 or decreased by 1 with equal probability (1/8, 1/8). The use of residual items is marginal. Simulated interdecile intervals are not reported for residual items.

* Denotes an instrument crossing the OECD median.

Table B5. **Instrument use and short- to medium-term consolidation with fixed weights in cluster No. 2**

Percentage of potential GDP except if otherwise mentioned

Description	USA
Subsidies	0
	(0.0-0.0)
Pensions	0.5
	(0.6-0.6)
Other property taxes	0.7*
	(0.7-0.7)
Unemployment benefits	0.2
	(0.2-0.2)
Corporate income taxes	0.2
	(0.2-0.2)
Environmental taxes	0.7
	(0.7-0.7)
Recurrent taxes on immovable property	0
	(0.0-0.0)
Personal income taxes	1
	(1.0-1.0)
Other government in kind consumption (excluding family policy)	0
	(0.0-0.0)
Sales of goods and services	0
	(0.0-0.0)
Sickness and disability payments	0
	(0.0-0.0)
Consumption taxes (other than environmental)	2.5
	(2.1-2.5)
Public investment	0.3
	(0.0-0.3)
Health services provided in kind	1.1*
	(1.1-1.5)
Family	0
	(0.0-0.0)
Social security contributions	0
	(0.0-0.6)
Education	0
	(0.0-0.0)
Used spending residual	0.3
Used revenue residual	0.2
Share spending efforts	31
	(31-39)
Achieved consolidation	7.7
Consolidation needs	7.7
Share top 9 instruments	46
	(36-80)
Instruments crossing the median	2
	(2-2)

Note: See note to Table B4.

Table B6. **Instrument use and short- to medium-term consolidation with fixed weights in cluster No. 3**

Percentage of potential GDP except if otherwise mentioned

Description	ESP	GRC	IRL
Pensions	0	0.6	0
	(0.0-0.0)	(0.6-0.6)	(0.0-0.0)
Subsidies	0.2	0	0
	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)
Other property taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Unemployment benefits	1.6	0	1.1
	(0.6-1.6)	(0.0-0.0)	(1.1-1.1)
Environmental taxes	0.5	0.2	0.3*
	(0.0-0.7)	(0.0-0.2)	(0.3-0.3)
Recurrent taxes on immovable property	0.3	1.0*	1.0*
	(0.0-0.4)	(0.3-1.0)	(0.9-1.0)
Corporate income taxes	0.9*	0	0.3*
	(0.0-0.9)	(0.0-0.0)	(0.3-0.3)
Sales of goods and services	0	0	0.3*
	(0.0-1.0)	(0.0-0.3)	(0.0-0.3)
Personal income taxes	1.7*	3.7*	0.1
	(0.0-2.5)	(2.8-3.7)	(0.1-0.1)
Sickness and disability payments	0	0	0.2
	(0.0-0.7)	(0.0-0.0)	(0.0-0.2)
Other government in kind consumption (excluding family policy)	0	2.1*	0
	(0.0-0.6)	(1.1-2.3)	(0.0-0.0)
Consumption taxes (other than environmental)	0	0	2.2*
	(0.0-1.9)	(0.0-0.7)	(0.6-2.5)
Public investment	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.3)
Health services provided in kind	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-1.5)
Family	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0.1	0.1	0.1
Used revenue residual	0.1	0.4	0.1
Share spending efforts	36	35	26
	(28-60)	(22-41)	(21-32)
Achieved consolidation	5.3	8.2	5.8
Consolidation needs	5.3	8.2	5.8
Share top 9 instruments	100	100	51
	(93-100)	(51-100)	(38-87)
Instruments crossing the median	2	3	5
	(1-2)	(2-3)	(4-5)

Note: See note to Table B4.

Table B7. **Instrument use and short- to medium-term consolidation with fixed weights in cluster No. 4**

Percentage of potential GDP except if otherwise mentioned

Description	AUT	BEL	CZE	FIN	FRA
Pensions	0	0	0.1	0	0.6
	(0.0-0.1)	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.6-0.6)
Subsidies	0.2	0.8	0.8	0.7*	0.7*
	(0.0-0.2)	(0.8-0.8)	(0.8-0.8)	(0.7-0.7)	(0.6-0.6)
Other property taxes	0	0	0.5*	0.4*	0
	(0.0-0.2)	(0.0-0.0)	(0.0-0.6)	(0.4-0.4)	(0.0-0.0)
Unemployment benefits	0	0.7	0	0.9	0.9*
	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)	(0.0-0.9)	(0.8-0.9)
Environmental taxes	0	0	0	0	0.7*
	(0.0-0.0)	(0.0-0.6)	(0.0-0.0)	(0.0-0.0)	(0.0-0.7)
Recurrent taxes on immovable property	0	0	0	0.8*	0
	(0.0-0.0)	(0.0-0.3)	(0.0-0.0)	(0.0-0.8)	(0.0-0.0)
Sales of goods and services	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Sickness and disability payments	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.7)	(0.0-0.0)
Personal income taxes	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Corporate income taxes	0	0	0	0.5*	0.6*
	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.5)	(0.2-0.6)
Other government in kind consumption (excluding family policy)	0	0	0	0.3	1.0*
	(0.0-0.0)	(0.0-0.5)	(0.0-0.0)	(0.0-2.1)	(0.0-1.8)
Consumption taxes (other than environmental)	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.1)	(0.0-1.0)
Health services provided in kind	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.5)
Education	0	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0	0.1	0.1	0.2	0.2
Used revenue residual	0	0	0	0	0
Share spending efforts	100	100	66	55	71
	(0-100)	(56-100)	(61-66)	(55-89)	(71-86)
Achieved consolidation	0.2	1.6	1.6	3.8	4.7
Consolidation needs	0.2	1.6	1.6	3.8	4.7
Share top 9 instruments	100	100	100	100	100
	(100-100)	(100-100)	(100-100)	(91-100)	(78-100)
Instruments crossing the median	0	0	1	4	5
	(0-0)	(0-1)	(1-1)	(2-4)	(4-5)

Note: See note to Table B4.

Table B7. **Instrument use and short- to medium-term consolidation with fixed weights in cluster No. 4** *(cont.)*

Percentage of potential GDP except if otherwise mentioned

Description	HUN	ISL	SVK	SVN
Pensions	0.2	0	0	0
	(0.2-0.2)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Subsidies	0.7*	0.8*	0.5*	0.5*
	(0.7-0.7)	(0.8-0.8)	(0.5-0.5)	(0.5-0.5)
Other property taxes	0.2	0.3	0.7*	0.7*
	(0.0-0.2)	(0.3-0.3)	(0.7-0.7)	(0.4-0.7)
Unemployment benefits	0.2*	0.6	0	0
	(0.0-0.2)	(0.3-0.6)	(0.0-0.0)	(0.0-0.0)
Environmental taxes	0	0.5*	0.3	0
	(0.0-0.0)	(0.0-0.4)	(0.0-0.7)	(0.0-0.0)
Recurrent taxes on immovable property	0.1	0	0.4*	0.2
	(0.0-0.7)	(0.0-0.0)	(0.0-1.0)	(0.0-0.9)
Sales of goods and services	0	0	0	0
	(0.0-0.0)	(0.0-0.3)	(0.0-1.0)	(0.0-0.0)
Sickness and disability payments	0	0.1	0	0
	(0.0-0.0)	(0.0-0.2)	(0.0-0.0)	(0.0-0.0)
Personal income taxes	0.3	0	2	0.5
	(0.0-0.7)	(0.0-0.0)	(0.0-2.8)	(0.0-0.9)
Corporate income taxes	0.1	0.6*	0	0.2
	(0.0-0.7)	(0.0-0.6)	(0.0-0.6)	(0.0-0.9)
Other government in kind consumption (excluding family policy)	0	0.5	0	0
	(0.0-0.6)	(0.1-1.6)	(0.0-0.0)	(0.0-0.9)
Consumption taxes (other than environmental)	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Social security contributions	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0	0
	(0.0-0.0)	(0.0-0.1)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0.1	0.1	0.1	0.1
Used revenue residual	0	0	0	0.1
Share spending efforts	68	61	13	23
	(51-90)	(54-85)	(13-13)	(23-66)
Achieved consolidation	1.8	3.6	4	2.2
Consolidation needs	1.8	3.6	4	2.2
Share top 9 instruments	100	99	100	100
	(100-100)	(83-100)	(96-100)	(100-100)
Instruments crossing the median	2	3	3	2
	(1-2)	(2-3)	(2-4)	(2-3)

Note: See note to Table B4.

Table B8. **Instrument use and short- to medium-term consolidation with fixed weights in cluster No. 5**

Percentage of potential GDP except if otherwise mentioned

Description	LUX	NLD	SWE
Subsidies	0.2	0.6*	0.7*
	(0.2-0.2)	(0.6-0.6)	(0.2-0.8)
Other property taxes	0	0.2	0.2
	(0.0-0.0)	(0.2-0.2)	(0.0-0.7)
Pensions	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.4)
Environmental taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Recurrent taxes on immovable property	0	0.8*	0
	(0.0-0.0)	(0.0-0.8)	(0.0-0.0)
Other government in kind consumption (excluding family policy)	0	1	0
	(0.0-0.0)	(0.0-1.9)	(0.0-0.0)
Sales of goods and services	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Personal income taxes	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Unemployment benefits	0	0.2	0
	(0.0-0.0)	(0.0-0.4)	(0.0-0.0)
Social security contributions	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Sickness and disability payments	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Corporate income taxes	0	0	0
	(0.0-0.0)	(0.0-0.9)	(0.0-0.0)
Consumption taxes (other than environmental)	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Family	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Public investment	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Health services provided in kind	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Education	0	0	0
	(0.0-0.0)	(0.0-0.0)	(0.0-0.0)
Used spending residual	0	0.1	0.1
Used revenue residual	0	0	0
Share spending efforts	100	65	85
	(100-100)	(32-92)	(31-100)
Achieved consolidation	0.2	2.8	1
Consolidation needs	0.2	2.8	1
Share top 9 instruments	100	100	100
	(100-100)	(100-100)	(100-100)
Instruments crossing the median	0	2	1
	(0-0)	(1-3)	(1-2)

Note: See note to Table B4.

Table B9. **Instrument use and short- to medium-term consolidation with additional room for manoeuvre in cluster No. 1**

Percentage of potential GDP except if otherwise mentioned

Description	AUS	CAN	GBR	ISR	ITA	JPN	NZL	POL	PRT
Subsidies	0.8*	0.6	0.2	0.3	0.6	0.2	0	0	0.2
Other property taxes	0.2	0.7*	0.4	0.4	0	0.6*	0.7*	0.7*	0.6*
Pensions	0	0	0.3	0.2	0	0.8	0	0.9	0.9
Corporate income taxes	0	0.6	0.8*	0	0	0	0	0.9*	0.6
Personal income taxes	0.9	0.7	2.5	0.8	0	4.6*	0	1.3	4.6*
Recurrent taxes on immovable property	0	0	0	0	0	0	0.2	0	0.1
Unemployment benefits	0	0	0	0	0	0.4	0	0	0
Environmental taxes	0	0	0.5*	0	0	0.7	0.7	0	0
Other government in kind consumption (excluding family policy)	0	0	2.3*	0	0	0.6	0	0	0
Sales of goods and services	0	0	1.0*	0	0	1.0*	0	0	0
Consumption taxes (other than environmental)	0	0	0.4*	0	0	2.5	0	0	0
Public investment	0	0	0	0	0	1.1	0	0	0
Sickness and disability payments	0	0	0	0	0	0	0	0	0
Health services provided in kind	0	0	0	0	0	1.5*	0	0	0
Social security contributions	0	0	0	0	0	0.9*	0	0	0
Family	0	0	0	0	0	0	0	0	0
Education	0	0	0	0	0	0.4	0	0	0
Used spending residual	0.1	0.1	0.5	0.1	0.1	0.8	0	0.1	0.1
Used revenue residual	0	0	0.3	0	0	0.6	0.1	0.1	0.3
Share spending efforts	46	25	35	31	100	35	0	25	16
Achieved consolidation	1.9	2.7	9.2	1.7	0.7	16.7	1.7	4	7.5
Consolidation needs	1.9	2.7	9.2	1.7	0.7	18.3	1.7	4	7.5
Share top 9 instruments	100	100	83	100	100	52	100	100	100
Instruments crossing the median	1	1	5	0	0	5	1	2	2

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Table B10. **Instrument use and short- to medium-term consolidation with additional room for manoeuvre in cluster No. 2**

Percentage of potential GDP except if otherwise mentioned

Description	USA
Subsidies	0
Pensions	0.5
Other property taxes	0.7*
Unemployment benefits	0.2
Corporate income taxes	0.5
Environmental taxes	0.7
Recurrent taxes on immovable property	0
Personal income taxes	3
Other government in kind consumption (excluding family policy)	0
Sales of goods and services	0
Sickness and disability payments	0
Consumption taxes (other than environmental)	1.7
Public investment	0
Health services provided in kind	0
Family	0
Social security contributions	0
Education	0
Used spending residual	0.1
Used revenue residual	0.2
Share spending efforts	11
Achieved consolidation	7.7
Consolidation needs	7.7
Share top 9 instruments	77
Instruments crossing the median	1

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Table B11. **Instrument use and short- to medium-term consolidation with additional room for manoeuvre in cluster No. 3**

Percentage of potential GDP except if otherwise mentioned

Description	ESP	GRC	IRL
Pensions	0	0.6	0
Subsidies	0.4	0	0
Other property taxes	0.2	0.3	0.2
Unemployment benefits	1.9	0	1.1
Environmental taxes	0.7	0.3	0.6*
Recurrent taxes on immovable property	1	1.0*	1.0*
Corporate income taxes	0.9*	0.4	0.9*
Sales of goods and services	0	0.5	0.7*
Personal income taxes	0	4.6*	1
Sickness and disability payments	0	0	0
Other government in kind consumption (excluding family policy)	0	0	0
Consumption taxes (other than environmental)	0	0	0
Public investment	0	0	0
Health services provided in kind	0	0	0
Social security contributions	0	0	0
Family	0	0	0
Education	0	0	0
Used spending residual	0.1	0.1	0.1
Used revenue residual	0.1	0.3	0.2
Share spending efforts	45	9	22
Achieved consolidation	5.3	8.2	5.8
Consolidation needs	5.3	8.2	5.8
Share top 9 instruments	100	100	100
Instruments crossing the median	1	2	4

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Table B12. **Instrument use and short- to medium-term consolidation with additional room for manoeuvre in cluster No. 4**

Percentage of potential GDP except if otherwise mentioned

Description	AUT	BEL	CZE	FIN	FRA	HUN	ISL	SVK	SVN
Pensions	0.2	0	0.1	0	0.6	0.2	0	0	0
Subsidies	0	0.8	0.8	0.8*	0.8*	0.8*	0.8*	0.7*	0.7*
Other property taxes	0	0	0.6*	0.7*	0	0.4	0.6	0.7*	0.7*
Unemployment benefits	0	0.7	0	0.9	0.9*	0.2	0.6	0	0.2
Environmental taxes	0	0	0	0	0.7*	0	0.7*	0.7	0
Recurrent taxes on immovable property	0	0	0	1.0*	0	0	0.4	1.0*	0.4*
Sales of goods and services	0	0	0	0	0	0	0.2	0.6	0
Sickness and disability payments	0	0	0	0.2	0.2	0	0	0	0
Personal income taxes	0	0	0	0	0	0	0	0	0
Corporate income taxes	0	0	0	0	0.9*	0	0	0	0
Other government in kind consumption (excluding family policy)	0	0	0	0	0.2	0	0	0	0
Consumption taxes (other than environmental)	0	0	0	0	0	0	0	0	0
Social security contributions	0	0	0	0	0	0	0	0	0
Family	0	0	0	0	0	0	0	0	0
Public investment	0	0	0	0	0	0	0	0	0
Health services provided in kind	0	0	0	0	0	0	0	0	0
Education	0	0	0	0	0	0	0	0	0
Used spending residual	0	0.1	0.1	0.2	0.3	0.2	0.1	0.1	0.1
Used revenue residual	0	0	0	0	0	0	0	0.1	0.1
Share spending efforts	100	100	64	56	65	78	45	21	47
Achieved consolidation	0.2	1.6	1.6	3.8	4.7	1.8	3.6	4	2.2
Consolidation needs	0.2	1.6	1.6	3.8	4.7	1.8	3.6	4	2.2
Share top 9 instruments	100	100	100	100	74	100	100	100	100
Instruments crossing the median	0	0	1	3	4	1	2	3	3

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Table B13. **Instrument use and short- to medium-term consolidation with additional room for manoeuvre in cluster No. 5**

Percentage of potential GDP except if otherwise mentioned

Description	LUX	NLD	SWE
Subsidies	0.2	0.8*	0.8*
Other property taxes	0	0.5	0.1
Pensions	0	0	0
Environmental taxes	0	0	0
Recurrent taxes on immovable property	0	1.0*	0
Other government in kind consumption (excluding family policy)	0	0.4	0
Sales of goods and services	0	0	0
Personal income taxes	0	0	0
Unemployment benefits	0	0	0
Social security contributions	0	0	0
Sickness and disability payments	0	0	0
Corporate income taxes	0	0	0
Consumption taxes (other than environmental)	0	0	0
Family	0	0	0
Public investment	0	0	0
Health services provided in kind	0	0	0
Education	0	0	0
Used spending residual	0	0.1	0.1
Used revenue residual	0	0	0
Share spending efforts	100	47	95
Achieved consolidation	0.2	2.8	1
Consolidation needs	0.2	2.8	1
Share top 9 instruments	100	100	100
Instruments crossing the median	0	2	1

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Table B14. **Instrument use and long-term consolidation with additional room for manoeuvre**

Percentage of potential GDP except if otherwise mentioned

Description	JPN	GBR	USA	SVK	AUS	POL	ESP	NZL	ISR	FRA	NLD	IRL	CAN	SVN	CZE
Subsidies	0.2	0.2	0	0.7*	0.8*	0	0.4	0	0.3	0.8*	0.8*	0	0.6	0.7*	0.8
Pensions	3.2	1	2.2	0	0	3.7	0	0	0.8	2.2	0	0	0	0	0.4
Other government in kind consumption (excluding family policy)	0.6	2.3*	0	0.6	1.5	1.5	0.6	0	2.3	1.3*	2.3	0.6	2.3*	2.3*	1.8*
Unemployment benefits	0.4	0	0.2	0	0.2	0	1.4	0	0	0	0.5	0.9	0.3*	0.2	0.3*
Environmental taxes	0.7	0.5*	0.7	0.7	0.7*	0.7	0.7	0.7	0	0	0	0.6*	0.2	0	0
Other property taxes	0.6*	0.4	0.7*	0.7*	0.2	0	0.2	0.7*	0.4	0	0.5	0.2	0	0.1	0
Sickness and disability payments	0	0.7*	0	0.2	0.2	0	0.7*	0.7	0.7*	0	0.1	0.5*	0	0	0
Recurrent taxes on immovable property	0	0	0	1.0*	0.6	0	1	0.2	0	0	0	1.0*	0	0	0
Sales of goods and services	1.0*	1.0*	0	1	0.3	0	0.9	0.3	0.1	0	0	0.3*	0	0	0
Personal income taxes	3.1	1.3	3	1.6	1.6	0	0	0	0	0	0	0	0	0	0
Corporate income taxes	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
Consumption taxes (other than environmental)	0	0	0.3	0	0	0	0	0	0	0	0	0	0	0	0
Public investment	0	0	0	0	0	0	0	1.1*	0	0	0	0	0	0	0
Health services provided in kind	0	0	0	0	0	0	0	1.3	0	0	0	0	0	0	0
Family	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Social security contributions	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Education	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Used spending residual	0.5	0.8	0.5	0.3	0.2	0.3	0.1	0.4	0.4	0.3	0.3	0.2	0.4	0.4	0.3
Used revenue residual	0.1	0.1	0.2	0.2	0	0	0	0.1	0	0	0	0.1	0	0	0
Share spending efforts	47	60	34	25	45	88	53	63	89	100	89	50	96	98	100
Achieved consolidation	10.5	8.5	8.3	7	6.3	6.3	6.1	5.6	5	4.7	4.5	4.4	3.7	3.6	3.6
Consolidation needs	10.5	8.5	8.3	7	6.3	6.3	6.1	5.6	5	4.7	4.5	4.4	3.7	3.6	3.6
Share top 9 instruments	69	82	50	75	73	100	100	52	100	100	100	100	100	100	100
Instruments crossing the median	2	4	1	3	2	0	1	2	1	2	1	4	2	2	2

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Table B14. **Instrument use and long-term consolidation with additional room for manoeuvre** *(cont.)*

Percentage of potential GDP except if otherwise mentioned

Description	PRT	LUX	BEL	HUN	DNK	AUT	CHE	FIN	SWE	DEU	EST	ISL	KOR	GRC
Subsidies	0.2	0.8*	0.8	0.8*	0.8	0.8	0.8	0.8*	0.8*	0.5	0.5	0.8*	0	0
Pensions	3.4	0	0	0.6	1	1.2	0.4	0	0.8*	0.8	0.8	0	0	0.9
Other government in kind consumption (excluding family policy)	0	0	1.9*	0.8	0.3	0	0.7	0.9	0	0	0.2	0.3	0	0
Unemployment benefits	0	0.5*	0	0	0	0	0	0	0	0.1	0	0	0.1	0
Environmental taxes	0	0.5	0	0	0	0	0	0	0	0	0	0	0.3	0
Other property taxes	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sickness and disability payments	0	0.2	0	0	0	0	0	0	0	0	0	0	0	0
Recurrent taxes on immovable property	0	0.4	0	0	0	0	0	0	0	0	0	0	0.5	0
Sales of goods and services	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Personal income taxes	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Corporate income taxes	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Consumption taxes (other than environmental)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Public investment	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Health services provided in kind	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Family	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Social security contributions	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Education	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Used spending residual	0	0.3	0.1	0.2	0.2	0.1	0.1	0.1	0.1	0.2	0	0.1	0	0
Used revenue residual	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Share spending efforts	100	66	100	100	100	100	100	100	100	100	100	100	17	100
Achieved consolidation	3.6	2.8	2.8	2.4	2.3	2.2	2.1	1.8	1.7	1.6	1.6	1.2	1	0.9
Consolidation needs	3.6	2.8	2.8	2.4	2.3	2.2	2.1	1.8	1.7	1.6	1.6	1.2	1	0.9
Share top 9 instruments	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Instruments crossing the median	0	2	1	1	0	0	0	1	2	0	0	1	0	0

Note: The room for manoeuvre is defined by the 20th percentile of OECD countries for spending items and the 80th percentile of OECD countries for revenue items.

* Denotes an instrument crossing the OECD median.

Notes

1. One exception is Norway where such projections are not available. A rough projection of Norwegian real interest rates to 2060 has been built by adding a premium to the projection for the real growth rate in the *OECD Economic Outlook*, May 2013 long-term baseline. This premium is calculated each year between 2015 and 2060 as a weighted average between two growth-interest rate differentials: the last (2014) projected one for Norway and the OECD average in the year under consideration, with the weight on the former gradually decreasing to zero over the period. The income return on Norway's financial assets is projected to 2060 by starting from its average over the last five years of data and making it gradually converge to the average rate of return OECD governments receive on their financial assets in the *OECD Economic Outlook*, May 2013 long-term projections.
2. Formula (D1) is adjusted in the special case of Japan where consolidation is assumed to proceed at a faster pace of 1½ per cent of potential GDP per year to avoid an overly long consolidation period.
3. In the wake of the previous paragraph, one should note that obtaining 2012 estimates of tax expenditures or potential efficiency gains would be virtually impossible.

Appendix references

Alesina, R. and S. Ardagna (2010), "Large Changes in Fiscal Policy: Taxes *versus* Spending", in *Tax Policy and the Economy*, Vol. 24, pp. 35-68, National Bureau of Economic Research.

Arjona, R., M. Ladaique and M. Pearson (2002), "Social Protection and Growth", *OECD Economic Studies*, No. 35, OECD Publishing, Paris.

Auerbach, A. and Y. Gorodnichenko (2012), "Fiscal Multipliers in Recession and Expansion", in *Fiscal Policy after the Financial Crisis*, National Bureau of Economic Research.

Barbiero, O. and B. Cournède (2013), "New Econometric Estimates of Long-Term Growth Effects of Different Areas of Public Spending", *OECD Economics Department Working Papers*, No. 1100.

Bastagli, F., D. Coady and S. Gupta (2012), "Income Inequality and Fiscal Policy", *IMF Staff Discussion Note*, SDN/12/08 (Revised), 27 September, IMF, Washington, DC.

Becker, J., C. Fuest and N. Riedel (2012), "Corporate Tax Effects on the Quality and Quantity of FDI", *European Economic Review*, No. 56.

Bom, P. and J. Ligthart (2013) "What Have We Learned from Three Decades Of Research on the Productivity of Public Capital?", *Journal of Economic Surveys*.

Bova, E., T. Kinda, J. Woo and S. Zhang (2012), "Distributional Consequences of Fiscal Consolidation and the Role of Fiscal Policy: What Do the Data Say?", forthcoming, IMF, Washington, DC.

Calderon, S. (2004), "The Effects of Infrastructure Development on Growth and Income Distribution", *World Bank Policy Research Working Paper Series*, No. 3400.

Chatterjee, S. and S. Turnovsky (2012), "Infrastructure and Inequality", *European Economic Review*, Vol. 56.

Causa, O. and Å. Johansson (2010), "Intergenerational Social Mobility in OECD Countries", *OECD Journal: Economic Studies*, OECD Publishing, Paris.

de la Maisonneuve, C. and J. Oliveira Martins (2013), "Public Spending on Health and Long-Term Care: A New Set of Projections", *OECD Economics Department Working Paper*, forthcoming.

De Mooij, R. (2012), "Tax Biases to Debt Finance: Assessing the Problem, Finding Solutions", *Fiscal Studies*, Vol. 33, No. 4.

Devries, P., J. Guajardo, D. Leigh and A. Pescatori (2011), "A New Action-Based Dataset on Fiscal Consolidation", *IMF Working Paper*, No. 11/128, Washington, DC.

Driscoll, J.C. and A.C. Kraay (1998), "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data", *The Review of Economics and Statistics*, Vol. 80(4), pp. 549-560, MIT Press, November.

Duval, R. (2003), "Retirement Behaviour in OECD Countries: Impact of Old-Age Pension Schemes and Other Social Transfer Programmes", *OECD Economic Studies*, No. 37, OECD Publishing, Paris.

Echevarria, C. (2012), "Income Tax Progressivity, Physical Capital, Aggregate Uncertainty and Long-Run Growth in an Old Economy", *Journal of Macroeconomics*, Vol. 34(4), pp. 955-974.

Ford, R. and W. Suyker (1990), "Industrial Subsidies in the OECD Economies", *OECD Economic Studies*, No. 15, OECD Publishing, Paris.

Fournier, J.-M. and I. Koske (2012), "Less Income Inequality and More Growth – Are they Compatible? Part 7. The Drivers of Labour Earnings Inequality – An Analysis Based on Conditional and Unconditional Quantile Regressions", *OECD Economics Department Working Papers*, No. 930.

Gemmell, N., R. Kneller and Ismael Sanz (2011), "The Timing and Persistence of Fiscal Policy Impacts on Growth: Evidence from OECD Countries", *Economic Journal*, Vol. 121(550), pp. 33-58.

Gonand, F. (2010), "Choosing A Pension Reform: A Framework for the Social Planner", *Pension Reform, Fiscal Policy and Economic Performance*, paper presented at the Banca d'Italia workshop held in Perugia, 26-28 March 2009, Printing Office of the Banca d'Italia, Rome.

Guajardo, J., D. Leigh and A. Pescatori (2011), "Expansionary Austerity: New International Evidence", *IMF Working Papers*, WP/11/158.

Gyimah-Brempong, K. and M. Wilson (2004), "Health Human Capital And Economic Growth In Sub-Saharan African and OECD Countries", *The Quarterly Review of Economics and Finance*, Vol. 44. No. 2.

Hoeller, P., I. Joumard, M. Pisu and D. Bloch (2012), "Less Income Inequality and More Growth: Are They Compatible? Part 1. Mapping Income Inequality Across the OECD", *OECD Economics Department Working Papers*, No. 924.

Johansson, Å, C. Heady, J. Arnold, B. Brys and L. Vartia (2008), "Taxation and Economic Growth", *OECD Economics Department Working Papers*, No. 620.

Johansson, Å., D. Turner, Y. Guillemette, F. Murtin, C. de la Maisonneuve, P. Bagnoli, G. Bousquet and F. Spinelli (2013), "Long-Term Growth Scenarios", *OECD Economics Department Working Papers*, No. 1000.

Joumard, I., P. Hoeller, C. André and C. Nicq (2010), *Health Care Systems: Efficiency and Policy Settings*, OECD Publishing, Paris.

Joumard, I., M. Pisu and D. Bloch (2012), "Less Income Inequality and More Growth – Are They Compatible? Part 3. Income Redistribution via Taxes and Transfers in OECD Countries", *OECD Economics Department Working Papers*, No. 926.

Leduc, S. and D. Wilson (forthcoming), "Roads to Prosperity or Bridges to Nowhere? Theory and Evidence on the Impact of Public Infrastructure Investment", *NBER Macroeconomics Annual 2012*.

OECD (2001), *Competition Policy in Subsidies and State Aid*, OECD Publishing, Paris.

OECD (2003a), *Sources of Economic Growth in OECD Countries*, OECD Publishing, Paris.

OECD (2003b), *Environmentally Harmful Subsidies: Policy Issues and Challenges*, OECD Publishing, Paris.

OECD (2010), *Sickness, Disability and Work: Breaking the Barriers*, OECD Publishing, Paris.

OECD (2010a), *Tax Expenditures in OECD Countries*, OECD Publishing, Paris.

OECD (2010b), "Health Care Systems: Getting More Value for Money", *OECD Economics Department Policy Notes*, No. 2, OECD Publishing, Paris.

OECD (2011), *Divided We Stand*, OECD Publishing, Paris.

OECD (2011a), *OECD Economic Surveys: Switzerland 2011*, OECD Publishing, Paris.

OECD (2012a), "Looking to 2060: Long-Term Global Growth Prospects – A Going for Growth Report", *OECD Economic Policy Papers*, No. 3, OECD Publishing, Paris.

OECD (2012b), "Special Feature: Trends in Personal Income Tax and Employee Social Security Contribution Schedules", *Taxing Wages 2011*, OECD Publishing, Paris.

OECD (2013), *OECD Economic Outlook*, Vol. 2013/1, No. 93, May, OECD Publishing, Paris.

Rawdanowicz, L., E. Wurzel and A.K. Christensen (2012), "The Equity Implications of Fiscal Consolidation", *OECD Economics Department Working Papers*, No. 1013.

Serret, Y. and N Johnstone (2006) (ed.), *The Distributional Effects of Environmental Policy*, OECD Publishing, Paris.

Sutherland, D., R. Price, I. Joumard and C. Nicq (2007), "Performance and Indicators for Public Spending Efficiency in Primary and Secondary Education", *OECD Economics Department Working Papers*, No. 546.

Sutherland, D., S. Araujo, B. Egert and T. Kozluk (2009), "Infrastructure Investment: Links to Growth and The Role of Public Policies", *OECD Economics Department Working Papers*, No. 686.

Sutherland, D., P. Hoeller, R. Merola and V. Ziemann (2012), "Debt and Macroeconomic Stability", *OECD Economics Department Working Papers,* No. 1003.

OECD Journal: Economic Studies
Volume 2013
© OECD 2014

Choosing the pace of fiscal consolidation

by
Łukasz Rawdanowicz*

In many OECD countries debt has soared to levels threatening fiscal sustainability, necessitating its reduction over the medium to longer term. This paper proposes a stylised model, featuring endogenous interactions between fiscal policy, growth and financial markets, to highlight how economic shocks and structural features of an economy can affect consolidation strategy and resulting growth and inflation developments. The fiscal authorities are assumed to choose a consolidation path from a predetermined set of possible paths by maximising cumulative GDP growth and minimising cumulative squared output gaps, with the objective to reach a given debt-to-GDP level within a finite horizon and stabilise debt afterwards under the assumption of the unchanged fiscal policy stance. Illustrative simulations for a hypothetical economy show, among other things, that by requiring debt to stabilise part of the initial adjustment can be reversed; some stepping up of the fiscal adjustment can be optimal if bond yields increase due to an exogenous shock; and for some debt reduction targets, high fiscal multipliers, hysteresis effects and higher government bond yields imply protracted deflation and large negative output gaps, stressing the need to select reasonable fiscal targets consistent with market conditions.

JEL classification: E61, E62, H6

Keywords: Fiscal consolidation, sovereign debt, government budget balance, fiscal rules

* Łukasz Rawdanowicz (e-mail: lukasz.rawdanowicz@oecd.org) works at the OECD Economics Department. The author would like to thank Rudiger Ahrend, Sveinbjörn Blöndal, Luc Eyraud, Yvan Guillemette, Patrice Ollivaud, Jean-Luc Schneider, Stéphane Sorbe, Douglas Sutherland, David Turner and Eckhard Wurzel for valuable comments, and Lyn Urmston for help in the final document preparation. The views expressed in this paper are those of the author and do not necessarily represent those of the OECD or its member countries.

Government deficits and debt increased in many OECD countries over recent years, often to record high levels. Such positions are not sustainable and fiscal consolidation has become a necessity. Merely stabilising debt at current high levels requires a substantial adjustment and even a larger and more protracted one is needed to reduce debt to lower, prudent levels (OECD, 2012). The latter option is desirable as high debt may weigh negatively on output growth,[1] expose a sovereign to market sentiment shifts, limit room for accommodation of future negative shocks and add to fiscal challenges resulting from the expected increase in government spending related to population ageing and health care.

While there is generally little controversy about the need of fiscal consolidation, its optimal pace is hotly debated (Corsetti, 2012), posing a key dilemma for policymakers in many OECD countries. Some argue for postponing consolidation as a large, frontloaded adjustment can reduce GDP growth with negative fallout for the fiscal situation. Such effects are more likely when – as is currently the case for several OECD countries – output and unemployment gaps are large, effectiveness of monetary policy is reduced with interest rates close to the zero bound, and credit constrains become more binding (Gali et al., 2007; Christiano et al., 2009; Woodford, 2011; Auerbach and Gorodnichenko, 2012). Consolidation-induced slowdowns may also reduce potential GDP via hysteresis effects and thus be counterproductive from a longer-term point of view (De Long and Summers, 2012). On the other hand, postponing consolidation may undermine market confidence in the solvency of a government, risking, in a benign scenario, sluggish growth due to the pass-through of higher sovereign risk to borrowing costs for the private sector and, in the worst case scenario, disruptive sovereign default.

The choice of optimal consolidation path depends crucially on the ultimate long-term objective of fiscal policy and market conditions. Estimating optimal consolidation pace is challenging given the nexus of interactions between fiscal policy, financial markets and economic growth. These interactions are likely to be non-linear and give rise to multiple equilibria. Consequently, their quantification is highly uncertain.

To shed some light on determinants of consolidation and their likely quantitative impact this paper uses stylised simulations of a small, calibrated model for a hypothetical economy where the fiscal adjustment is determined via an optimisation of a government loss function given the assumed objective of stabilising debt at the target at a finite horizon. The model accounts for endogenous feedback among fiscal policy, growth and financial markets as well as a simple monetary policy rule and debt maturity structure under a perfect foresight. The hypothetical economy is calibrated so as to reflect key characteristics of fiscally troubled OECD countries in 2011, including high debt, structural and headline budget deficits and initial large negative output gaps. The exercise, although only illustrative given its stylised nature, involves a number of sensitivity tests with respect to key parameters of the model and alternative scenarios. This is necessary given large uncertainty about interactions between fiscal, real and financial variables. The simulations investigate in particular how optimal consolidation changes with interest rates shocks,

different fiscal multipliers, hysteresis effects and under delayed consolidation. In addition, conditions for a self-defeating fiscal adjustment are investigated. The model simulations are not meant to prescribe concrete policy recommendations – in particular regarding debt targets and the timing to reach them – but rather to offer a framework, that can be tailored to country-specific circumstances, to inform general policy choices.

1. Model's framework

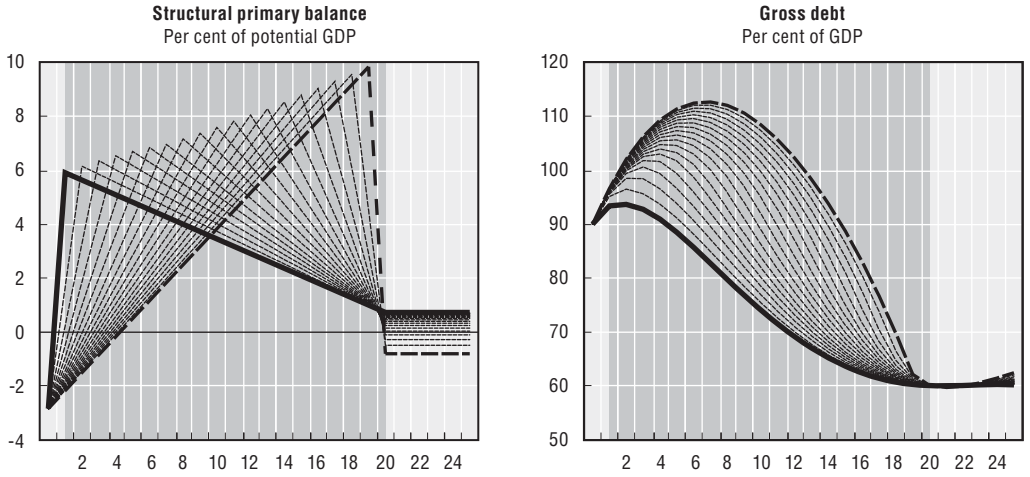
1.1. Consolidation objective

Given an initial high debt level, it is assumed that fiscal authorities are fully determined to reduce the debt-to-GDP ratio to a lower and prudent level within a finite horizon and then stabilise it at this level indefinitely with an unchanged fiscal policy stance. Debt stabilisation – around what is deemed a sustainable and prudent level – can be desirable as it creates room for absorbing future shocks and at the same time allows for a less stringent fiscal policy stance (see below). It also imposes a longer-term perspective, beyond the analysed period, requiring no further fiscal adjustment with unchanged fundamentals.

The requirement of debt stabilisation has implications for calculating consolidation needs. As explained in Annex A in greater detail, approaches adopted in the literature (assuming either a one-off adjustment in the first year or a gradual and equal in size adjustment over the entire period)[2] to reach a given debt target within a finite horizon, imply that with unchanged fiscal policy beyond the analysed horizon the debt-to-GDP ratio would continue to fall. This stems from the fact that with high initial debt and deficit ratios, a large improvement in the budget balance is needed to reduce debt within a given period, but debt stabilisation, once the lower debt level is achieved, requires a lower budget balance. Thus, consolidation needs derived from the methods employed in the literature imply that consolidation is permanent and they ignore debt evolution after the end of the period analysed. In contrast, consolidation aimed at debt stabilisation implies that some of the initial adjustment (i.e. a difference between the peak surplus and initial budget balances) can be reversed and total consolidation (i.e. a difference between the end and initial budget balances) will be lower (Annex A). There are numerous possible consolidation paths that are consistent with stabilising the debt ratio at a lower level within a finite horizon. The simulations in this paper focus on a subset of consolidation paths that involves a constant annual adjustment in the structural primary balance, both during fiscal tightening and easing, and assume an immediate start of consolidation. This implies a kinky, linear path of the structural primary balance (Figure 1). For this set of functions, there are $k - 1$ possible paths, where k is the number of years to reach the debt target, which differ with respect to the time of attaining the peak and the size of annual adjustment (i.e. slopes of the lines). Annex A proposes two alternative paths that ensure debt stabilisation.

Given the set of consolidation paths above, an optimal path can be selected by optimising a fiscal policy objective function. Several plausible objectives are possible.[3] In this paper, since the assumed objective is to stabilise debt at the target within a finite period, minimising the impact of consolidation on economic growth seems a natural criterion. Such an objective is appealing, especially when the effectiveness of monetary policy is reduced and there are large negative output gaps. The pursuit of growth objectives should not, however, lead to undesirable large output gaps. Thus, it is assumed that the government maximises a discounted sum of real GDP growth and minimises a discounted sum of squared output gaps (implying that positive and negative gaps are treated

Figure 1. **Possible consolidation paths to stabilise debt within a finite horizon**



Note: Different levels of the structural primary balance after reaching the debt target (non-shaded area) stem from differences in nominal GDP growth and in interest rates at the end of the simulation period as a result of endogenous feedback assumed in the baseline specification of the model (see Section 1.2).

Source: Author's calculations.

symmetrically).[4] Both components are summed from the start of the simulations until four years after the end of simulations. The extension of the sample is meant to capture possible large positive growth rates/output gaps resulting from fiscal easing at the end of the simulation period. The government's loss function is given by:

$$LF = min \sum_{i=1}^{k+4} \beta^i \left(-w1 * r_i + w2 * gap_i^2\right) \quad [1]$$

where $w1$ and $w2$ are weights attached to the first and second component, r is real GDP growth rate, gap is the output gap, and β is the time discount factor based on the 1% discount rate. The function ignores other possible important objectives of fiscal policy, like income redistribution or promoting potential growth.

Fiscal consolidation is defined as a change in the structural primary budget balance. The structural balance is corrected for cyclical effects and thus tends to capture discretionary fiscal measures,[5] and the primary balance eliminates effects of changes in interest rates on net costs of debt servicing, which may not always reflect government discretionary decisions.

1.2. The model

The fiscal simulations are based on a stylised, calibrated, small macroeconomic model. It consists of four main blocks determining real GDP growth, inflation, monetary policy and long-term interest rates. It also contains standard accounting identities for the budget balance and debt, and simplified calculation of structural (cyclically-adjusted) budget balances.[6] The description of the main equations focuses on baseline specifications – alternative scenarios are discussed in the next section.

1.2.1. Real GDP growth

Following the assumption adopted in Johansson et al. (2013), the cyclical component of real GDP growth is driven by the closure of the output gap. This implies that with a negative output gap, GDP grows faster than potential GDP, and when the output gap is closed it grows at the potential rate (equation [2]). The closing of the output gap is based on

an error-correction specification with an elasticity of –0.2, implying that, all other things being equal, an output gap of 5% will be largely closed after 12 years. This relatively slow pace of convergence is selected to reflect the fact that financial crises can be followed by a prolonged period of slow economic growth (Reinhart and Rogoff, 2009). Potential real GDP growth is assumed exogenous in the baseline (this assumption is relaxed in a scenario with hysteresis – Section 2.4). In addition, real output growth is affected by fiscal policy via a short-term fiscal multiplier (0.5 in the baseline scenario), and by changes in real long-term interest rates with a short-term elasticity of –0.3. The latter draws on the long-term elasticity estimated by Cournède (2010).[7] The full GDP equation is given by:

$$dlog(GDPV) = dlog(GDPVTR) - 0.2 * log(GDPV[-1]/GDPVTR[-1]) + FM * d(NLGXQU) - 0.3 * d(RIRL)/100 \quad [2]$$

where *GDPV* and *GDPVTR* are real actual and potential GDP, *d(NLGXQU)* stands for the change in the structural primary balance as a per cent of GDP, *FM* is the fiscal multiplier and *d(RIRL)* stands for changes in the real long-term interest rate.

1.2.2. *Inflation*

Inflation, defined in terms of the GDP deflator, is modelled as an inflation expectations augmented Philips curve based on the output gap:

$$pi = 0.3 * pi(-1) + 0.7 * infl\_target + 0.2 * GAP \quad [3]$$

where *pi* is the rate of inflation and *GAP* is the output gap. Inflations expectations are assumed to be a weighted average of the past inflation rate and the inflation target (*infl_target*) which is set at 2%.

1.2.3. *Monetary policy*

The monetary policy stance is driven by the standard Taylor rule.[8] A target short-term interest rate (*IRS_ss*) is set in response to deviations of actual inflation (*pi*) from the target (*infl_target*) and the output gap (*GAP*), and the natural (nominal) short-term interest rate is assumed to be 4%:[9]

$$IRS\_ss = 4 + 0.5 * GAP + 1.5 * (pi - infl\_target) \quad [4]$$

Monetary policy is characterised by interest rate smoothing. Consequently, the actual policy interest rate adjusts only gradually to the Taylor rate:

$$IRS = 0.5 * IRS\_ss + 0.5 * IRS(-1) \quad [5]$$

It is assumed that the short-term interest rate cannot fall below zero.

1.2.4. *Long-term interest rates and implied cost of debt*

Following the assumptions in Johansson et al. (2013), the long-term interest rate is modelled as a 10-year average of future short-term policy rates (under a perfect foresight), a term premium (*Tprem*) fixed at 0.7%, and a fiscal risk premium (*Frisk*):

$$\mathrm{IRL} = \mathrm{Tprem} + \sum_{i=0}^{9} \mathrm{IRS}_i/10 + \mathrm{Frisk}_i \quad [6]$$

The fiscal risk premium depends on gross debt. It increases by 0.02 percentage point for each percentage point of the excess of gross debt over 75% of GDP and by additional 0.02 percentage point when gross debt exceeds 125% of GDP.

The implied cost of debt (*IRP*) accounts, in a simplified way, for the maturity structure of outstanding debt. It is assumed that the cost of debt is a weighted average of the last period cost of debt and a weighted average of current market interest rates:

$$IRP = (1 - RFSH) * IRP(-1) + RFSH * (0.25 * IRS + 0.75 * IRL) \quad [7]$$

The weights (*RFSH*) are the shares of debt which matures within one year. In the baseline scenario, *RFSH* is set at 0.2, which is close to the 2011 average observed for selected OECD countries.[10]

1.3. Hypothetical economy

The model – as outlined above – is solved for a hypothetical economy which features many characteristics of OECD countries in difficult fiscal positions in 2011. In particular, the initial conditions involve the high overall budget deficit and gross debt, and the large negative output gap (Table 1).[11] In addition, short and long-term interest and inflation rates are assumed to be initially low (only in the baseline scenario). In an illustrative scenario, the policy objective is to reduce debt from 90% to 60% of GDP within 20 years and to ensure its stabilisation at 60% of GDP. The target is set in the vicinity of a level conventionally deemed as safe (Sutherland et al., 2012), and the length and the size of debt reductions are within the post-World War II historic averages of past big debt reductions in the OECD countries (Figure 2).[12] For the baseline simulation, weights of 0.2 and 0.8 in the government loss function are selected.

Table 1. **Main characteristics of the hypothetical economy**

Initial conditions		Constant assumptions over the entire simulation period	
Budget balance	–8.5% of GDP	Potential real GDP growth	2%
Primary budget balance	–5.5% of GDP	Government financial assets	30% of GDP
Structural budget balance	–5.7% of GDPTR	Interest rate earned on financial assets	2%
Structural primary budget balance	–2.9% of GDPTR		
Gross debt	90% of GDP		
Inflation rate	1.5%		
Short-term interest rate	0.5%		
Long-term interest rate	2%		
Output gap	–5% of GDPTR		

Note: GDPTR is nominal potential GDP.
Source: Author's calculations.

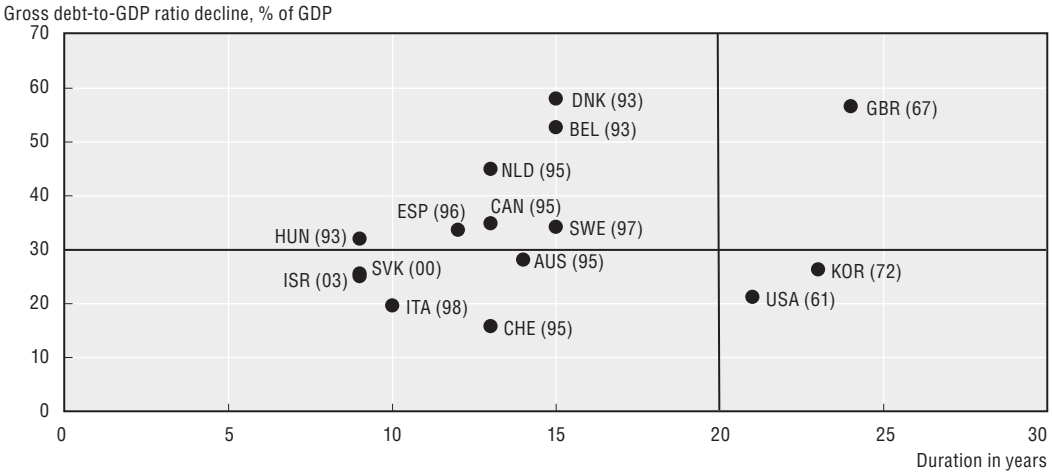
2. Simulation results

2.1. Baseline consolidation path

The parameters of the government loss function in the baseline scenario imply that initial consolidation is gradual and protracted, lasting 16 years (Table 2 and Figure 3). This stems from the dominance of the GDP growth maximisation objective since the paths with large fiscal easing at the end of the simulation period result in very high GDP growth rates (Section 2.2). The long duration of the consolidation phase makes the initial fiscal adjustment sizeable, nearly 10% of GDP. With shorter consolidation, the maximum adjustment would be smaller (Figure 1). As discussed above, initial consolidation is significantly higher than total consolidation of only around 3% of GDP, as it is followed by four years of fiscal easing. The baseline simulation demonstrates that, even with benign assumptions and a rather long period to stabilise debt, high budget deficit and debt make

a significant reduction in debt challenging. Thus, unless helped by stronger GDP growth and lower interest rates, the process would be very difficult politically.

Figure 2. **Past debt reductions**



Note: Numbers in parentheses after the country abbreviation refer to the year when debt reduction started.
Source: OECD Economic Outlook Database No. 91.

Table 2. **Summary of different consolidation scenarios**

	Scenario	Model assumptions							Initial consolidation		Fiscal easing		Structural budget balance		
									Length	Size	Length	Size	Initial	Max.	End
		W1	W2	FM	RFSH	IR	Hyst.	AS	Years	% of GDP	Years	% of GDP	% of GDP		
1	Baseline	0.2	0.8	0.5	0.2	0	0.0	0.5	16	10.1	4	-6.9	-2.8	7.2	0.3
2	Alternative weights 1	1.0	0.0	0.5	0.2	0	0.0	0.5	19	10.7	1	-8.2	-2.8	7.9	-0.3
3	Alternative weights 2	0.0	1.0	0.5	0.2	0	0.0	0.5	14	9.6	6	-6.3	-2.8	6.8	0.5
4	Interest rate shock	0.2	0.8	0.5	0.2	3	0.0	0.5	11	10.8	9	-7.2	-2.8	7.9	0.7
5	Interest rate shock with higher debt turnover	0.2	0.8	0.5	0.4	3	0.0	0.5	11	10.8	9	-7.0	-2.8	8.0	1.0
6	Hysteresis	0.2	0.8	0.5	0.2	3	0.0	0.5	16	10.4	4	-7.3	-2.8	7.5	0.2
7	Lower automatic stabilisers	0.2	0.8	0.5	0.2	0	0.1	0.5	17	11.1	3	-6.9	-3.8	7.3	0.4
8	Higher fiscal multiplier	0.2	0.8	1.5	0.2	0	0.0	0.3	17	11.3	3	-8.7	-2.8	8.5	-0.3
9	Higher fiscal multiplier with hysteresis	0.2	0.8	1.5	0.2	0	0.1	0.5	17	12.4	3	-9.9	-2.8	9.5	-0.4
10	Higher fiscal multiplier with hysteresis and interest rate shock	0.2	0.8	1.5	0.2	3	0.1	0.5	16	17.0	4	-14.6	-2.8	14.2	-0.4
11	Delayed consolidation under baseline	0.2	0.8	0.5	0.2	0	0.0	0.5	14	10.4	6	- 7.1	-2.9	7.6	0.4
12	Delayed consolidation with interest rate shock	0.2	0.8	0.5	0.2	3	0.0	0.5	9	11.5	11	-7.8	-2.9	8.6	0.8
13	Delayed consolidation with higher fiscal multiplier, hysteresis and interest rate shock	0.2	0.8	1.5	0.2	3	0.1	0.5	16	19.4	4	-17.1	-2.9	16.5	-0.6

Note: *W1/W2* are weights of the government loss function, *FM* is the size of fiscal multiplier, *RFHS* is the share of government debt that matures within one year, *IR* indicates the size of interest rate shock, *Hyst.* is the size of the hysteresis effect, and *AS* is the size of automatic stabilisers.
Source: Author's calculations.

2.2. Alternative weights in the loss function

The choice of weights of the loss function affects the consolidation path (Table 2 – scenarios 1-3 – and Table 3). Under the baseline specification of the model and initial conditions (Table 1), maximising cumulative GDP growth always favours a more gradual

Table 3. **Timing of the budget balance peak for different weights of the loss function**

In years

	W2				
	0.00	0.20	0.50	0.80	1.00
W1 = 1 – W2	19	19	19	16	14

Note: The loss function is: $LF = min \sum_{i=1}^{k+4} \beta^i(-w1 * r_i + w2 * gap_i^2)$.

Source: Author's calculations.

and longer-lasting consolidation (i.e. a back-loaded consolidation). This is primarily due to very high GDP growth rate at the end of the simulation period, resulting from extremely large fiscal easing when implemented within one or two years (Figure 3).[13] In contrast, minimising the sum of squared output gaps tends to favour the path with the peak closer to the middle of the sample as large initial consolidation and large fiscal loosening at the end imply large output gaps, which are penalised. Consequently, increasing the weight of the first component prolongs consolidation, whereas raising the weight of the second component shortens it to a minimum of 14 years (Table 3).

Given the model's assumptions and initial conditions, a smaller annual adjustment implies a longer and larger cumulative fiscal adjustment (Figure 3). Back-loaded consolidation results in higher cumulative growth, but the closing of the output gap is very sluggish and there is a big change in the output gap at the end of the sample. In addition, the initial increase in gross debt is more pronounced and protracted (Section 2.8).

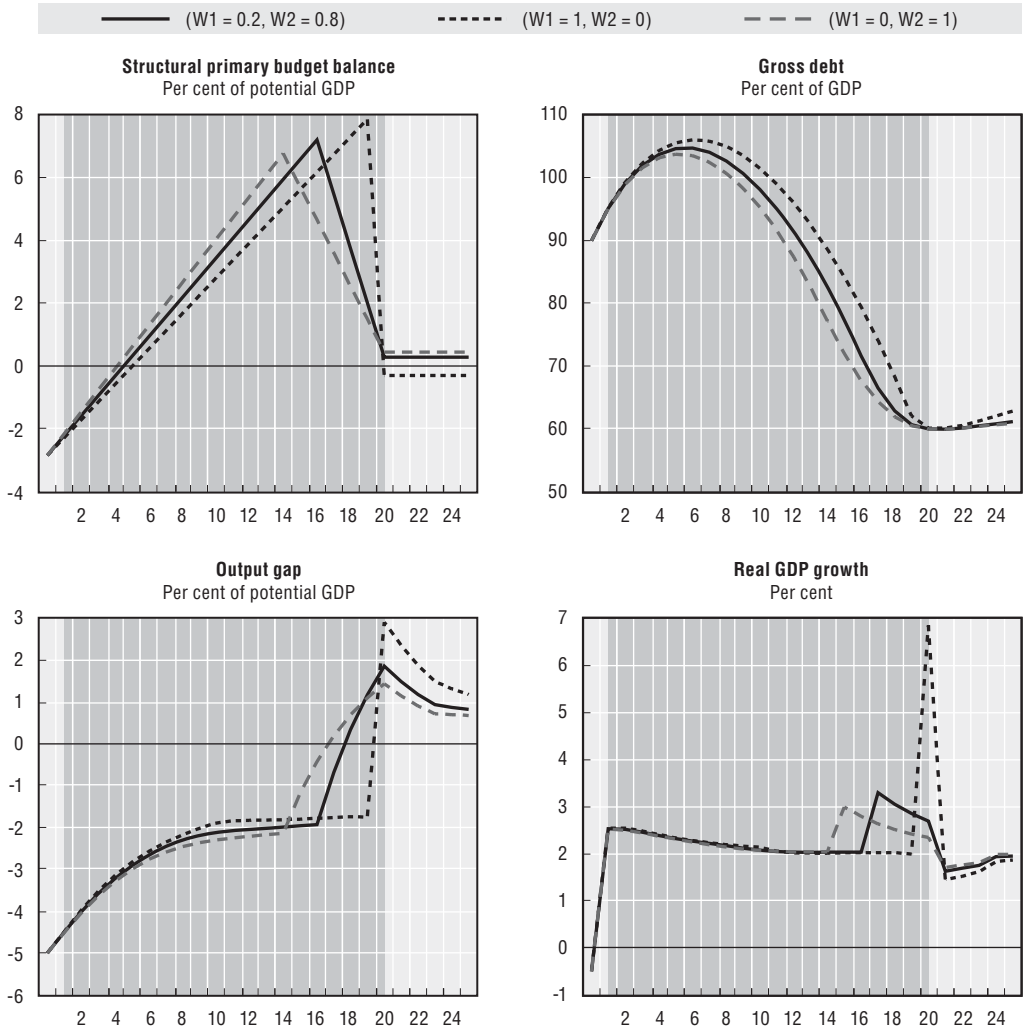
2.3. Interest rate shocks

Contrary to the baseline assumption, some OECD countries in difficult fiscal situations are faced with high market long-term interest rates. This section investigates how a long-term interest rate shock affects fiscal consolidation strategy.

Modelling market reactions is difficult. The model arbitrarily assumes a 3 percentage point increase in long-term interest rates. The shock persists at this level as long as the lagged overall budget deficit is above 6% of GDP and then it declines linearly to zero in line with narrowing of the deficit (it is zero when the budget is balanced).[14] The interest rate shock is transmitted to the implied cost of debt via long-term interest rates as explained in equation [7]. Monetary policy short-term interest rates are not affected by the interest rate shock.

The simulation results suggest that a country faced with adverse and protracted market reactions may find it optimal to shorten and steepen fiscal consolidation compared with the baseline (i.e. opt for front-loaded consolidation). Cumulative consolidation under the interest shock is marginally higher than in the baseline, but it is shorter by five years and the annual adjustment almost doubles (Table 2, scenario 4; Figure A.3). The interest rate shock necessitates larger annual consolidation for all possible paths due to higher debt servicing costs and, initially, a bigger cyclical deficit due to lower GDP growth resulting from higher interest rates and the larger fiscal adjustment (equation [2]). Given the criterion of minimising squared output gaps, very front-loaded and back-loaded consolidation strategies are increasingly penalised, leading to a selection of the structural primary budget balance path with the peak closer to the middle of the sample. Doubling the share of debt which matures within a year (RFSH) from 0.2 to 0.4 changes little the

Figure 3. **Consolidation under different weights of the loss function**



Note: The loss function is: $LF = min \sum_{i=1}^{k+4} \beta^i(-w1 * r_i + w2 * gap_i^2)$.
Source: Author's calculations.

consolidation path. The interest rate shock does not change much the level of real output and of prices at the end of simulation sample (Section 2.9).

It is worth noting that market interest rates pass through to the actual cost of debt with a lag. Even with a higher debt turnover the initial interest rate shock is not fully passed to the cost of debt and this takes around five years (Figure A.3).[15] To some extent this stems from endogenous decline in the short-term interest rate in the context of subdued growth and inflation. This implies that governments may have some time to take policy action when faced with adverse market reactions before debt financing becomes unsustainable. However, the interest rate shock which is limited by assumption to 3 percentage points and this may be viewed as a benign assumption given the observed sovereign risk premia in countries such as Greece and Portugal, and some evidence that the premia can increase nonlinearly with rising fiscal deficits and debt and negative real GDP growth (Haugh et al., 2009; and Cottarelli and Jaramillo, 2012).

2.4. Hysteresis

Prolonged negative output gaps can impact negatively the level of potential GDP, for instance, by discouraging labour participation or by depreciating human capital – the so-called hysteresis effects. Following De Long and Summers (2012), it is assumed that each percentage point of the (lagged) output gap reduces the growth rate of potential GDP by 0.1 percentage point (i.e. the level of potential GDP is permanently reduced). The hysteresis effect is assumed symmetric, i.e. positive output gaps increase the level of potential output. Under the assumed parameterisation, the presence of hysteresis does not change significantly the optimal consolidation path compared with the baseline (Table 2, scenario 6), but when it can affect the adjustment path when interacted with higher fiscal multipliers and interest rate shocks and have a considerable and permanent impact on the GDP level – see Section 2.9.

2.5. Automatic stabilisers

The presence of automatic stabilisers complicates consolidation as part of the structural adjustment is offset by the deterioration of fiscal balances due to the cycle. The size of automatic stabilisers also impacts the estimate of the initial structural balance. For a given overall budget deficit and a negative output gap, higher automatic stabilisers imply a higher structural balance. The opposite is true for lower values. Since the baseline assumption about the size of automatic stabilisers (0.5) corresponds to the upper range of estimates for the OECD countries observed mainly in western, continental European countries (Girouard and André, 2005), a sensitivity test assumes a lower value of 0.3 (which equals the estimates for Japan and the United States).

It turns out that lowering automatic stabilisers has a very small impact on the optimal consolidation path (Table 2, scenario 7). Total and initial consolidations are slightly larger than in the baseline since the initial level of the structural primary deficit is higher with lower automatic stabilisers.

2.6. Fiscal multipliers

The size of fiscal multipliers is at the centre of discussions about the pace of fiscal consolidation. The fiscal multiplier assumed in the baseline scenario of 0.5 is around the average of what is viewed as normal aggregate multipliers in the OECD countries (IMF, 2010; Barrell et al., 2012). However, there are several arguments that in current circumstances fiscal multipliers can be much higher. For instance, the multipliers tend to be larger during recessions as expansionary spending is less likely to crowd out private demand (Auerbach and Gorodnichenko, 2012), when monetary policy is constrained by the zero bound (Christiano et al., 2009; Woodford, 2011), or when consumers have no access to credit and are constrained to consume only out of their current income (Gali et al., 2007). Thus, this section investigates an alternative scenario where fiscal multiplier is increased to 1.5 as long as the output gap is below –3% of GDP. The latter condition is added to avoid a situation that fiscal easing excessively stimulates the economy when the output gap is positive and monetary policy functions normally (i.e. monetary policy interest rates are significantly above zero).

The higher fiscal multiplier does not change much fiscal consolidation strategy in terms of the duration and size of the fiscal effort, with initial consolidation longer by one year and slightly larger than in the baseline. This reflects the fixed deadline for reaching

the debt target and the choice of loss function. However, the GDP cost of such consolidation is very high, resulting in large and persistent negative output gaps and in inflation and monetary policy rates close to zero for an extended period (Section 2.9).

2.7. Delayed consolidation

Given possible large negative consequences of consolidation on growth and weaker recovery, postponing consolidation (or even pursuing fiscal stimulus) has been suggested. To illustrate potential consequences of such a strategy, three scenarios (the baseline, interest rate shock and combined higher fiscal multiplier, hysteresis and the interest rate shock) are re-run assuming that the start of consolidation and the timing of reaching the debt target are delayed for two years. This implies that the structural primary balance is left unchanged for two initial years and consolidation starts only in the third year. The deadline to stabilise debt is kept unchanged at 20 years, so the total simulation sample is 22 years. Total consolidation is slightly higher with a shorter duration and a larger annual adjustment than in the respective scenarios with an immediate start of the adjustment (Table 2, scenarios 11-13; and Figure A.5). The build-up of debt is however larger (Section 2.8).

2.8. Self-defeating consolidation

Recent announcements of consolidation plans in the OECD countries caused concerns that fiscal adjustments may actually increase government budget deficits and debt (i.e. may result in a self-defeating consolidation). This may occur, for instance, when increased tax rates reduce economic activity and in turn the tax base, which – in contrast to the intentions – diminishes tax revenues and in turn – via higher deficit – increases debt.

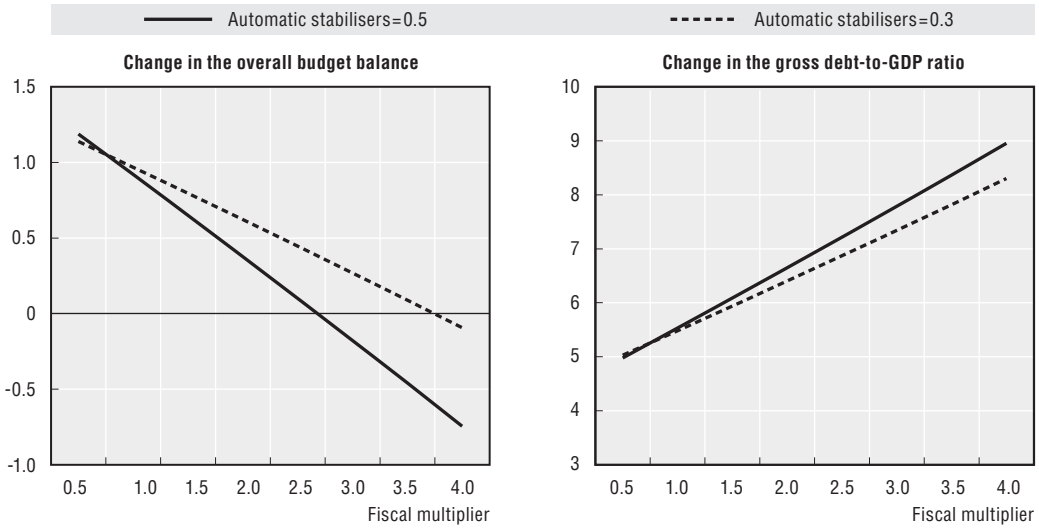
In the logic of the model presented, self-defeating consolidation arises when the improvement in the structural primary balance lowers GDP growth (via the fiscal multiplier), which in turn deteriorates the cyclical component of the budget balance (via automatic stabilisers) and results in no improvement or even a deterioration in the overall budget balance. This effect depends on the size of the fiscal multiplier and of automatic stabilisers.[16] The higher the fiscal multiplier for a given size of automatic stabilisers, the bigger the effect. Lowering the automatic stabilisers' parameter weakens this relation. Model simulations under the baseline specification[17] show that for reasonable ranges of automatic stabilisers and fiscal multipliers, a structural fiscal adjustment of 1% of GDP is not likely to lead to a decline in the overall budget balance in the first year (Figure 4).

The impact of structural consolidation on the debt-to-GDP ratio depends on the erosion in the ratio by nominal GDP growth and the resulting budget balance.[18] For a given structural adjustment, a higher fiscal multiplier dampens GDP growth by more and thus the gross debt ratio increase is larger (Figure 4).

Table 4 summarises changes in the overall budget balance, the primary structural balance and gross debt (all as ratios to GDP) for selected scenarios which differ largely from the baseline (detailed results are in Table A.2). In three scenarios involving the interest rate shock, especially when combined with the higher fiscal multiplier, the increase in the overall budget balance is indeed smaller than the structural improvement during the first three years, at most by half. This reflects raising net interest payments and the widening of the output gap (i.e. the increase in the cyclical deficit). In none of the scenarios, there is a full or negative offset though.[19]

Figure 4. **First-year change in the budget balance and debt following 1% of GDP structural consolidation**

Per cent of GDP



Note: Calculations are based on model simulations with the initial conditions as assumed in Table 1, the assumption that inflation and interest rates remain unchanged and GDP growth is determined as in equation [2].
Source: Author's calculations.

Table 4. **Changes in the structural primary balance, overall balance and debt in selected scenarios**

	Scenario	Cumulative change during first 3 years of consolidation in:		Gross debt	
		Structural primary balance	Overall balance	Maximum	Reached in year
		% of potential GDP	% of GDP	% of GDP	
1	Baseline	1.9	3.4	105	6
4	Interest rate shock	2.9	2.7	109	6
5	Interest rate shock with higher debt turnover	3.0	2.6	109	6
8	Higher fiscal multiplier	2.0	3.2	108	6
9	Higher fiscal multiplier with hysteresis	2.2	3.2	111	7
10	Higher fiscal multiplier with hysteresis and interest rate shock	3.2	1.6	126	8
11	Delayed consolidation under baseline	2.2	2.7	110	7
12	Delayed consolidation with interest rate shock	3.8	3.1	116	7
13	Delayed consolidation with higher fiscal multiplier, hysteresis and interest rate shock	3.6	1.5	137	10

Note: For specifications of scenarios, see Table 2.
Source: Author's calculations.

In case of the immediate start of consolidation, the debt-to-GDP ratio begins to decline only after seven or eight years and in most scenarios gross debt increases by around 15-20% of GDP, with the exception of the scenario involving the higher fiscal multiplier, hysteresis and the interest rate shock when debt increases initially by 36% of GDP. The continued increase in debt after the start of the consolidation was observed in the past consolidation episodes in the OECD countries (Blöchliger et al., 2012). Delaying the start of consolidation results in a higher and longer debt build-up, in the worst case by up to 47% of GDP, lasting more than a decade.

By and large, the tested specifications of the parameters suggest that a fully-fledged self-defeating consolidation is rather unlikely, but a large offset of the structural adjustment can be expected in initial years as well as only a delayed debt reduction.

2.9. GDP costs of consolidation

GDP growth and inflation vary among different consolidation paths of the alternative scenarios. Under the assumptions of the model, involving an endogenous reaction of the central bank, hysteresis is the main factor impacting negatively real growth. When combined with the higher fiscal multiplier and interest rate shock, real GDP can be lower after 20 years by 14% compared with the baseline scenario and even by 20% when this scenario involves delayed consolidation (Table 5). Most of these losses would be permanent due to the decline in potential output. Given very weak GDP growth, a country would experience almost two decades of large and negative output gaps (Figures A.4 and A.5). This is due to a combined effect of large and protracted consolidation (reflecting delayed consolidation and a weak cyclical improvement of the budget balance) and hysteresis magnified by the higher fiscal multiplier (as it results in larger output gaps). For scenarios without hysteresis, the differences in real GDP growth are small and will be ultimately eliminated with the closing of the output gap. The scenarios with the higher fiscal multiplier, and thus larger output gaps, have also much lower inflation – after 20 years, the price level would be by more than 12% lower than in the baseline. Inflation growth is even weaker when fiscal multipliers interact with hysteresis and the interest rate shock, resulting in the price level of up to 20% lower compared with the baseline. In the latter case, the inflation rates are close to or below zero for around a decade (Figures A.4 and A.5).

Table 5. **GDP costs of different consolidation scenarios**

	Scenario	Model assumptions							Real GDP	GDP deflator	Nominal GDP
		W1	W2	FM	RFSH	IR	Hyst.	AS	Level change after 20 years, in %		
1	Baseline	0.2	0.8	0.5	0.2	0	0.0	0.5	59.4	28.4	87.8
2	Alternative weights 1	1.0	0.0	0.5	0.2	0	0.0	0.5	61.0	27.6	88.6
3	Alternative weights 2	0.0	1.0	0.5	0.2	0	0.0	0.5	58.7	28.8	87.5
4	Interest rate shock	0.2	0.8	0.5	0.2	3	0.0	0.5	58.6	28.7	87.4
5	Interest rate shock with higher debt turnover	0.2	0.8	0.5	0.4	3	0.0	0.5	58.6	28.6	87.2
6	Hysteresis	0.2	0.8	0.5	0.2	3	0.0	0.5	52.4	28.3	80.7
7	Lower automatic stabilisers	0.2	0.8	0.5	0.2	0	0.1	0.5	59.4	27.6	87.0
8	Higher fiscal multiplier	0.2	0.8	1.5	0.2	0	0.0	0.3	60.9	16.0	76.9
9	Higher fiscal multiplier with hysteresis	0.2	0.8	1.5	0.2	0	0.1	0.5	47.3	14.5	61.8
10	Higher fiscal multiplier with hysteresis and interest rate shock	0.2	0.8	1.5	0.2	3	0.1	0.5	45.1	7.8	52.9
11	Delayed consolidation under baseline	0.2	0.8	0.5	0.2	0	0.0	0.5	57.8	32.5	90.3
12	Delayed consolidation with interest rate shock	0.2	0.8	0.5	0.2	3	0.0	0.5	58.2	32.7	90.9
13	Delayed consolidation with higher fiscal multiplier, hysteresis and interest rate shock	0.2	0.8	1.5	0.2	3	0.1	0.5	39.0	8.4	47.4

Note: *W1/W2* are weights of the government loss function, *FM* is the size of fiscal multiplier, *RFHS* is the share of government debt that matures within one year, *IR* indicates the size of interest rate shock, *Hyst.* is the size of the hysteresis effect and *AS* is the size of automatic stabilisers. For the sake of comparison, real GDP and price levels refer to the 20th year also for the scenarios with delayed consolidation, even though stabilisation of debt is achieved only in the 22nd year.
Source: Author's calculations.

Reaching the assumed debt target of 60% of GDP within 20 years in the presence of hysteresis and higher fiscal multipliers, especially in the case of delayed consolidation, can thus be very costly for the economy. This stresses the need of setting a realistic debt target

within a realistic deadline. This also implies that lowering debt to prudent levels will likely be very long process in many OECD countries, since achieving high budget surpluses over many years, even under the benign baseline scenario, would be politically difficult unless debt dynamics is helped by the improvement in interest-growth differentials.

3. Conclusions

The main implications of the analysis of this paper are as follows:

- Lowering high debt requires large and protracted consolidation, but if the aim is to stabilise debt at a lower level not all consolidation has to be permanent. This stems from the fact that the budget balance needed to stabilise debt at a low target is usually smaller than the level of budget balance needed to reduce initial high debt. In the scenarios considered, on average three-quarters of initial budget tightening can be reversed. The method employed and the two alternative *ad hoc* methods presented in Annex A ensure that the budget balance is consistent with debt stabilisation and they imply lower overall consolidation needs than the methods adopted in the literature which do not meet this condition.
- If a government is fully committed to meeting the debt target within a finite horizon, it could be optimal, from the growth maximisation perspective subject to the minimisation of extreme output gaps, to front-load somewhat the fiscal adjustment in case of an exogenous shock which raises government bond yields.
- Delaying the start of consolidation and the timing of reaching the debt target by two years has generally little implications for the size of initial consolidation. However, with high fiscal multipliers, interest rate shocks and hysteresis effects, this strategy would result in much higher initial debt and much slower nominal GDP growth.
- Ambitious debt targets in the context of adverse circumstances (large fiscal multipliers, hysteresis effects and interest rate shocks) can lead to a protracted spell of negative output gaps and deflation. This highlights the need to select realistic fiscal targets within realistic deadlines consistent with prevailing market conditions.
- Under the model specifications, a fully-fledged self-defeating consolidation, when the improvement in the structural primary balance results in no improvement or even a deterioration in the overall budget balance, is rather unlikely, but a large offset of the structural adjustment can be expected in initial years. Moreover, consistent with past experience, debt usually starts to decline several years after the start of consolidation.

The results presented are specific to the model assumptions, especially GDP growth dynamics and market reactions, but the sensitivity analysis of key parameters can still be informative about the expected effects of fiscal consolidation. The analysis could be further extended by investigating the effects of particular consolidation instruments as fiscal multipliers and implications for potential growth are likely to differ across various instruments, as well as of uncertainty, for instance by applying stochastic simulations. Moreover, in the context of concerted consolidation in many OECD countries, accounting for international spillovers would be desirable.

Notes

1. Several studies find negative correlation between high government debt, above 80-90% of GDP, and real GDP growth (Caner et al., 2010; Checherita and Rother, 2010; Kumar and Woo, 2010; Cecchetti et al., 2011), however the causality from debt to growth has been questioned (Panizza and Presbitero, 2012) and the threshold effect and its level hotly debated (Égert, 2013).

2. See for instance European Commission (2009), Sutherland et al. (2012) and OECD (2011).
3. For instance, Kanda (2011) uses a loss function based on the squared output gap and the squared fiscal sustainability gap (the latter is measured as the European Commission's S2 sustainability indicator – European Commission, 2009).
4. An alternative specification could include a maximisation of potential growth. This would, however, require feedback from fiscal policy to potential growth. This is not the case in the baseline scenario of this paper, but one form of such a feedback is assumed in simulations with hysteresis effects – see Section 2.4.
5. Calculating structural primary balances in practice is challenging and there is large uncertainty about derived consolidation measures (Bi et al., 2013).
6. The cyclical adjustment applies an aggregate elasticity (0.5 in the baseline scenario) to the level of the output gap.
7. This is somewhat lower than the estimated elasticity for the euro area of –0.19 by Guichard et al. (2009).
8. An alternative assumption would have to be made to account for a situation when monetary policy is not tailored to a particular economy, for instance as in the case of members of the euro area. A simple solution would be to assume an exogenous path of policy interest rates.
9. The nominal natural rate of 4% is assumed to be equal to the sum of real potential GDP growth (2%) and an inflation target (2%).
10. The average maturity in G10 countries in 2011 ranged from five years in the United States to 13 years in the United Kingdom (Rawdanowicz et al., 2011).
11. For instance, the average of budget deficit in the top quartile of OECD countries was around 8.5% of GDP in 2011, while average gross debt was around 110% of GDP, i.e. higher than assumed for the hypothetical economy.
12. Prudent debt levels are often selected in pragmatic and somewhat arbitrary ways. They are likely to vary across countries, reflecting diversified fundamentals, current conditions, preferences and future contingent and contractual liabilities (like pension obligations). Past debt reductions were, however, not only driven by the improvement in the structural primary balance.
13. This result does not change even if the time discount rate is increased from 1% to 5%, reducing the weight of growth rates at the end of the 20-year period. However, the higher discount rate would imply more back-loaded consolidation, according to the output gap criterion, as positive output gaps driven by fiscal easing at the end of the period would be discounted by more and thus less penalised. Conversely, without any time discounting, the output gap criterion would pick consolidation lasting 13 years instead of 14 under the baseline assumptions.
14. The budget deficit is found empirically to be one of the determinants of sovereign risk premia, along with gross debt and real GDP growth (Haugh et al., 2009; and Cottarelli and Jaramillo, 2012).
15. This low pass-through partly owes to the assumption that the government refinances short-term debt at the level of monetary policy rates.
16. The assumed cyclical adjustment is based on the level of the output gap and thus it impacts the size of the offset as well. The offset increases as the negative output gap narrows.
17. The real interest rate effect in the GDP equation is omitted as it depends on the path of structural budget balance and monetary policy rates over the following ten years.
18. This stems from the debt dynamics equation: $\Delta gd_t = -g_t/(1 + g_t) * nd_{t-1} + \Delta a_t - b_t$, where gd/nd is gross/net debt, g is the real GDP growth rate, a is financial assets and b is the overall budget balance. The erosion of the debt ratio is a function of initial net debt and a given growth rate. Consequently, the increase in debt would be smaller for higher initial debt (higher erosion effect).
19. In fact, for four scenarios (the baseline and the scenario with the higher fiscal multiplier but without the interest rate shock) the overall balance even increases initially by more than the structural primary balance. This primarily stems from lower interest payments on debt, following the decline in long- and short-term market interest rates, and in the baseline scenario, also from a cyclical improvement of the budget balance as the negative output gap closes.

References

Auerbach, A.J. and Y. Gorodnichenko (2012), "Measuring the Output Responses to Fiscal Policy", *American Economic Journal: Economic Policy*, Vol. 4(2), pp. 1-27.

Barnes, S., D. Davidsson and Ł. Rawdanowicz (2012), "Europe's New Fiscal Rules", *OECD Economics Department Working Papers*, No. 972.

Barrell, R., D. Holland and I. Hurst (2012), "Fiscal Consolidation: Part 2. Fiscal Multipliers and Fiscal Consolidations", *OECD Economics Department Working Papers*, No. 933.

Bi, R., H. Qu and J. Roaf (2013), "Assessing the Impact and Phasing of Multi-Year Fiscal Adjustment: A General Framework", *IMF Working Papers*, No. 13/182, International Monetary Fund.

Blöchliger, H., D. Song and D. Sutherland (2012), "Fiscal Consolidation: Part 4. Case Studies of Large Fiscal Consolidation Episodes", *OECD Economics Department Working Papers*, No. 935.

Bouis, R., B. Cournède and A.K. Christensen (2012), "Implications of Output Gap Uncertainty in Times of Crisis", *OECD Economics Department Working Papers*, No. 977.

Caner, M., T. Grennes and F. Koehler-Geib (2010), "Finding the Tipping Point – When Sovereign Debt Turns Bad", *World Bank Policy Research Working Papers*, No. 5391.

Cecchetti, S., M. Mohanty and F. Zampolli (2011), "The Real Effects of Debt", *BIS Working Papers*, No. 352, Bank for International Settelments.

Checherita, C. and P. Rother (2010), "The Impact of High and Growing Government Debt on Economic Growth. An Empirical Investigation for the Euro Area", *ECB Working Papers*, No. 1237, European Central Bank.

Christiano, L., M. Eichenbaum and S. Rebelo (2009), "When is the Government Spending Multiplier Large?", *NBER Working Papers*, No. 15394.

Cournède, B. (2010), "Gauging the Impact of Higher Capital and Oil Costs on Potential Output", *OECD Economics Department Working Papers*, No. 789.

Cottarelli, C. and L. Jaramillo (2012), "Walking Hand in Hand: Fiscal Policy and Growth in Advanced Economies", *IMF Working Papers*, No. WP/12/137, International Monetary Fund.

Corsetti, G. (ed.) (2012), *Austerity: Too Much of a Good Thing? A VoxEU.org eCollection of Views by Leading Economists*, CEPR eBook.

De Long, J.B and L.H. Summers (2012), "Fiscal Policy in a Depressed Economy", paper prepared for the Spring 2012 Brookings Panel on Economic Activity, 22-23 March, Washington, DC.

Égert, B. (2013), "The 90% Public Debt Threshold: The Rise and Fall of a Stylised Fact", *OECD Economics Department Working Papers*, No. 1055.

European Commission (2009), "Sustainability Report 2009", *European Economy*, No. 9/2009, Economic and Financial Affairs Directorate General, European Commission.

Gali, J., D. Lopez-Salido and J. Valles (2007), "Understanding the Effects of Government Spending on Consumption", *Journal of the European Economic Association*, Vol. 5(1), pp. 227-270.

Girouard, N. and C. André (2005), "Measuring Cyclically-Adjusted Budget Balances for OECD Countries", *OECD Economics Department Working Papers*, No. 434.

Guichard, S., D. Haugh and D. Turner (2009), "Quantifying the Effect of Financial Conditions in the Euro Area, Japan, United Kingdom and United States", *OECD Economics Department Working Papers*, No. 677.

Haugh, D., P. Ollivaud and D. Turner (2009), "What Drives Sovereign Risk Premiums?: An Analysis of Recent Evidence from the Euro Area", *OECD Economics Department Working Papers*, No. 718.

IMF (2010), *World Economic Outlook*, October, International Monetary Fund.

Johansson, A. et al. (2013), "Long-Term Growth Scenarios", *OECD Economics Department Working Papers*, No. 1000.

Kanda, D. (2011), "Modeling Optimal Fiscal Consolidation Paths in a Selection of European Countries", *IMF Working Papers*, No. WP/11/164, International Monetary Fund.

Koske, I. and N. Pain (2008), "The Usefulness of Output Gaps for Policy Analysis", *OECD Economics Department Working Papers*, No. 621.

Kumar, M.S. and J. Woo (2010), "Public Debt and Growth", *IMF Working Papers*, WP/10/174, International Monetary Fund.

OECD (2011), "Medium and Long-Term Developments: Challenges and Risks", Chapter 4 in *OECD Economic Outlook,* No. 89, OECD Publishing.

OECD (2012), *OECD Economic Outlook*, No. 91, Vol. 2012/1, May, OECD Publishing.

Panizza, U. and A. Presbitero (2012), "Public Debt and Economic Growth: Is There a Causal Effect?", *MoFiR Working Papers*, No. 65.

Rawdanowicz, Ł., E. Wurzel and P. Ollivaud (2011), "Issues in Government Debt and Asset Management", *OECD Economics Department Working Papers*, No. 923.

Reinhart, C.M. and K.S. Rogoff (2009), "The Aftermath of Financial Crises", *American Economic Review*, American Economic Association, Vol. 99(2), pp. 466-72.

Sutherland, D., P. Hoeller and R. Merola (2012), "Fiscal Consolidation. Part 1. How Much Is Needed and How to Reduce Debt to a Prudent Level?", *OECD Economics Department Working Papers*, No. 932.

Woodford, M. (2011), "Simple Analytics of the Government Expenditure Multiplier", *American Economic Journal: Macroeconomics*, Vol. 3(1), pp. 1-35.

ANNEX A

Calculating consolidation needs

Consolidation needs to achieve a given debt target can be calculated in several ways. The choice of a particular method has implications for the results and thus for the on-going policy debates on the size and pace of a fiscal adjustment. This annex highlights main determinants of such calculations, explains *pros* and *cons* of the existing methods and proposes two alternative solutions in addition to the approach discussed in the main paper.

Debt dynamics

Consolidation requirements are derived from the standard debt dynamics equation:[1]

$$d_t = \frac{1+i}{1+g} d_{t-1} - pb_t \qquad [A1]$$

where d_t is the debt-to-GDP ratio, pb_t is the primary balance-to-GDP ratio, i is the nominal interest rate paid on debt, g is the growth rate of the nominal GDP; i and g are assumed not to vary over time. This assumption is relaxed below. Given equation [A1], the debt ratio in year $N(d_N)$ can be expressed as a function of initial debt (d_0):

$$d_N = d_0 \left(\frac{1+i}{1+g}\right)^N - \sum_{t=1}^{N} pb_t \left(\frac{1+i}{1+g}\right)^{N-t} \qquad [A2]$$

The equation implies that the debt level in year N is a sum of discounted initial debt and a discounted sum of primary budget balances between the initial year and year N.

The consolidation requirement to reach a certain debt level after N years (the so-called fiscal gap) can thus be given by:

$$pb_N - pb_0 = d_0 \left(\frac{1+i}{1+g}\right)^N - \sum_{t=1}^{N-1} pb_t \left(\frac{1+i}{1+g}\right)^{N-t} - d_N - pb_0 \qquad [A3]$$

Hence, consolidation needs increase with higher initial debt and interest-growth rate differential ($\frac{1+i}{1+g} \cong 1 + i - g$), and with the lower debt target and initial primary balance.

With time-varying interest rates and nominal GDP growth rates, debt dynamics equations [A1]-[A3] become:

$$d_t = \frac{1+i_t}{1+g_t} d_{t-1} - pb_t \qquad [A1']$$

$$d_N = d_0 \prod_{t=1}^{N} \frac{1+i_t}{1+g_t} - \sum_{t=1}^{N-1} pb_t \prod_{k=t+1}^{N} \frac{1+i_k}{1+g_k} - pb_N \qquad [A2']$$

$$pb_N - pb_0 = d_0 \prod_{t=1}^{N} \frac{1+i_t}{1+g_t} - \sum_{t=1}^{N-1} pb_t \prod_{k=t+1}^{N} \frac{1+i_k}{1+g_k} - d_N - pb_0 \qquad [A3']$$

One-off and constant adjustment approaches

The magnitude of the fiscal gap is usually analysed in terms of the starting and end conditions as well as the interest-growth rate differential, but less so in terms of the evolution of the primary balance over time. In particular, some studies assume that the entire consolidation takes place in the first year and the primary balance is kept constant at the level that ensures reaching the debt target at the end of the simulation period (European Commission, 2009; Sutherland et al., 2012) – the one-off adjustment in Figure A.1.

There are two problems with this fiscal gap measure. *First*, for certain values of d_N, N, and interest and GDP growth rates, consolidation can be very large, especially if initial debt is considerably above the target. Such an adjustment can be neither desirable due to possible strong recessionary effects, with knock-on effects on government finances in the short term, nor feasible due to political constraints. *Second*, the primary balance at the end of the period is likely to differ sizeably from the one that stabilises debt at the target, and keeping it unchanged will continue lowering debt instead of stabilising. To address the first problem a more gradual fiscal tightening would be required, implying larger and longer initial cumulative consolidation. One simple solution is to assume constant annual consolidation lasting over the entire simulation period with the size of the adjustment implied by the condition to reach the debt target at the end of the simulation period (the constant adjustment in Figure A.1). This form of the adjustment implies a linear path of the primary balance and was used in the *OECD Economic Outlook 89* to show consolidation requirements with alternative debt targets (OECD, 2011).[2] The gradual consolidation implies a bigger total fiscal adjustment than under one-off consolidation (see Figure A.1 and further below). Although, in most cases the pace of consolidation will be moderate, this method does not address the second problem.

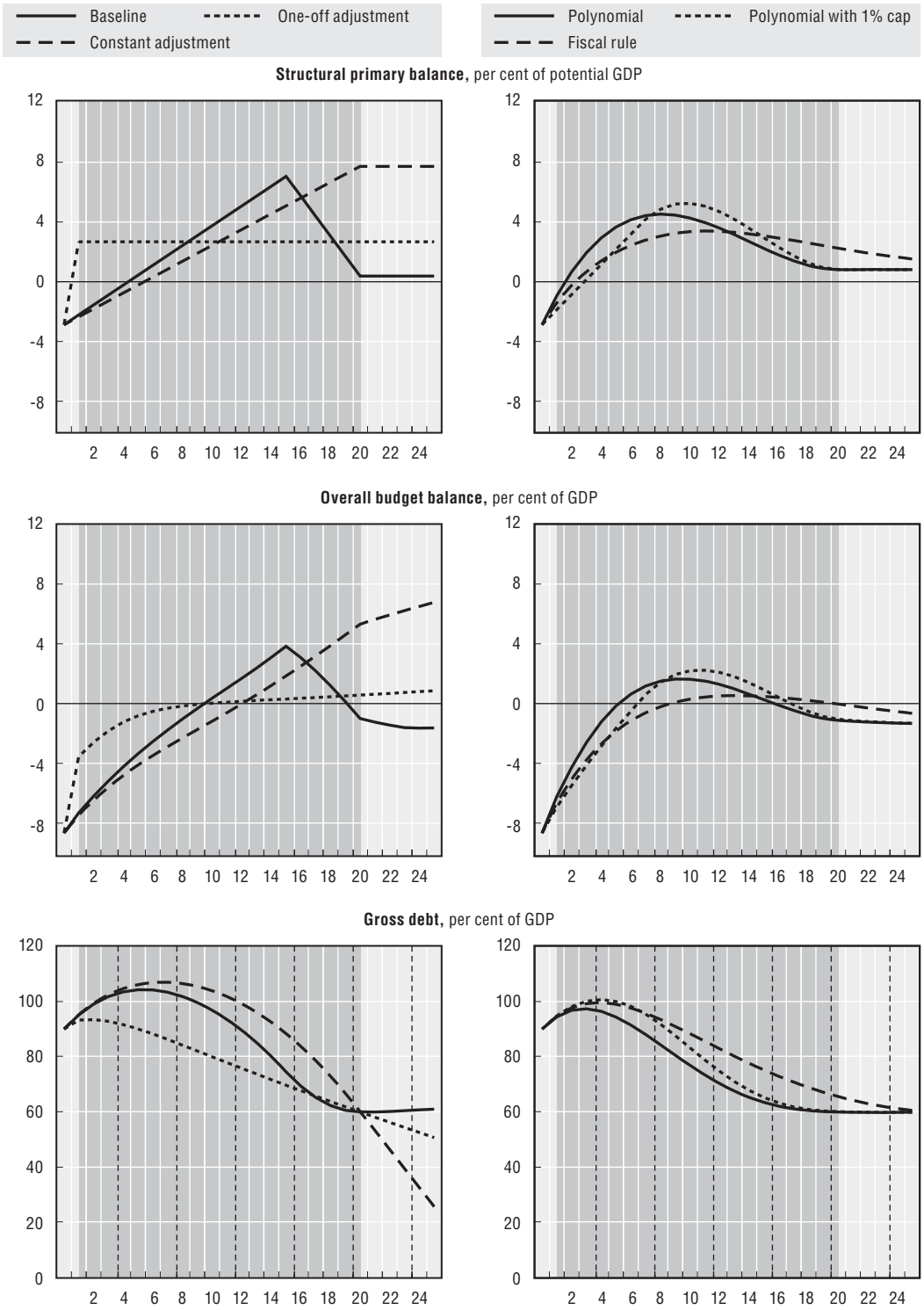
Alternative approaches

The consolidation path proposed in the main paper, which assumes constant annual adjustment during the consolidation and fiscal easing phase, ensures debt stabilisation at the end of that simulation period and avoids large swings in the fiscal policy stance. Other, more pragmatic and ad hoc approaches suited for simple illustrative calculations of consolidation needs are possible. Two of them are proposed in this annex. One adopts a mathematical approach based on a third-order polynomial and the other utilises a debt-stabilising fiscal rule. Both approaches address the two weaknesses of the fiscal gaps discussed above by imposing on the primary balance a gradual initial adjustment and convergence to a level guaranteeing debt stabilisation at the target.

Third-order polynomial approach

Many mathematical functions of the primary balance are possible to ensure its convergence to the level that stabilises debt at its target and to avoid large changes in the fiscal policy stance at the beginning and at the end of the adjustment period. In this annex, a third-order polynomial function is chosen. It not only meets these two conditions but also provides a plausible adjustment path which is easy to implement. In addition, it implies a smooth adjustment – a feature that is not essential from the policy point of view though. For a given (finite) duration of consolidation, a third-order polynomial function ensures that the primary balance improves at a diminishing rate (i.e. is concave), implying front-loaded consolidation, and that at the end it converges at a diminishing rate to a local

Figure A.1. **Different consolidation paths**



Note: Assumptions behind the presented adjustment paths are explained in the text. Baseline refers to the approach based on the optimisation of the policy loss function as described in the main paper (Section 1.2).
Source: Author's calculations.

minimum (i.e. is convex), implying smooth primary balance convergence (see Figure A.1 for the shape of such a function).

In practical terms, one has to find parameters of the following third-order polynomial so that the path of the primary balance guarantees reaching the debt target given the model's assumptions and debt accounting identities:

$$pb_t = \beta_1 t^3 + \beta_2 t^2 + \beta_3 t + \beta_4 \quad [A4]$$

where t stands for time (1… N).[3] Given projections of interest rates (i_t) and nominal GDP growth (g_t), four unknown parameters of the polynomial can be obtained by solving a system of four equations requiring that: *i)* the primary balance equals its initial level (pb_1) in the first year ($t = 1$); *ii)* in year t_x the primary balance is at a level pb_x that ensures reaching the debt target in year N; *iii)* the primary balance equals its target ($\text{pb}_N = \frac{i_N - g_N}{1+g_N} \text{d}_N$)[4] in the last year ($t = N$); and *iv)* the last period's primary balance is a local minimum ($3\beta_1 N^2 + 2\beta_2 N + \beta_3 = 0$). pb_x is chosen iteratively by solving equation [A2'] so that debt settles at its target in year N.

The system is formally presented in a matrix form:

$$\begin{bmatrix} pb_1 \\ pb_x \\ pb_N \\ pb_N \end{bmatrix} = \begin{bmatrix} 1^3 & 1^2 & 1 & 1 \\ x^3 & x^2 & x & 1 \\ N^3 & N^2 & N & 1 \\ 3N^2 & 2N & 1 & 0 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$$

Its solution is found as $\beta = X^{-1}PB$, where β is a vector of estimated parameters, X is a matrix of data points as given above, and PB is a vector of primary balances.

For given starting and end conditions and the time within which the debt target has to be reached, fitting equation [A4] does not ensure on its own that initial consolidation will not be very large (though it definitely avoids large fiscal loosening at the end). To this end, the polynomial has to be combined with a consolidation cap, limiting the annual adjustment to a certain threshold. The threshold can be selected based on past episodes of fiscal adjustments, taking into account a current state of the economy regarding the output gap, expected fiscal multipliers, etc. If the solution of the model implies a change in the primary balance in the first year which is bigger than the consolidation cap, the change is limited to the cap and equation [A2'] has to be solved again but starting from year 2 and with initial conditions as of the previous year. This process is repeated until the consolidation cap does not bind in any subsequent year.

The consolidation needs calculated by the polynomial method to be fully informative would be better presented by at least two numbers, expressing initial and total consolidation. For high starting values of debt relative to the target, initial cumulative consolidation (i.e. a cumulated improvement in the primary balance from its initial to the maximum level) is likely to be larger than total consolidation (i.e. a difference between initial and the last period's primary balance) – see the illustration below. This is in contrast to the fiscal gaps calculated based on the one-off and constant adjustments approaches, which can be summarised in one number.[5] In addition, in the polynomial method it could be instructive to provide information on the duration of fiscal tightening. In some cases, however, the smooth profile of the primary budget balance implies very small changes lasting several years which are not policy relevant. As a result the duration of fiscal tightening may be exaggerated.

Debt-stabilising fiscal rule

Gradual consolidation consistent with debt stabilisation at the target can also be worked out from a debt-stabilising fiscal rule. Given the aim to stabilise the debt-to-GDP ratio around a target, the fiscal rule has to be derived from the debt dynamics equation and consequently it has to simultaneously control for debt and the fiscal balance. After reaching the debt target, the debt ratio will stabilise only if the overall budget balance ratio settles at a particular level determined by nominal GDP growth and net debt: $b^* = -g/(1+g)(d^* - a^*)$, where b^* is the debt-stabilising budget balance, g is the growth rate of nominal GDP, d^* is target gross debt and a^* is target financial assets (the last bracket shows net debt; all variables but g are expressed in per cent of GDP). This condition is equivalent to the debt stabilisation requirement for the primary balance of the polynomial approach (see requirement *iii)* above).

The rule can thus be given by:

$$\Delta b_t = \alpha(b^* - b_{t-1}) + \beta(d_{t-1} - d^*) \qquad [A5]$$

where b_t is the current year's overall budget balance, b^* is the debt-stabilising overall budget balance (which is a function of the projected GDP growth rate and debt and asset targets), d_{t-1} is previous year debt, d^* is the debt target (all expressed in per cent of GDP), α and β are positive parameters determining the speed of convergence of the budget balance and debt to their targets.

If the rule were to be used in a policy context, it should allow for cyclical fluctuations of fiscal balances. The adjustment for automatic stabilisers can be done based on the output gap (the standard practice) or, alternatively, based on the difference between actual and potential GDP growth rates.[6] The latter method avoids using output gaps, whose estimates are usually revised more than the estimates of actual and potential GDP growth rates (Koske and Pain, 2008; Bouis et al., 2012), and could facilitate monitoring of the rule and communication with the public. Moreover, it could allow for the use of nominal growth, ensuring consistency with the calculation of b^* and requiring the authorities to implicitly reveal their inflation target.

The two approaches of accounting for automatic stabilisers have different implications for the pace of consolidation over the cycle. For instance, the use of the output gap implies less fiscal tightening than the use of the growth rate differential when the output gap is negative and closing as the actual GDP growth rate already exceeds the potential growth rate (Figure A.2). Consequently, after a recession the former method would allow for larger and more prolonged automatic stabilisers than the latter method. For this reason, the output gap proxy is chosen in this annex.

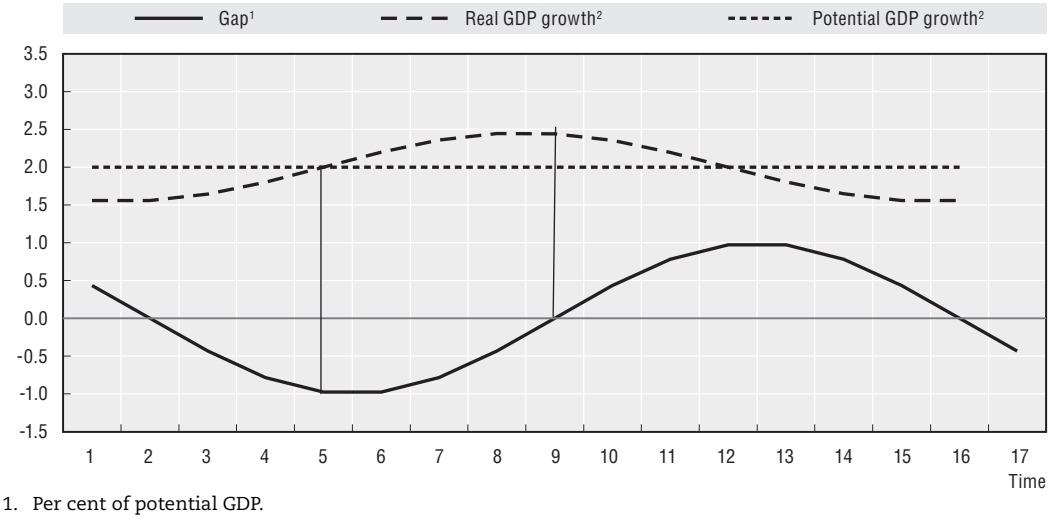
The complete rule is given by:

$$\Delta b_t = \alpha(b^* - b_{t-1}) + \beta(d_{t-1} - d^*) + \gamma gap_t \qquad [A6]$$

where gap_t is the output gap and γ is a positive parameter reflecting the size of automatic stabilisers. The rule refers to the overall balance, and not to the primary balance, as this makes the rule easier to monitor and more internally consistent. This is in contrast to the approaches discussed above, which focus on the structural primary balance.

The rule requires the overall budget balance to improve when, in the previous year, it was below its target, debt was above its target or when the (projected) output gap for the current year is positive. However, when the output gap is closed and the fiscal balance and

Figure A.2. **Relation between the output gap and the potential and actual GDP growth rates**



1. Per cent of potential GDP.
2. Per cent.

Source: Author's calculations.

debt are at their targets, the budget balance has to remain unchanged. When α and β are below 1, the adjustment to targets is gradual (see below). By construction, the rule would allow for a full operation of automatic stabilisers only when debt and the fiscal balance are at their targets.[7] Consequently, the rule can be used to measure fiscal policy room to accommodate negative shocks (the so-called fiscal space). The rule can be effective in preventing debt ratchet effects. No budgetary slippages would be forgotten, as deviating further from the targets will require more consolidation in the future. In contrast, a better budget outcome will lower the consolidation need in the following year.

Despite allowing for automatic stabilisers, the fiscal rule implies large initial consolidation when the budget balance and gross debt are far from their targets. This may not be desirable, especially when a large output gap persists or when there is a sizeable increase in gross debt resulting from the purchase of financial assets (without changing net debt as it can be the case after a financial crisis).[8] Consequently, a consolidation cap can be applied to the fiscal rule similarly to the polynomial approach.

The fiscal rule cannot however control precisely the timing of reaching a debt target and the size of initial consolidation, as these outcomes depend not only on parameters α and β but also on initial conditions. This may complicate comparing the size of consolidation requirements with other methods, but is less of a problem from the point of view of fiscal policy guidance.

Parameters α and β can be chosen pragmatically in model simulations under different assumptions (including starting and end values) as well as economic shocks so as to render reasonable and robust results. They could also be based on the existing fiscal rules. In the case of EU countries, for example, parameter β could be set at 0.05 in line with the new debt convergence rule, requiring the EU countries to reduce the excess of the gross debt ratio over the Treaty limit of 60% by 1/20th on average over three years (Barnes et al., 2012). Parameter γ could reflect existing estimates of the size of automatic stabilisers.

Illustration

To illustrate differences in the size of consolidation needs resulting from various possible adjustment paths, stylised calculations are done for the hypothetical economy under the baseline specification describe in Section 1.2 of the main paper. The country is expected to reach the debt target of 60% of GDP within 20 years. Simulations assume feedback from consolidation to growth, from interest rates to growth and from debt to interest rates.

The model specifications imply the following parameters of the polynomial:

$$pb_t = 0.004t^3 - 0.196t^2 + 2.481t - 5.139, \qquad t \in [1,21] \qquad \text{[A7]}$$

and the fiscal rule is parameterised as follows:

$$\Delta b_t = 0.4(b^* - b_{t-1}) + 0.05(d_{t-1} - d^*) + 0.5gap_t \qquad \text{[A8]}$$

Parameter β (0.05) is selected to reflect the new EU debt convergence rule. Parameter α (0.4) is set so as to ensure relative fast convergence to the debt-stabilising balance and thus to avoid debt undershooting.[9] Parameter γ (0.5) equals the OECD average size of automatic stabilisers (Girouard and André, 2005).

Figure A.1 and Table A.1 show consolidation paths and requirements under the one-off permanent fiscal adjustment, constant adjustment, gradual front-loaded consolidation without any cap (Polynomial), gradual front-loaded consolidation with an annual cap of 1.5% of GDP (Polynomial with 1% cap), the illustrative fiscal rule and the baseline specification with the optimised kinky-linear underlying balance path from the main paper.

Table A.1. **Consolidation requirements under different consolidation paths**

	Consolidation			
	1st year, % of GDP	Initial, % of GDP	Total, % of GDP	Duration, years
One-off adjustment	5.5	5.5	5.5	1
Constant adjustment	0.5	10.5	10.5	20
Polynomial	1.9	7.3	3.6	8
Polynomial with 1% cap	1.0	8.0	3.6	10
Fiscal rule	1.4	6.2	5.0	11
Baseline	0.7	9.8	3.2	15

Note: Initial consolidation is defined as the improvement in the structural primary budget balance between its initial and maximum level, whereas total consolidation refers to the difference between the initial and the final primary budget balance. Baseline refers to the approach based on the optimisation of the policy loss function as described in the main paper (Section 1.2).
Source: Author's calculations.

The one-off adjustment under the standard method is very large, in this example nearly 6% of GDP (Table A.1, Figure A.1). All other methods imply more gradual consolidation (the most gradual with the constant adjustment and the least gradual with the polynomial), resulting in a longer and larger initial but not total adjustment, respectively. The polynomial approach without the cap and the fiscal rule require 1.8 and 0.7% of GDP more of initial fiscal tightening which last 7 to 10 years longer than under the one-off adjustment. However, the total adjustment is lower by almost 2% of GDP (in the case of the fiscal rule the difference is only 0.5% of GDP but the rule does not stabilise debt at the target within 20 years). The cap of 1% of GDP does not change the results much. Consolidation needs under the constant adjustment are the largest, double those under the one-off adjustment.

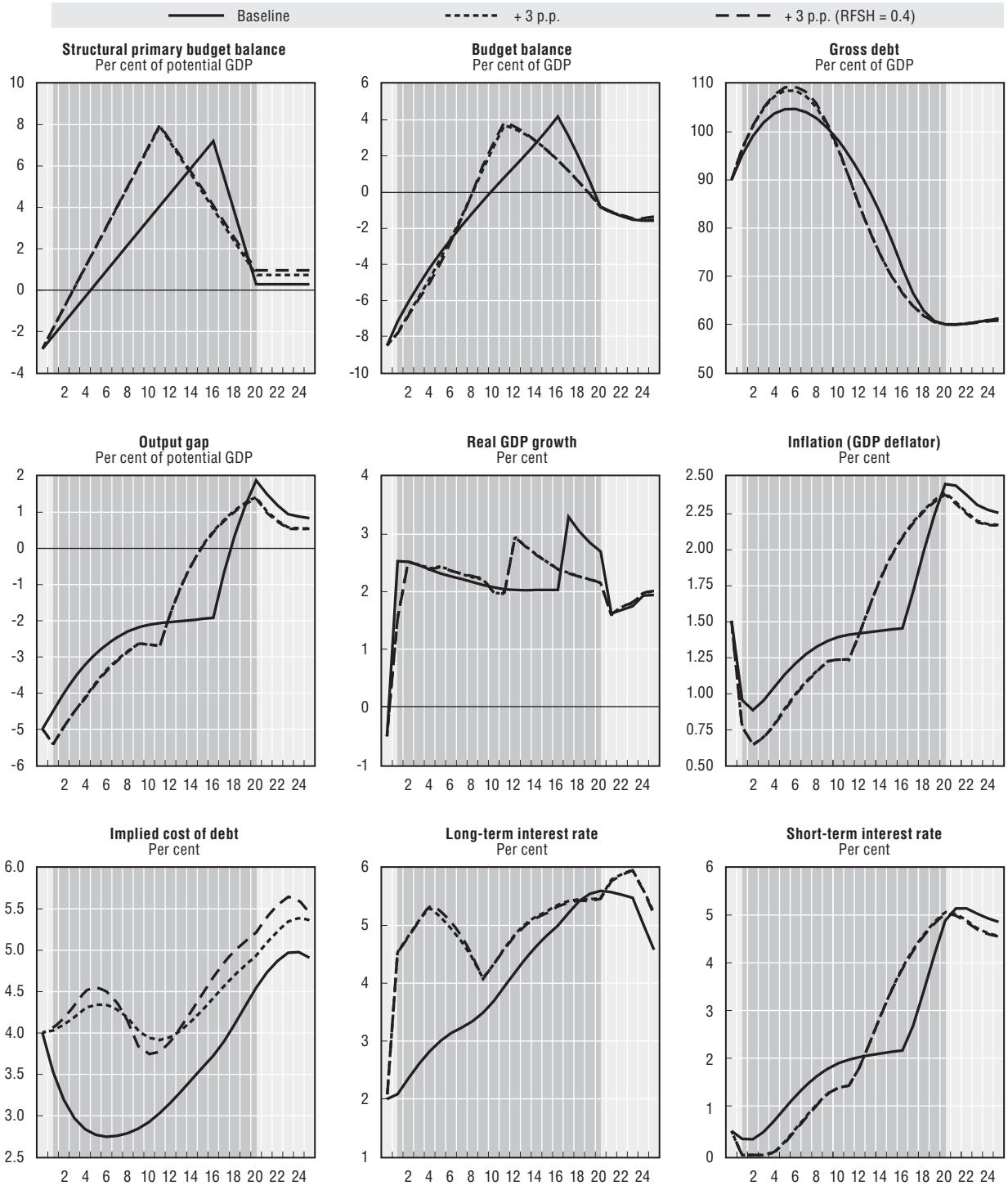
Table A.2. **Changes in the structural primary balance, overall balance and debt in selected scenarios**

Per cent of (potential) GDP

	Years									
	1	2	3	4	5	6	7	8	9	10
1. Baseline										
Structural primary balance	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Overall balance	1.3	1.1	1.0	0.9	0.8	0.8	0.7	0.7	0.7	0.7
Gross debt	5.1	4.0	2.8	1.8	0.9	0.1	-0.7	-1.4	-2.0	-2.6
Gross debt (level)	95	99	102	104	105	105	104	103	101	98
4. Interest rate shock										
Structural primary balance	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Overall balance	0.7	1.0	1.0	1.0	1.1	1.1	1.2	1.3	1.4	1.2
Gross debt	6.4	4.7	3.6	2.5	1.2	0.0	-1.2	-2.5	-3.8	-4.7
Gross debt (level)	96	101	105	107	109	109	107	105	101	96
5. Interest rate shock with higher debt turnover										
Structural primary balance	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Overall balance	0.7	1.0	0.9	0.9	1.1	1.2	1.3	1.4	1.5	1.2
Gross debt	6.5	4.8	3.7	2.7	1.5	0.2	-1.1	-2.6	-4.0	-5.0
Gross debt (level)	96	101	105	108	109	109	108	106	102	97
8. Higher fiscal multiplier										
Structural primary balance	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Overall balance	1.3	0.9	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.7
Gross debt	5.4	4.6	3.6	2.6	1.6	0.7	-0.2	-1.0	-1.7	-2.3
Gross debt (level)	95	100	104	106	108	108	108	107	106	103
9. Higher fiscal multiplier with hysteresis										
Structural primary balance	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Overall balance	1.3	0.9	1.0	1.0	0.9	0.9	0.9	0.9	0.8	0.7
Gross debt	5.8	5.0	4.0	2.9	1.9	1.0	0.1	-0.8	-1.6	-2.3
Gross debt (level)	96	101	105	108	110	111	111	110	108	106
10. Higher fiscal multiplier with hysteresis and interest rate shock										
Structural primary balance	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
Overall balance	0.3	0.6	0.7	0.8	0.8	1.0	1.2	1.3	1.4	1.5
Gross debt	7.9	6.9	6.1	5.3	4.4	3.2	1.9	0.6	-0.8	-2.3
Gross debt (level)	98	105	111	116	121	124	126	126	125	123
11. Delayed consolidation under baseline										
Structural primary balance	0.0	0.0	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Overall balance	0.8	0.5	1.0	0.9	0.9	0.8	0.8	0.8	0.8	0.8
Gross debt	5.5	4.7	3.9	2.9	1.9	1.0	0.1	-0.7	-1.4	-2.1
Gross debt (level)	95	100	104	107	109	110	110	109	108	106
12. Delayed consolidation with interest rate shock										
Structural primary balance	0.0	0.0	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Overall balance	0.0	0.2	1.0	1.0	1.0	1.2	1.3	1.5	1.6	1.7
Gross debt	6.8	5.9	5.2	4.0	2.8	1.5	0.0	-1.5	-3.1	-4.7
Gross debt (level)	97	103	108	112	115	116	116	115	112	107
13. Delayed consolidation with higher fiscal multiplier, hysteresis and interest rate shock										
Structural primary balance	0.0	0.0	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
Overall balance	0.1	0.3	0.4	0.5	0.6	0.7	0.8	1.0	1.2	1.4
Gross debt	6.9	5.9	6.9	6.5	5.9	5.2	4.4	3.2	1.9	0.4
Gross debt (level)	97	103	110	116	122	127	132	135	137	137

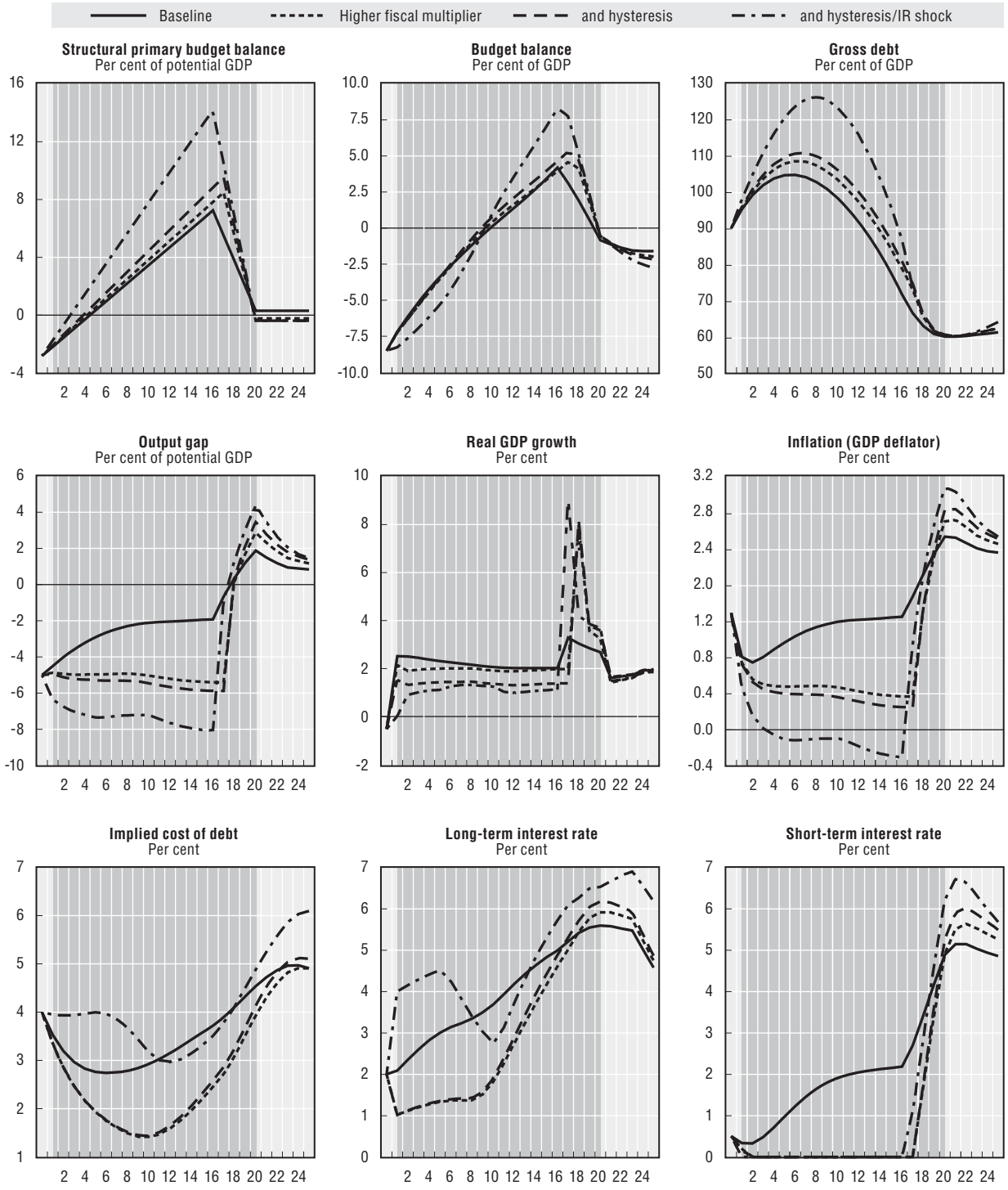
Note: Structural balances are expressed in per cent of potential GDP, whereas overall balances and gross debt as per cent of actual GDP. For specifications of scenarios see Table 2.
Source: Author's calculations.

Figure A.3. **Long-term interest rate shock**



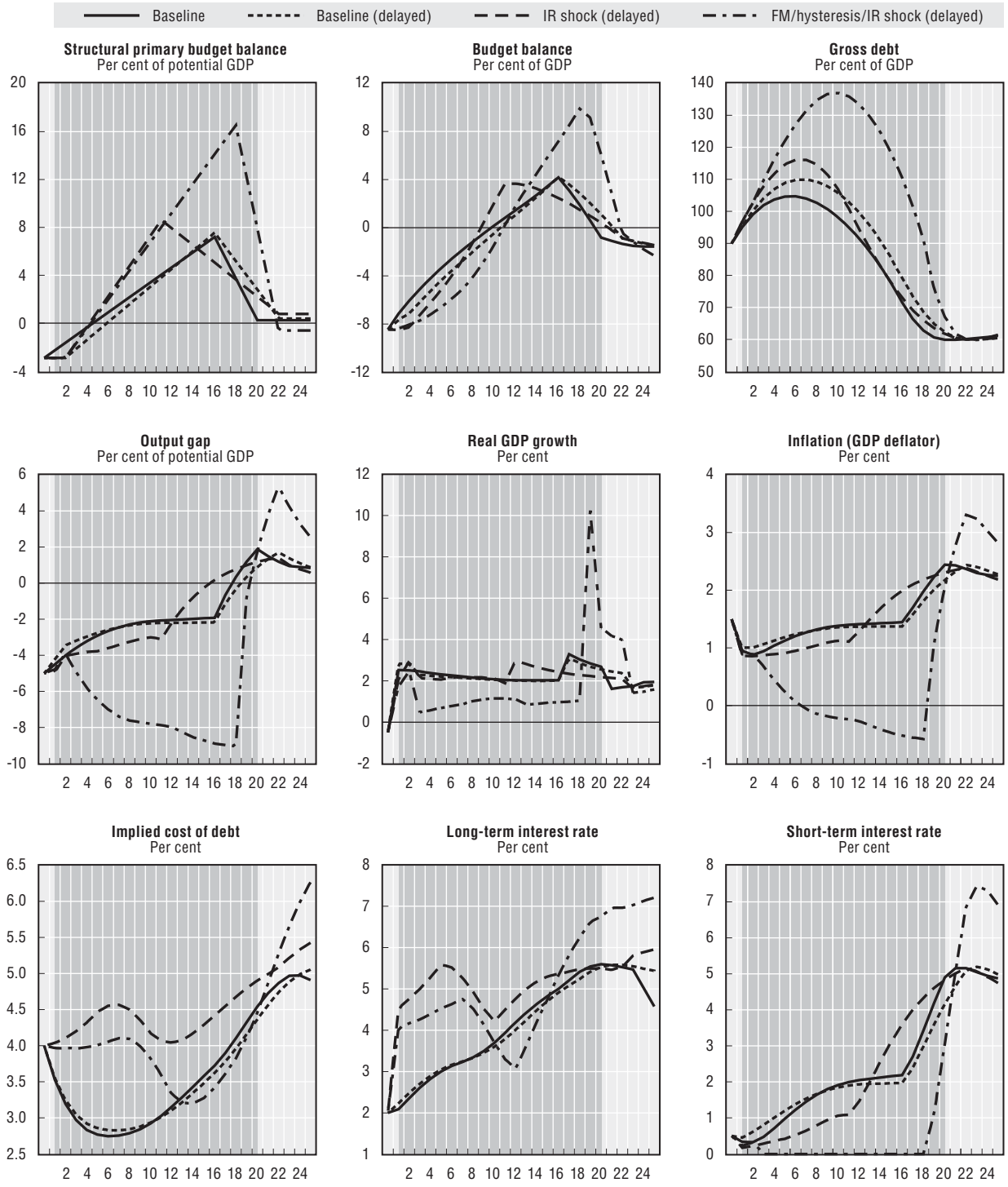
Source: Author's calculations.

Figure A.4. **Higher fiscal multipliers**



Source: Author's calculations.

Figure A.5. **Delayed consolidation**



Source: Author's calculations.

As expected, the one-off and constant adjustment methods do not ensure stabilisation of debt at its target (Figure A.1) and the primary budget balances are significantly higher than required to stabilise debt at the target (by between around 2% and 7% of GDP). Consequently, sustaining the balances at this level results in a continued decline in the debt ratio. In contrast, all other methods ensure debt stabilisation, though for the fiscal rule it takes more than 20 years.

Notes

1. Under the assumption of no government financial assets (i.e. when gross equals net debt). With financial assets, the gross debt equation becomes: $d_t = (1 + i_p)/(1 + g)d_{t-1} - (1 + i_r)/(1 + g)a_{t-1} - pb_t + a_t$, where a_t is the ratio of financial assets to GDP, i_p is the interest rate paid on debt, and i_r is the interest rate earned on assets.
2. See Figure A.2. The baseline calculations of consolidation needs in *OECD Economic Outlook* No. 89 assumed a constant consolidation of 0.5% of GDP per year until the primary balance reached a level consistent with debt stabilisation at its prevailing level, resulting in debt stabilisation in OECD countries at different points in time and at different levels. Thus, this method has a different objective (i.e. it does not require reaching a specific debt target within a common, finite horizon) and is not appropriate for a comparison here.
3. To ensure consistency with equations [A1]-[A3], *N* in equation [4] must be higher than *N* in equations [A1]-[A3] by 1, so that initial period 0 in equations [A1]-[A3] corresponds to year 1 in equation [A4], as the polynomial cannot be solved for $t = 0$.
4. With assets the condition becomes: $\mathrm{pb_N} = \frac{i_{pN} - g_N}{1+g_N}\mathrm{d_N} - \frac{i_{rN} - g_N}{1+g_N}\mathrm{a_N}$.
5. For these two approaches, the last value of the primary balance is equal to its maximum, which is not the case with the polynomial approach.
6. If the adjustment is based on the change (not level) of the output gap – which is sometimes the case – then it is approximately equivalent to the one based on the difference between actual and potential GDP growth rates.
7. For certain values of parameters α and β and large negative shocks, almost a full operation will be possible even when debt and deficits diverge (within some limits) from their targets.
8. To avoid the latter problem, the fiscal rule could be defined in terms of net and not gross debt.
9. Very high initial debt relative to the target requires large budgetary surpluses and low α will bring the surplus down only slowly, potentially undershooting the debt target.

OECD Journal: Economic Studies
Volume 2013
© OECD 2014

Policy incentives for private innovation and maximising the returns

by
Ben Westmore*

This paper uses panel regression techniques to assess the policy determinants of private-sector innovative activity – proxied by R&D expenditure and the number of new patents – across 19 OECD countries. The relationship between innovation indicators and multifactor productivity (MFP) growth is also examined with a particular focus on the role of public policies in influencing the returns to new knowledge. The results establish an empirical link between R&D and patenting, as well as between these measures of innovation intensity and MFP growth. Innovation-specific policies such as R&D tax incentives, direct government support and patent rights are found to be successful in encouraging the innovative activities associated with higher productivity growth. However, direct empirical evidence of the positive effects of these policies on productivity is less forthcoming. A pervasive theme from the analysis is the importance of coupling policies aimed at encouraging innovation or technological adoption with well-designed framework policies that allow knowledge spillovers to proliferate. In particular, the settings of framework policies relating to product market regulation, openness to trade and debtor protection in bankruptcy provisions are found to be important for the diffusion of new technologies.

JEL classification: L20, O30, O40

Keywords: Intangible assets, innovation, productivity growth, public policy

* The author works in the OECD Economics Department (ben.westmore@oecd.org). This paper was originally part of the Working Party No. 1 project on "Knowledge-based capital, innovation and resource allocation" conducted jointly with the Science, Technology and Industry Directorate. Without implication, the author would like to thank Dan Andrews, Jørgen Elmeskov, Giuseppe Nicoletti, Dominic Guellec, Benjamin Montmartin, Chiara Criscuolo and Carlo Menon, for valuable comments and suggestions, Catherine Chapuis for technical assistance and Irene Sinha for editorial support. The views expressed in this paper are those of the author and do not necessarily reflect those of the OECD and its member countries.

1. Context and purpose

Beginning with the work of Solow (1956), the importance of innovation as a driver of long-term economic growth has long been recognised. Such activity may be characterised by positive externalities, as innovation by one entity can produce knowledge spillovers which benefit follow-on innovators. The fact that these spillovers are often non-rival and non-excludable results in the social rate of return to many innovative activities exceeding the private rate of return (Arrow, 1962). As a consequence, private firms tend to under-invest in innovation, potentially motivating public policy intervention to boost innovative activity to more socially optimal levels.

This paper investigates the influence of public policies on private sector innovation and the returns to new knowledge. The analysis uses country-level data to assess: *i)* the policy determinants of business research and development (R&D); *ii)* the policy determinants of the number of new patents; and *iii)* the link between innovative activity and multifactor productivity (MFP). One benefit of studying innovation at the aggregate instead of the sectoral or firm level is that the spillover effects between firms, industries and countries that are associated with innovation are better identified (Guellec and Van Pottelsberghe, 2003). Spillovers are critical in determining the degree to which new technologies, production processes and organisational methods translate into economy-wide productivity gains, and public policies may play a role in shaping this process. Nevertheless, a possible downside to an aggregate-level study is that the estimated relationships may be more sensitive to changes in sectoral specialisation patterns and less informative about the channels through which policies influence innovation and productivity growth than analysis based on more disaggregated data.

This study is conducted over three separate panels of 19 countries[1] from the mid-1980s to 2008 and is an update and extension of earlier OECD work by Jaumotte and Pain (2005a). A number of key findings emerge from the analysis:

- Both business and non-business R&D have a positive influence on patenting activity. In turn, business R&D and the stock of patents are positively associated with MFP growth. There is some evidence that innovative activity has a larger effect on MFP growth when the share of business researchers in total employment is high, reflecting the importance of absorptive capacity for the proliferation of knowledge spillovers.
- More generous innovation-specific policies are found to encourage private sector innovative activity, though empirical evidence regarding the direct impact of such policies on MFP growth is less forthcoming. In particular:
  - R&D tax incentives are positively associated with recorded private sector R&D spending. However, this does not hold in countries that frequently reverse policy settings, suggesting that a predictable policy environment may be important for the efficacy of such measures.

- ❖ In contrast with some earlier studies, direct government support for private R&D is found to be positively associated with business R&D spending. This finding relies on data from the post-2000 period, raising the possibility that the design of such policies has improved through time.

- Framework policies that promote competition have a pervasive influence on incentives to innovate and the returns to new ideas. For example:
  - ❖ Pro-competition reforms to product market regulation are associated with an increase in the number of patents. There is also some, albeit weaker, evidence that such reforms encourage higher business R&D spending.
  - ❖ Strengthening patent rights is more likely to encourage patenting activity when regulatory barriers to firm entry are low.
  - ❖ Domestic patenting activity appears to benefit from access to foreign R&D, but only when combined with low barriers to firm entry. Such an environment mitigates the opportunity for incumbent firms with monopoly rights over existing technologies to resist the adoption of innovations from abroad. This idea aligns with the finding that countries are slower to converge to the productivity frontier when product market regulation is strict.
  - ❖ Less stringent product market regulation is associated with higher MFP growth. This may reflect easier implementation and commercialisation of new ideas in more competitive markets.
  - ❖ Provided that domestic barriers to firm entry are low, countries with greater geographical proximity to technological leader countries exhibit stronger MFP growth.
- Other framework policies relating to trade and bankruptcy laws may also be important for the diffusion of innovations from abroad. In particular:
  - ❖ Firms with access to a greater quantity of foreign R&D undertake less domestic R&D, suggesting a degree of substitutability between the two sources of innovation.
  - ❖ Countries behind the productivity frontier tend to converge more rapidly the more open they are to trade.
  - ❖ Convergence is faster in countries with bankruptcy regimes that are less punishing to debtors. This is consistent with firm level evidence suggesting that low exit costs make it less likely that resources are tied up in inefficient firms (Andrews and Cingano, 2012).

1.1. Roadmap

The next section examines the existing evidence of the links between innovation and productivity growth, outlining a framework for thinking about the relationship between the variables of interest and the influence of public policies. Section 3 describes three empirical models that seek to explore the links between a range of potentially relevant policy variables and the key outcome variables relating to business R&D, the number of new patents per capita and multifactor productivity growth, respectively. Section 4 describes the data used for the analysis and Section 5 presents the econometric results. Section 6 summarises the findings and offers some concluding thoughts.

2. Innovation and public policies in OECD countries: Existing evidence

2.1. The relationship between innovation and productivity growth

A large volume of literature has found that the majority of income differences in developed economies cannot be explained by the stock of labour and tangible capital resources (Caselli, 2005). The residual is ascribed to differences in productivity that depend on countries' ability to accumulate intangible assets, not least through the process of innovation and its diffusion. For each of the 19 OECD countries in the sample, Figure 1 plots the evolution of a measure of MFP that accounts for the capital stock and employment adjusted for human capital (Panel A), along with the business R&D stock to GDP (Panel B) and the stock of triadic patent families (Panel C).[2] Before proceeding, it is important to recognise that such measures of MFP are prone to measurement error as they are derived as the residual once the other (imperfectly measured) production factors are taken into account. That said, common patterns in measured productivity have recently been exhibited across the countries in the sample, with average MFP growth generally slowing since 2000 relative to the previous decade.[3]

Empirical studies tend to find a positive link between business R&D and measures of innovative outputs such as new patents (see Danguy et al., 2009 for a review) and more importantly, productivity growth (see Nadiri, 1993; or Hall et al., 2009 for a review).[4] Examining trends in these variables indicates that in Belgium, Canada, Ireland and the United States, a slowing in the pace of productivity growth since 2000 has occurred at the same time as weaker growth in the stock of business R&D (to GDP) and patents. Nevertheless, movements in productivity growth in a number of countries do not seem to accord with trends in R&D and patents. Certainly, recent declines in productivity in many regions are difficult to reconcile with increases in the stock of business R&D and patents over the period.[5] Taken together, these trends suggest that while it is important to gain a better understanding of the policy determinants of business R&D and patenting activity, it is also necessary to investigate the factors that may influence the degree to which such innovative activity is reflected in movements in measured productivity growth.

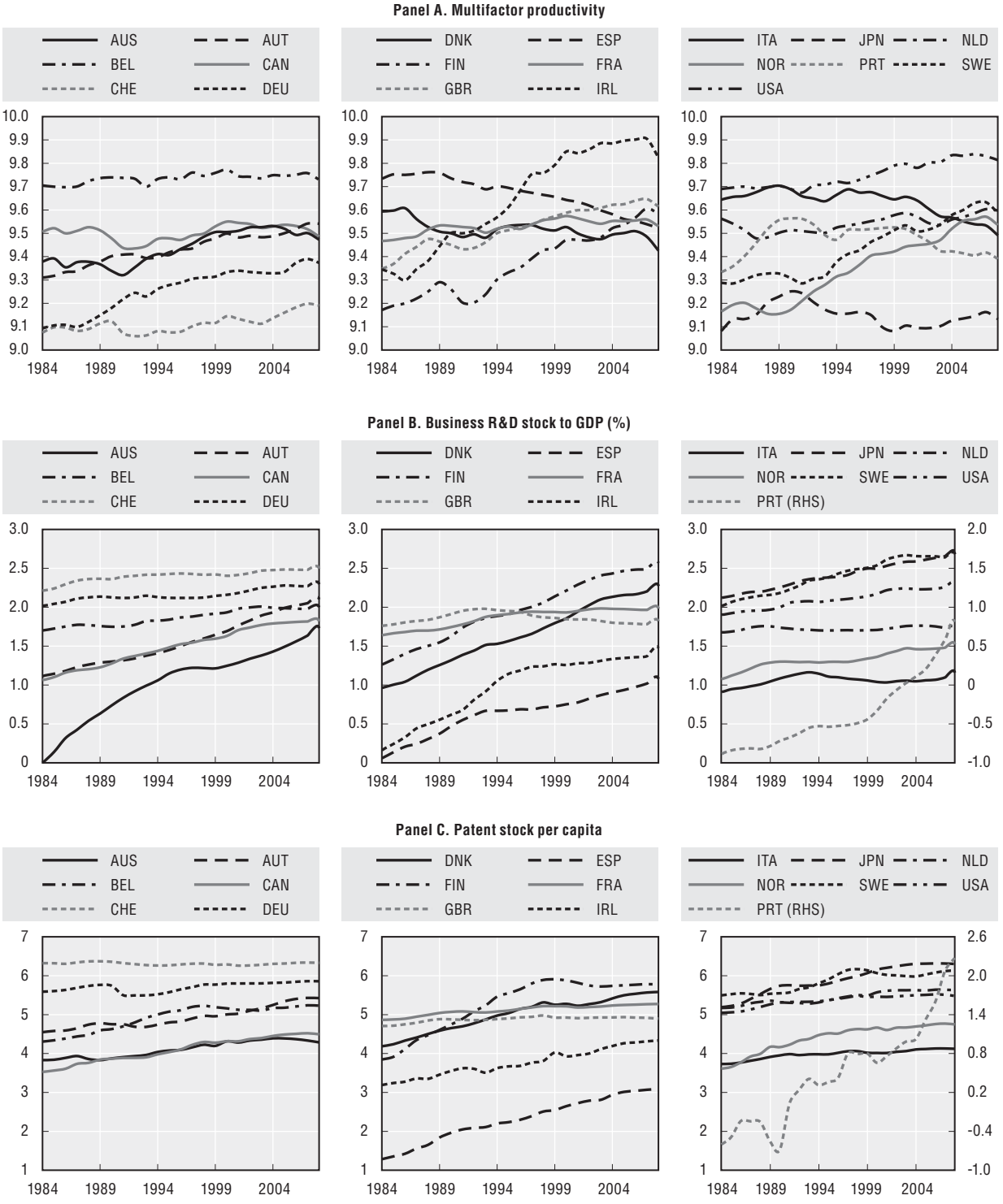
The innovation process is complex and varies greatly across industries and technologies. With this in mind, Figure 2 portrays a general – albeit heavily stylised – framework that maps the linkages between government policies, measures of innovation and productivity growth.

From an initial knowledge state that is drawn-upon throughout the innovation process, firms invest in R&D in the conviction that growth in the stock of such activity improves the probability of innovative success – which can be imperfectly proxied by the number of new patents.[6] The non-rivalrous nature of knowledge means that it is possible for novel ideas generated from R&D to be used by other innovative entities at no additional cost, enabling increasing returns to R&D in terms of the number of new patents. Likewise, a subsequent rise in the stock of patented ideas can benefit multiple follow-on innovators, as a condition of being conferred patent rights is that the patent holder publicly discloses technical information regarding the underlying innovation. As made explicit in Figure 2, R&D may also contribute to increasing the stock of technological advances that are not patented. In the case of non-patent intangibles, firms can use other methods of appropriating returns such as secrecy, barriers to entry and long lead times.

New ideas – patented or not – are absorbed into the knowledge stock and can contribute to higher productivity as they are utilised to produce a higher volume of output

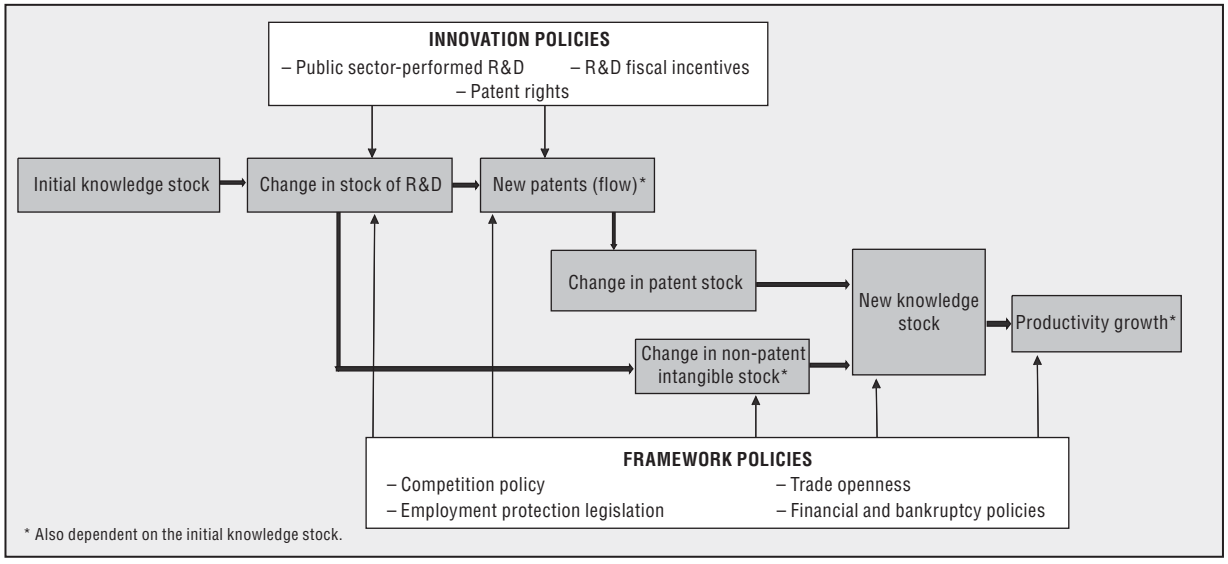
Figure 1. **Cross country differences in MFP, the business R&D stock to GDP and the patent stock**

1984-2008, log scale



Source: The stock of R&D and patents per capita (million working age population) are author's calculations based on data from OECD *Science and Technology Indicators*. MFP data are taken from Johansson and Murtin (2012) and, along with the business R&D stock to GDP, are calculated from data in constant 2005 PPP USD terms. Panel C shows signs of a possible data break in Germany around the time of the country's reunification. However, unreported regressions controlling for this break confirm that the estimation results reported in Section 5 are robust to this feature of the data.

Figure 2. **A stylised depiction of the private sector innovation process**



Note: Non-R&D expenditures and informal innovation at the workplace level are also methods for augmenting the existing knowledge stock and driving productivity growth not illustrated here.

for the same volume of inputs.[7] The effect on productivity is the focus of Figure 2 and the analysis that follows, but it should be acknowledged that, in the short to medium term, the gains from innovation could come in the form of higher firm profits as innovations may lower input costs or allow firms to raise finished good prices without increasing the volume of output. The magnitude of the effect on productivity is likely to be influenced by the quality and execution of new ideas as well as the size of the spillovers that exist. Each stage of the innovation process will be influenced by government policies (the lighter arrows in Figure 2) relating specifically to innovation as well as economy-wide framework policies, which are now discussed in turn.

2.2. Innovation policies

Innovation-specific policies strengthen business incentives to innovate, counteracting the private under-investment in such activity. A number of such policies are targeted to particular types of firms or industries, and hence, may be best analysed using event studies or data at a highly disaggregated level. The analysis in this paper focuses on a few larger scale innovation policies for which relatively reliable cross-country data exist and that are suspected to directly influence the dependent variables of interest.

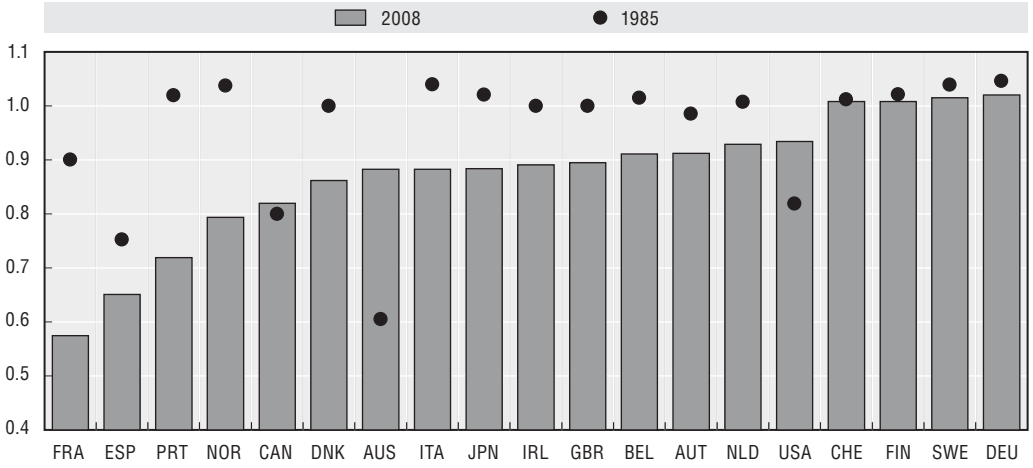
2.2.1. R&D tax incentives

Tax incentives for R&D can take various forms but are market-based instruments that leave firms to decide which R&D activities to fund. Either expenditure-based schemes (i.e. R&D tax credits, tax allowances and payroll withholding tax credits for R&D wages), or income-based schemes (i.e. taxing income derived from knowledge-based capital at a preferential rate), can be implemented. The overall generosity of R&D tax incentives can be summarised by the B-index (Warda, 2001), which represents the required rate of pre-tax return to justify $ 1 of R&D outlay taking account of both R&D tax incentives and the corporate income tax rate. Consequently, an increase in R&D tax incentives is reflected by a decline in the B-index.

A comparison of the level of the B-index for the sampled countries in 1985 and 2008 highlights that R&D tax incentives in most countries have become more generous in recent years, but important cross-country differences remain (Figure 3).

Figure 3. **R&D tax incentives**

The B-index at 1985 and 2008



The majority of existing cross-country studies find that tax incentives are effective in explaining the evolution of business R&D with a long-run elasticity around unity: on average, one dollar of tax incentive eventually induces one dollar of recorded private R&D spending (Hall and Van Reenen, 2000). This may vary with the design of the incentive, with incremental tax credits (i.e. R&D expenditures are only eligible if they exceed some baseline amount) sometimes found to be more effective in inducing business R&D spending than general – volume-based – tax credits (i.e. all R&D expenditures are eligible; Lokshin and Mohnen, 2009). Regardless, the implications for economic welfare are unclear since the introduction of an R&D tax incentive raises administrative and compliance costs as well as the need for financing through distortionary taxation.

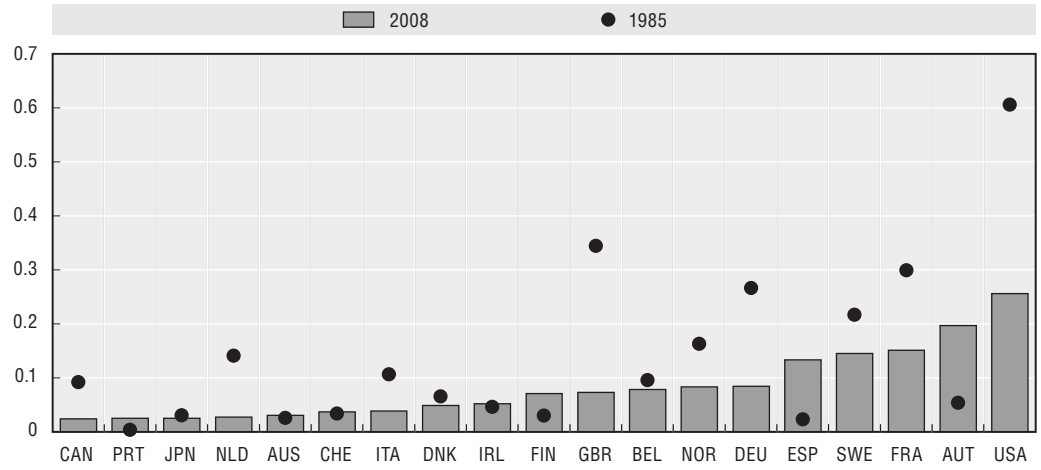
2.2.2. Government financed business R&D

Direct government funding of private R&D can come in a number of different guises and can vary significantly in the way it is administered (for an overview, see Blanco Armas et al., 2006). Grants, loans, loan guarantees or procurement contracts are all ways that governments can directly support business R&D, though each may have quite different objectives. For instance, in many countries, government loans support private R&D activities that are focused on the needs of civilians, while the type of private R&D procured by government is often defence-related.

Taking all direct support measures together indicates that while there is significant cross-country variation in the ratio of government financed business R&D to GDP, direct funding has generally become less supportive of R&D over the past few decades (Figure 4). One reason for this trend in some countries is a gradual decline in the intensity of military spending and, hence, government procurement of private defence-related R&D. However, in almost all of these regions, the pace of decline in government support for private R&D has markedly exceeded that in military spending.

Figure 4. **Government financed business R&D**

Proportion of GDP



In general, there appears to have been a decrease in the reliance of governments in the sampled countries on direct support policies in favour of R&D tax incentives. This may reflect a growing acknowledgment that although, in principle, direct funding schemes allow governments to select R&D activities with the highest marginal social returns, in practice, identifying such projects can be complicated by information asymmetries and decisions may be influenced by rent-seeking entities.

The tendency for governments to move away from direct support has been reinforced by a body of empirical research that remains inconclusive as to whether such schemes encourage additional private R&D (David et al., 2000). Instead, it may be that firms substitute public R&D funds for their own or that subsidies feed into higher wages rather than greater innovation intensity (Goolsbee, 1998; Jaumotte and Pain, 2005a).

Whether government assistance is provided through tax incentives or direct funding, there is some evidence that the effect of such R&D policies can be undermined if they are particularly "unstable" – proxied by the standard deviation of policy variables (Guellec and Van Pottelsberghe, 2003). Cost certainty is important for many R&D projects given that they often have a multi-year horizon and the investment decision is difficult to reverse once the project is commenced (Pindyck, 1992). It may be, however, that unstable but predictable policy movements still allow private firms to confidently plan and implement R&D projects. As such, a better proxy for policy instability may be a measure of the frequency with which policy has reversed course, rather than simply an indicator of the dispersion of policy from its mean over a given period.

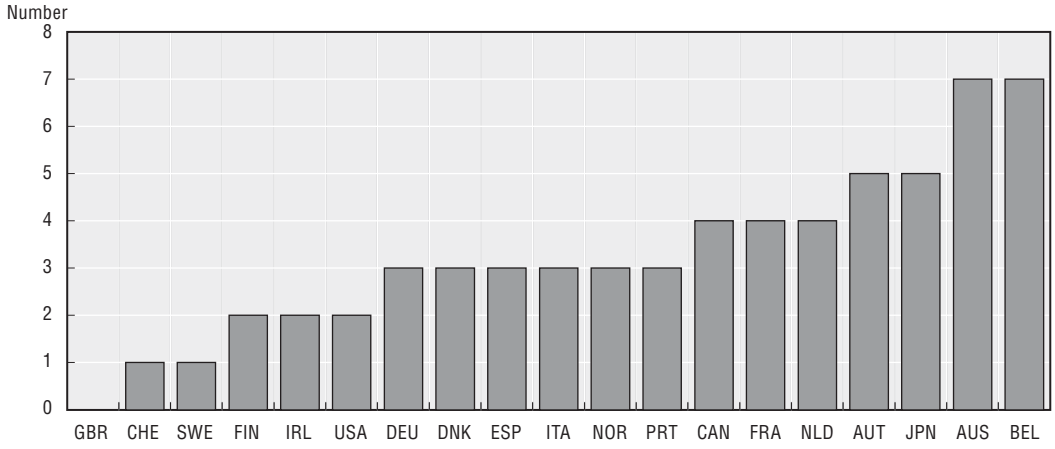
Measuring the unpredictability of R&D tax policy as the number of instances in which the B-index immediately reversed course over the 1981-2008 period highlights significant dispersion across the countries in the sample (Figure 5).[8] Moreover, a simple regression between this variable and the within-country standard deviation of the B-index suggests that the two measures are not statistically related to one another.

2.2.3. Government performed research

Recent country level empirical studies find that non-defence related public research tends to have a neutral effect on business R&D; not encouraging nor substituting for

Figure 5. **Cumulative "reversals" in R&D tax policy**

1982-2008



Source: Authors calculations based on data from the OECD *Science, Technology and Industry Outlook*.

private research (Guellec and Van Pottelsberghe, 2003; Montmartin, 2013). While history provides numerous instances of government research being at the root of revolutionary commercial technologies, the lags can be long and unpredictable and thus difficult to identify empirically (Sheehan and Wyckoff, 2003). A common example is the Internet revolution of the past few decades, which evolved from government investments that were made as far back as the 1960s.

2.2.4. Patent protection

Patent rights temporarily grant the holder the ability to limit others from using an invention in exchange for the innovation being made public. In strengthening such rights, a delicate balance exists between granting firms some market power to incentivise innovation on the one hand, and ensuring that the competitive forces that prevent abuse and also motivate on-going innovation and diffusion of ideas are not stifled on the other. For many years there was a general tendency for countries to increase the amount of protection given to patent holders (Figure 6). However, in the past decade or so, most OECD countries have maintained a relatively stable intellectual property rights (IPR) regime. One consequence of this for empirical analysis is that it may be difficult to identify a causal effect of patent rights (Falk, 2004b).

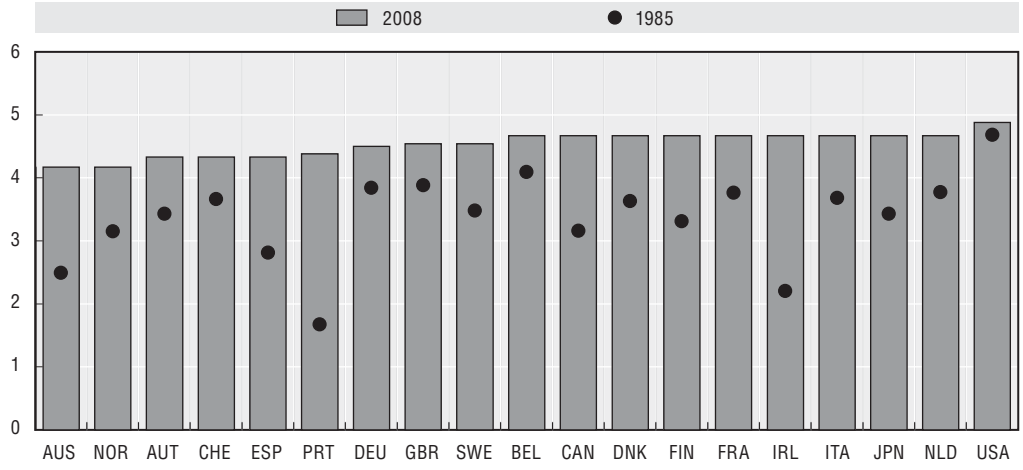
2.3. Framework policies

Framework policies that promote domestic and international competition (Conway et al., 2006; Coe and Helpman, 1995) and adaptable labour markets (Bassanini et al., 2009) have been identified as important for fostering productivity growth. A channel through which this may occur is via the effect on the innovation process. In particular, framework policies may influence the ability and incentive of firms to undertake innovative activity, attract resources to commercialise new ideas, scale back outdated activities and to take advantage of knowledge spillovers from earlier innovations.

2.3.1. Policy affecting competition

The relationship between competition and private sector innovation is likely to vary depending on the particular attributes of the firm or market. While some models advocate

Figure 6. **Patent rights**
Ginarte-Park index of patent protection



Source: Park (2008).

competition on the basis that firms with high market power may be disinclined to pursue innovation that may displace existing rents (Arrow, 1962; Reinganum, 1983), early Schumpeterian growth models (Aghion and Howitt, 1992) highlight that the incentives for innovation (in terms of prospective post-innovation rents) may increase at lower levels of competition. Indeed, recent empirical work indicates heterogeneity across industries. Using panel data from the United Kingdom, Aghion et al. (2005) find that innovation will benefit from a reduction in product market regulation (PMR) in initially high-PMR industries, but may be harmed in those industries that are already highly competitive. Thus, the aggregate impact of PMR on innovation is theoretically ambiguous and becomes an empirical issue.

2.3.2. Employment protection legislation

In allowing resources to flow to their most productive uses, easier employment protection legislation (EPL) is thought to benefit entrepreneurial firms undertaking radical innovations (Saint Paul, 1997, 2002; Bartelsman et al., 2010), which necessitate relatively large employment adjustments (Aghion and Howitt, 1997). This is supported by recent empirical evidence that higher EPL reduces R&D expenditure in innovative industries and that this effect is disproportionally large in sectors that are particularly turbulent (Andrews and Criscuolo, 2013).

Lowering EPL may not be unequivocally good for all firms. Acharya et al. (2010) find a positive relationship between stricter labour laws governing dismissal of employees and patenting, arguing that employees may be incentivised to undertake innovative activities that are value-maximising in the long run if they do not fear losing their job. Similarly, greater employment protection may encourage firms to invest in firm-specific human capital that benefits productivity (Autor, 2003; Wasmer, 2006). Griffith and Macartney (2010) find that multinational firms do more patenting of incremental innovation in relatively high EPL countries and more patenting of radical innovation in low EPL countries. Such heterogeneous effects indicate that it may be difficult to identify the impact of a change in EPL on R&D expenditure or patenting in a cross-country study based on aggregate data.

2.3.3. *Trade policy*

Countries with policies that promote international trade are more exposed to the foreign stock of R&D and new technologies, production processes and organisational methods. Past studies have found that such exposure has a positive relationship with measures of domestic innovation (Jaumotte and Pain, 2005a; Falk, 2004a). Nevertheless, it is possible that access to foreign R&D may provide a disincentive for domestic R&D expenditure and that the two sources of R&D are partly substitutable.

2.3.4. *Financial and bankruptcy policies*

R&D projects often require substantial investments that need to be funded from retained earnings or external sources. However, many innovative firms are small and young with limited internal income or assets and have difficulty obtaining external private finance given that the outputs to innovation are relatively uncertain. Consequently, financial sector policies that influence the availability of credit or equity-financing to businesses may affect the existence and success of innovative firms.

Government policies that relate to bankruptcy procedures are likely to be important for firms engaged in radical innovation activities that expand the technological frontier. Such innovation is synonymous with entrepreneurship and experimentation strategies that imply frequent and repeated failure (Bartelsman et al., 2008). Countries that impose high exit costs for failed firms – through more punishing bankruptcy regimes – may deter experimentation, lowering the capacity for innovation. However, some investor protection is necessary, especially given that the availability of finance may be positively related to the degree of creditor recourse.

2.3.5. *Policy complementarities*

It should now be apparent that there are a range of policy tools that may influence the volume of private R&D, patenting and other types of innovation activity. However, a number of these policy instruments are likely to interact, potentially complementing one another or leading to policy conflicts. One such interaction may arise when the success of a policy in meeting its objective is influenced by the settings of other policy instruments. Parente and Prescott (2000) highlight the case of trade openness which enables the adoption of new innovations from abroad, though such activity may be limited if high barriers to firm entry in the domestic market lead to incumbents that own existing technology resisting the adoption process. Another form of interaction may occur when a policy adjustment creates distortions unless other policy settings are correctly calibrated. For example, an increase in patent rights designed to provide firms an incentive to engage in innovative activities may result in innovators creating monopolistic positions. In this case, it may be important to ensure that IPR policy settings are complemented by pro-competition product market regulations (OECD, 2006; Uppenberg, 2009). Such policy interactions are an issue that has received relatively limited attention in the existing empirical literature but which this paper investigates further.

3. Empirical model

3.1. *R&D model*

The approach to modelling R&D corresponds to that of Jaumotte and Pain (2005a) in the spirit of earlier work by Bloom et al. (2002). This assumes that the stock of real R&D is

one input into a production process approximated by a CES production function. Taking the first order condition that R&D investment is undertaken up to the point where marginal revenue equals marginal cost, the following long-run relationship is derived:

$$\ln(RDS_{it}) = \alpha_i + \beta_i \ln(Y_{it}) + \tau \ln(user\ cost_{it}) + \sum_{j=1}^{n} \varphi_{ji} Z_{jit} + u_{it} \quad [1]$$

For country *i* at time *t*, the R&D stock is determined by real output (*Y*), the real user cost of R&D and a vector of additional influences (*Z*). Long-run constant returns to scale would imply $\beta_1 = 1$.

An empirical specification is then derived explaining privately-funded business R&D, reflecting both short-run dynamics and divergences from an underlying long-run relationship. This acknowledges that it may take time for R&D to react to changes in its determinants as businesses develop R&D plans and then obtain the necessary capital and labour resources. The dynamic non-linear error correction relationship is represented by the equation:

$$\Delta \ln(RDS_{it}) = \alpha_{1i} \Delta \ln(RDS_{it-1}) + \sum_{j=1}^{m} \rho_{ji} \bar{Z}_{jit} + \theta_i \left[\ln\left(\frac{RDS_{it-1}}{Y_{it-1}}\right) - \delta \ln(\text{user cost}_{it-1}) - \sum_{k=m+1}^{n-m} \gamma_{ki} \widetilde{Z}_{kit} \right] + \alpha_i + \alpha_t + \mu_{it} \quad [2]$$

The dependent variable is the stock of business R&D in constant price US dollar purchasing power parity (PPP) terms and excludes expenditures financed by government. The vector *Z* includes *m* variables that may affect the short-run evolution of R&D and *n* variables that explain long-run cross-country differences in R&D (e.g. R&D fiscal incentives, framework conditions). *Y* is real GDP in constant US dollar PPP terms. Long-run constant returns to scale are imposed, so that while the dependent variable is the growth of the R&D stock the long-run parameters ultimately determine R&D intensity. This is reflected in the specification of the error correction term. Factors that vary across countries but not across time are captured through country fixed effects (α_i), while time fixed effects (α_t) are included to absorb common global shocks. The lag length is chosen by running country-specific regressions of a base specification and consulting the Akaike and Schwarz' Bayesian information criterion.[9] The equation is specified as a one-stage error correction procedure instead of an Engle-Granger two-step approach due to the presence of theoretically relevant short-run, often contemporaneous, variables. This specification also allows the flexibility to include a variety of factors believed to share a long-run relationship with the R&D stock, without imposing strong restrictions on the data generating process (De Boef et al., 1999).

While either the flow or the stock of R&D can be modelled as the dependent variable, the primary focus of the analysis is the R&D stock. This owes to the fact that, as outlined in Figure 2, changes in the R&D stock are most relevant for augmenting the pool of existing knowledge and hence affecting productivity growth.

The construction of the real user cost of R&D follows Jaumotte and Pain (2005a) using an approach similar to that pioneered by Hall and Jorgenson (1967):

$$user\ cost_{it} = (B\text{-}index) * (r_{it} + \delta) \quad [3]$$

where r is the long-term real interest rate and δ the depreciation rate on R&D capital, assumed to be 15% per annum in all countries and time periods (Guellec and van Pottelsberghe, 2004; Hall and Mairesse, 1995).[10]

3.2. Patenting model

The model describing the evolution of the number of patents is broadly similar to that for movements in R&D detailed above. In contrast with the R&D model, but in keeping with the depiction of the innovation process in Figure 2, here the dependent variable is based on a flow rather than a stock measure, annual growth in the number of triadic patent families per capita:

$$\Delta \ln\left(\frac{Pat_{it}}{Pop_{it}}\right) = \alpha_1 \Delta \ln\left(\frac{Pat_{it-1}}{Pop_{it-1}}\right) + \sum_{j=1}^{m} \emptyset_{ji} \bar{Z}_{jit} + \alpha_2 \left[\ln\left(\frac{Pat_{it-1}}{Pop_{it-1}}\right) - \alpha_3 \ln\left(\frac{RDS_{it-1}}{Y_{it-1}}\right) - \alpha_4 \ln\left(1 + \frac{RDS^{G}_{it-1}}{RDS_{it-1}}\right) - \sum_{k=m+1}^{n} \partial_{ki} \tilde{Z}_{kit} \right] + \alpha_i + \alpha_t + \mu_{it} \quad [4]$$

The stock of business sector R&D and the stock of non-business sector R&D (RDS^G) are included as regressors (as in Figure 2), reflecting that innovation output is likely to be a function of the knowledge that has built up in an economy over time. The forms in which these terms enter the equation are taken from Jaumotte and Pain (2005a), with their separate inclusion designed to highlight any differential effects of the source of R&D on patenting activity. *Pop* is millions of working age population (15-64) and the *Z* terms represent other short-run ($\bar{Z}$) or long-run ($\tilde{Z}$) influences on patenting including government policies. Indeed, it is possible that policies that influence R&D may have an additional effect on the success of such activity. Country and time fixed effects are also included.

Given the theoretical relationship outlined in Figure 2, a two-stage empirical setup whereby a first stage estimate of the stock of R&D is used to explain variations in a second stage patent stock equation may be appealing. Nonetheless, unreported two-stage least square regressions suggest imprecise parameter estimates in the second stage. This may owe to the first stage equation explaining only a portion of the variation in the R&D stock, leading to changes in R&D that may be relevant for explaining movements in the patent stock being omitted from the model.

3.3. MFP model

In the spirit of past models of productivity growth (Nicoletti and Scarpetta, 2003; Griffith et al., 2004; Conway et al., 2006), the MFP equation is specified to allow for the role of technology transfer as a source of productivity growth for countries behind the technological frontier:

$$\Delta lnMFP_{it} = \delta(\Delta \ln MFP_{Lt}) - \rho \ln\left(\frac{MFP_{it-1}}{MFP_{Lt-1}}\right) + X'_{it}\beta + h_{it} + \alpha_i + \widetilde{\alpha}_t + \varepsilon_{jt} \quad [5]$$

The first term in equation [5] represents MFP growth of the frontier country, while the second term is the productivity gap (in terms of productivity level) and reflects the expectation that productivity growth is higher the further a country is from the frontier. The vector *X* represents the stock of R&D or the number of patents as well as other sources of knowledge diffusion that may affect aggregate MFP. The influence of framework policies can be assessed through interacting policy variables with these indicators as well as the

frontier gap term (though this is not explicitly illustrated in equation [5]). Because the labour input for the calculation of MFP is the number of employed (see the Data Annex for more details), a control variable for labour utilisation is included (*h*) which represents the number of hours worked per employee. Country fixed effects are included and following Bouis et al. (2011), five-year period dummies ($\widetilde{\alpha}_t$) are used to control for common variations in trend productivity growth. Equation [5] can be seen as an error correction model derived from a first-order autoregressive distributed lag process in which the MFP level in each country is cointegrated with that of the leader (Nicoletti and Scarpetta, 2003). The model implies heterogeneity in equilibrium MFP levels, with the productivity gap for each non-frontier country converging to a constant value. As a result, changes in the factors captured by vector *X* impact the steady state levels of MFP across countries, but do not result in permanent changes in MFP growth rates.

The models represented by equations [2], [4] and [5] are estimated across the 19 countries over the 1983-2008 period for the R&D equation and 1986-2008 for the patent and MFP equations.[11] This is the same sample of countries used by Jaumotte and Pain (2005a), with seven additional years of data.

Assumptions of independent and identically distributed standard errors within countries are often violated, leading to the clustering of standard errors at the country level. Such a technique allows for a general form of heteroskedasticity as well as intra-cluster correlation (Primo et al., 2007). The fixed effects estimator is used which pools data across the individual countries and in doing so assumes that the long and short-run determinants of R&D are homogeneous across regions.[12]

4. Data

Descriptive statistics are found in the Data Annex along with details of the construction of the R&D stock and MFP level variables used in the analysis. As mentioned, the MFP measure accounts for the contribution of the stock of physical capital and a labour input that is adjusted for changes in human capital. The main constraints to broadening the scope of the analysis (along the country or time dimension) are insufficient data for either R&D expenditure or the B-index.

4.1. *Indicators of innovation*

The OECD *Main Science and Technology Indicators Database* is the source of constant price R&D expenditure, employment and patent data. R&D expenditure series are available for most countries from 1981 and are deflated using the implicit GDP deflator from the OECD *National Accounts Database*. For patents, the number of triadic patent families by inventor's country of residence is used. As is generally the case (unless specified), linear interpolation is performed in instances of missing observations.

The foreign R&D stock variable is calculated for each country as the trade-weighted R&D of all other countries in the sample, employing the weighting scheme outlined by Lichtenberg and Van Pottelsberghe (1998). The weights are calculated from bilateral trade flow data from the IMF *Direction of Trade Statistics*.[13]

4.2. *Innovation policies*

Government financed business R&D is derived from the OECD *Science, Technology and R&D Statistics Database*. The measure of R&D tax generosity, the B-index, is available for large firms from 1981 with the historical series updated with data from the 2009 OECD

Science Technology and Industry Outlook. In years where data are not available, the series is held constant at the last observed value. In calculating the user cost of R&D capital, the long-term interest rate is measured from secondary market yields of 10 year government bonds and is taken from the OECD *Key Economic Indicators Database*.

The measure of patent rights is the cross-country index updated by Park (2008) which increases at higher levels of patent protection. The index takes into account five aspects of patent laws: extent of coverage, enforcement mechanisms, duration of protection, provisions for loss of protection and membership in international patent agreements. The raw data are available at five-year intervals between 1980 and 2005 and, for the purposes of the empirical analysis, the value is assumed to remain at the 2005 level from 2006 to 2008.

4.3. *Framework policies*

The indicator of PMR is the OECD measure of regulation in energy, transport and communications. This index summarises the regulatory provisions in the telecommunications, electricity, gas, post, rail, air passenger transport, and road freight sectors and has a longer time series than other such indicators. While the measure does not cover all industries, it captures those in which anti-competitive regulation tends to be most concentrated in the sample countries. The "high barriers dummy" appearing in the patent and MFP regressions captures regulatory or administrative barriers to firm entry (Wölfl et al., 2009).[14] The measure of EPL relates to both regular and temporary contracts from the OECD *Employment and Labour Market Statistics* (for further details, see Nicoletti et al., 2000).

The index of banking regulation used in the R&D regressions does not have a time dimension and is taken from De Serres et al. (2006) and compiled from the World Bank's *Bank, Regulation and Supervision Database* that draws on the responses to a survey conducted in 2003. The measure increases with the level of regulation in two broad areas: barriers to competition (relating to regulations for domestic entry, foreign entry, banking activity and government ownership) and stability (relating to accounting standards, auditing requirements, capital adequacy, liquidity and diversification, provisioning, internal management, ownership, discipline and enforcement, deposit insurance and supervisory structure).

The bankruptcy cost variable is a time-invariant indicator of the cost to close a business in 2004 sourced from the World Bank *Doing Business* project.

4.4. *Control variables*

Macroeconomic indicators include GDP growth, calculated from a measure of GDP in constant price US dollar PPP terms sourced from the OECD *National Accounts Database*, and inflation, which is measured by the Consumer Price Index from the OECD *Main Economic Indicators Database*. In the MFP regressions, average annual hours per worker is taken from the OECD *Employment and Labour Market Statistics*.[15]

5. Estimation results

5.1. *Private research and development*

Table 1 presents the results from the estimation of equation [2]. As discussed in Section 3.1, the analysis is focused on the determinants of growth in the R&D stock, though the results of a base specification with the R&D flow as dependent variable is also presented (column 8). The regressions suggest that government innovation policies as well as some framework conditions are important in explaining the stock of private sector R&D.

Table 1. **The determinants of business R&D**

1982-2008	Dependent variable = Δln(R&D stock)							Dependent variable = Δln(R&D flow)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Short-run dynamics								
$\Delta\ln(\text{R\&D stock})_{t-1}$	0.84***	0.85***	0.87***	0.84***	0.85***	0.84***	0.84***	
	(28.28)	(26.01)	(22.95)	(29.54)	(26.94)	(33.22)	(29.06)	
$\Delta\ln(\text{R\&D flow})_{t-1}$								0.28***
								(5.24)
ECM_{t-1}	-0.013***	-0.012***	-0.012***	-0.013***	-0.014***	-0.027***	-0.013***	-0.06***
	(2.84)	(2.60)	(3.64)	(3.09)	(3.31)	(5.54)	(3.18)	(2.88)
ECM_{t-1} * (banking reg.)						0.006**		
						(2.55)		
$\Delta\ln(\text{B-index})_t$	-0.03**	-0.03*		-0.04*		-0.04**	-0.03*	-0.08
	(1.96)	(1.84)		(1.95)		(2.09)	(1.87)	(1.25)
Inflation_{t-1}	-0.13***	-0.12***		-0.14***	-0.13***	-0.13***	-0.14***	-0.50***
	(3.47)	(2.78)		(3.30)	(3.10)	(3.80)	(3.20)	(3.61)
GDP growth_t	0.10**	0.098**		0.11***	0.11**	0.12***	0.11***	0.46***
	(2.36)	(2.24)		(2.65)	(2.34)	(2.88)	(2.71)	(2.66)
Long-run parameters								
$\ln(\text{Government financed private R\&D})_{t-1}$	0.36**	0.41**	0.56**	0.45***	0.34**	0.18**	0.47***	0.40***
	(2.02)	(2.08)	(2.16)	(2.73)	(2.17)	(2.26)	(2.72)	(3.03)
$\ln(\text{User cost})_{t-1}$	-0.98*			-1.15**	-0.80**	-0.49**	-1.00**	-1.06
	(1.65)			(2.06)	(2.08)	(2.44)	(2.04)	(1.60)
PMR_{t-1}		-0.23	-0.26*				-0.14	
		(1.35)	(1.76)				(1.04)	
$\ln(\text{Foreign R\&D stock})_{t-1}$				-1.02*			-0.98*	
				(1.74)			(1.72)	
$\ln(\text{User cost})_{t-1}$ * (policy reversals dummy)					1.33**			
					(2.17)			
Constant	-0.07***	-0.05***	-0.06***	0.02	-0.07**	-0.07**	0.04	-0.51***
	(2.84)	(2.72)	(4.54)	(0.43)	(2.41)	(2.39)	(0.62)	(3.78)
Observations	512	512	512	512	512	512	512	493
Adjusted R^2	0.76	0.76	0.74	0.76	0.76	0.76	0.76	0.29

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Note: t-statistics in parentheses. The regressions include country and time fixed effects and standard errors clustered at the country level. The long-run parameters and associated standard errors are calculated according to the technique outlined by Bardsen (1989). *banking reg* is a time invariant measure of banking regulation. *Policy reversals dummy* is a dummy for those countries in the top 10% of the distribution for reversals in R&D tax policy (Australia and Belgium).

5.1.1. *Innovation policies*

Consistent with past studies, the long-run effect[16] of the user cost of R&D on the stock of private R&D is estimated to be around unity.[17] There is also estimated to be a significant short-run influence of changes in the B-index, though this effect is estimated to be small relative to the long-run impact of the user cost. The coefficient on the user cost term in column 7 implies that a decrease in the B-index by 0.05 units (holding long-term real interest rates and the depreciation rate constant) would encourage an increase in the long-run stock of business R&D by just under 6%. Such an increase in the generosity of R&D tax policy (decrease in the B-index) is equivalent to a shift in policy settings from those in the United States to the more generous ones in Japan in 2008 (for partial effects of the policy variables, see Table 3).

The coefficient on an interaction term between the long-run user cost and a dummy for countries with a very high number of instances in which R&D tax policy has reversed course is estimated to be positive and statistically significant (column 5).[18] The sign and magnitude of the coefficient suggests that countries with very high policy volatility gain negligible benefit from a policy-induced reduction in the user cost of R&D.[19]

In contrast to Jaumotte and Pain (2005a), direct government support is found to be positively related to business-funded R&D in the long-run.[20] This implies that government grant and subsidy schemes are complements rather than substitutes for private sector R&D funding.[21] While there is some instability in the parameter estimates across the specifications in Table 1, the elasticity in column 7 implies that a 10% increase in direct government support for business R&D is met by an increase in the privately funded R&D stock by just under 5% in the long run. To give some perspective, 10% is around the average annual increase in government support across countries during the first six years of the sample. Unreported regressions suggest that, unlike R&D tax incentives, direct support measures are not associated with movements in private R&D in the short run.

Re-estimating the specification in column 1 over the 1982-2001 period, the time frame used by Jaumotte and Pain (2005a), Table 2 highlights that the direct government support variable is the only one to become statistically insignificant.[22] The sensitivity of this parameter to the sample period may reflect variation in the design of direct support policies through time. For example, Bloch and Graversen (2008) note that past government support for R&D was often through contracts where the government would fund as well as purchase the output of firms' R&D activity. This necessitated a sizeable role for government

Table 2. **Robustness of the estimates for the R&D equation**

	Dependent variable = Δln(R&D stock)	
	(1)	(2)
	1982-2008	1982-2001
Short-run dynamics		
$\Delta\ln(\text{R\&D stock})_{t-1}$	0.842**	0.81***
	(28.28)	(18.83)
ECM_{t-1}	-0.013***	-0.022***
	(2.84)	(4.74)
$\Delta\ln(\text{B-index})_t$	-0.03**	-0.04**
	(1.96)	(2.10)
Inflation_{t-1}	-0.13***	-0.15***
	(3.47)	(2.80)
GDP growth_t	0.10**	0.15***
	(2.36)	(2.92)
Long-run parameters		
$\ln(\text{Government financed private R\&D})_{t-1}$	0.36**	0.120
	(2.02)	(0.88)
$\ln(\text{User cost})_{t-1}$	-0.98*	-1.03**
	(1.65)	(1.96)
Constant	-0.07***	-0.11***
	(2.84)	(5.48)
Observations	512	379
Adjusted R^2	0.76	0.75

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Note: t-statistics in parentheses. The regression includes country and time fixed effects and standard errors are clustered at the country level.

in selecting the R&D projects to be undertaken and may have meant that much of the R&D performed was not widely commercially applicable or substituted for privately funded R&D as a result of rent seeking activity. Recently, however, the allocation of support has become more sensitive to market signals, with matching grants (where a firm commits to match, in a given proportion, the direct support received) an increasingly common feature of government funding programmes (see Blanco Armas et al., 2006; Hall and Maffioli, 2008). Given that a matching scheme stipulates input additionality, it may not be surprising that the more recent time period is important to the empirical finding that government financing encourages additional private R&D.[23]

Another factor behind the observed increasing impact of direct government support may be the recent scaling-back of such schemes across countries (discussed in Section 2.2.2 of this paper). Notwithstanding information asymmetries and the power of some rent-seeking firms, in rationalising direct support for private R&D, governments may first scrap those programmes perceived to be crowding out private sector R&D or those deemed to be least effective in encouraging further private innovation. Additionally, the apparent increase in the potency of direct government support may be related to the decline in the share of such spending directed to defence related private R&D.

Focusing on another innovation policy, a tightening of patent rights should normally raise the potential returns to R&D investment for firms (Falk, 2004a; Varsakelis, 2001), but the coefficient on the IPR term is not found to be statistically different from zero (see column 2, Appendix C).

5.1.2. *Framework policies*

The empirical results provide only weak evidence that a decrease in product market regulation has a positive effect on the stock of R&D. While the coefficient on the PMR term is estimated with a negative sign across the specifications in Table 1, the estimate is only weakly significant in one of the three regressions. With higher PMR limiting new entrants, the estimated (albeit at most weakly significant) sign accords with the theoretical prediction of patent race models that incumbents will invest less in R&D than entrants as successful innovations may replace incumbent-owned technology that is relatively profitable (Reinganum, 1983). In contrast to the industry based findings of Aghion et al. (2005), unreported regressions find little evidence at the country level that the innovation effects of a decline in PMR are more modest (or harmful) for countries with initially higher levels of competition.

There is some limited evidence that the financial regulatory environment influences the speed with which movements in the long-run parameters affect R&D. This is tested, and presented in column 6, by including an interaction variable between the ECM term and a time-invariant indicator of banking regulation that captures regulatory measures relating to bank entry and stability (described in Section 4.3). The positive coefficient estimate suggests that the stock of R&D may be slower to adjust to changes in the user cost or direct government support in those countries with stricter banking regulation. In such countries financing options may be fewer, slowing the response of firms to a new policy initiative. However, this result should be treated with caution given that the indicator of banking regulation is only observed at one point in time and a time-varying index of financial reform was not found to have a statistically significant effect on business R&D.[24]

In contrast with the finding of Jaumotte and Pain (2005a), columns 4 and 7 highlight the negative and statistically significant coefficient estimate for the foreign R&D stock. This implies that domestic and foreign R&D are substitutes, a result in keeping with the evidence of Bloom and Griffith (2001) that domestic R&D is positively related to the cost of doing such activity abroad. This idea was further investigated with the inclusion of the foreign user cost as a determinant of domestic R&D.[25] However, the estimated coefficient was not statistically significant despite having the expected positive sign (see column 3 of Appendix C).

EPL is not found to have any discernable effect on business R&D (see column 4, Appendix C). This is unsurprising given the theoretically ambiguous direction of the aggregate-level relationship already discussed.

5.1.3. *Dynamic effects*

The dynamic parameter estimates highlight the importance of macroeconomic factors for the evolution of business R&D in the short run. The positive coefficient on the GDP growth term suggests pro-cyclicality in R&D spending.[26] Higher inflation is estimated to reduce private R&D, possibly because price instability may create uncertainty over the real value of future cash flows and cause firms to defer investment decisions.[27]

The ECM term is significant in all of the specifications in Table 1, giving support for a model that allows short-run adjustments that correct for disequilibrium in the co-integrating relationship with the long-run factors. Consistent with past work (Jaumotte and Pain, 2005a) and the idea that it may take time for firms to obtain the necessary capital and labour resources to markedly change R&D expenditure, the impact of the long-run parameters is estimated to be very protracted. The ECM coefficients in columns 1-5 imply that just 1.2-1.3% of the adjustment of R&D in response to a shock to one of the long-run parameters occurs each year.

The lagged dependent variable is statistically significant in all of the specifications presented in Table 1. While a least squares model with a lagged dependent variable may create bias in the estimated coefficients when the time dimension of the panel is small (Nickell, 1981), the 28 years of data utilised in this model is large relative to most cross-country panel studies. Indeed, re-estimation with the Arellano-Bond estimator (Arellano and Bond, 1991), designed to correct for such bias, suggests that there is no statistically significant bias present in the Table 1 coefficient estimates.

Table 3. **The partial effects of policy determinants on business R&D**

	(1)	(2)	(3)	(4)	(5)
	Sample mean	Standard deviation	Assumed Δ in policy	Long-run coefficient	Partial effect (%)
B-index	0.94	0.05	-0.05	-1.00	5.8
Government financed business R&D	2 246	657	230 (+10%↑)	0.47	4.8
PMR	3.55	1.18	-0.35	-0.14	4.8

Note: The long-run coefficients are taken from column 7 of Table 1. The magnitude of the assumed change in policy is in line with frequently-observed annual policy movements across countries through the sample period. The standard deviation is calculated as the average of the country standard deviations (rather than the standard deviation of the pooled sample). For the B-index the long-run coefficient is for the user cost term but the partial effect is calculated based on a rise in the B-index, holding all other components of the user cost constant.

5.1.4. *Unreported results*

Unreported results of the analysis include further exploration of the role of innovation policies on R&D. No evidence was found that public research, either performed by the higher education sector or government labs, encouraged private R&D. This may reflect the broad range of research that is undertaken by such institutions, a large part of which is not immediately commercially applicable. Differential effects from the design of R&D tax incentives were explored by interacting country dummies for the type of R&D tax regime (e.g. incremental or volume-based tax credits and whether tax incentives existed for collaboration between business and public institutions) with the user cost variable. However, the estimated effects were not statistically significant. This is likely to partly owe to differences in the design of R&D tax regimes that cannot be adequately summarised by a time-invariant dummy.

Following some past evidence of a positive relationship between domestic R&D and foreign direct investment (FDI; Kuemmerle, 1999), a variable reflecting policy barriers to FDI was included.[28] While the estimated coefficient was not statistically significant, such policies may be most relevant for certain types of firms or industries and best investigated using more disaggregated data.

While differences in industry composition are likely to partly explain cross-country variation in R&D intensity, such differences should be captured by the country fixed effects. Although country dummies will not control for the impact of changes in industry composition through time, several measures of industry composition were included in the R&D regressions but were not found to be statistically significant.[29]

5.2. *Patenting*

The estimation results from the patent model (equation [4]) provide evidence of the posited relationship between the stock of R&D and the flow of new patents outlined in Figure 2 as well as the influence of innovation-specific and framework policies.

5.2.1. *R&D as a determinant of new patents*

The stocks of business and non-business R&D are estimated to have a positive and statistically significant influence on the number of patents per capita (see Table 4). Evaluated at the sample mean with the estimated coefficients from column 3 of Table 4, a 1% increase in business R&D will increase the number of new patents per capita by just under 1% in the long-run, while a 1% increase in non-business R&D will lead to a 0.4% rise. The differential effect of business compared with non-business R&D is consistent with the finding of de Rassenfosse and van Pottelsberghe (2009) that the number of patents per researcher increases with the share of total R&D performed by the business sector.

Column 7 presents a specification where the distinction between business and non-business sector R&D is relaxed, with the estimated coefficient suggesting that a 1% rise in the total R&D stock is associated with a 0.6% increase in patenting in the long-run. When reporting such results, the potential for reverse causality should be acknowledged as some "development" may occur after a patent application is made (Czarnitzki et al., 2008).

5.2.2. *Innovation policies*

There is evidence that tax incentives for R&D have a positive effect on patenting, even after controlling for the impact of R&D. The estimated long-run elasticity in column 3 of

Table 4. **The determinants of patents per capita**

1986-2008	Dependent variable = Δln(Patents per capita)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Short-run dynamics							
$\Delta\ln(\text{Patents per capita})_{t-1}$	0.025						
	(0.22)						
ECM_{t-1}	-0.49***	-0.51***	-0.51***	-0.46***	-0.47***	-0.52***	-0.48***
	(3.93)	(6.75)	(7.04)	(6.19)	(6.49)	(7.15)	(7.40)
$\Delta\ln(\text{B-index})_t$	-0.62**	-0.55**				-0.53**	-0.50**
	(2.17)	(2.24)				(2.22)	(2.22)
Inflation_{t-1}				-1.52*			
				(1.92)			
Long-run parameters							
$\ln(\text{Business R\&D/GDP})_{t-1}$	1.40***	0.99***	0.93***	1.30***	1.54***	1.14***	
	(3.34)	(3.30)	(3.05)	(3.08)	(3.63)	(3.42)	
$\ln(1 + \text{Non-business R\&D/Business R\&D})_{t-1}$	1.21***	0.90**	0.89**	1.36***	2.18***	1.23***	
	(2.84)	(2.13)	(2.13)	(3.71)	(6.70)	(3.06)	
$\ln(\text{Total R\&D})_{t-1}$							0.64***
							(3.11)
$\ln(\text{B-index})_{t-1}$	-0.95***	-0.82***	-0.44**			-0.69***	-0.74***
	(3.24)	(3.31)	(2.39)			(2.97)	(4.44)
$\ln(\text{IPR})_{t-1}$		0.84***	0.99***			0.72**	0.99*
		(2.68)	(3.01)			(2.11)	(1.77)
PMR_{t-1}			-0.09**			-0.07*	-0.10**
			(1.99)			(1.70)	(2.54)
$\ln(\text{Foreign R\&D stock})_{t-1}$				0.33**	0.38***		
				(2.35)	(2.93)		
$\ln(\text{Foreign R\&D stock})_{t-1}$ * (High barriers dummy)					-0.30***		
					(2.68)		
$\ln(\text{IPR})_{t-1}$ * (High barriers dummy)						-0.78**	
						(2.50)	
Constant	3.33**	2.36***	2.27***	1.61	2.16*	2.97***	-2.34
	(2.51)	(3.04)	(3.055)	(1.56)	(1.68)	(2.89)	(1.61)
Observations	418	418	418	418	418	418	418
Adjusted R^2	0.24	0.26	0.25	0.23	0.25	0.27	0.24

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Note: t-statistics in parentheses. The regressions include country and time fixed effects and standard errors clustered at the country level. The long-run parameters and associated standard errors are calculated according to the technique outlined by Bardsen (1989). *High Barriers dummy* equals one for those countries with average barriers to entry over the sample period in the top 50% of the distribution. These countries are Austria, Belgium, Switzerland, Germany, Spain, Finland, France, Italy and Japan.

Table 4 implies that the flow of patents per capita increases by around 2.5% following an increase in the generosity of R&D tax incentives equivalent to a 0.05 unit decline in the B-index (for partial effects of the policy variables, see Table 5). Such a reduction corresponds to the average change in the B-index in the sampled countries since 2000. Furthermore, columns 1, 2, 6 and 7 of Table 4 suggest that the B-index has a short run effect on patenting activity.

The empirical link between R&D tax incentives and patenting may reflect that the internal funds available to a firm for developing patents or for undertaking non-R&D expenditures that contribute to innovation are influenced by the size of R&D tax incentives received. In addition, some countries allow patent-related expenditures to be eligible under R&D tax incentive schemes (Warda, 2006).[30] One of the risks identified with introducing an R&D tax credit is that firms reclassify unrelated operating expenses as R&D

Table 5. **The partial effects of policy determinants on patents per capita**

	(1)	(2)	(3)	(4)	(5)
	Sample mean	Standard deviation	Assumed Δ in policy	Long-run coefficient	Partial effect (%)
B-index	0.94	0.05	-0.05	-0.44	2.5
IPR	4.10	0.50	0.15	0.99	3.5
PMR	3.55	1.18	-0.35	-0.09	3.1

Note: The long-run coefficients are taken from column 3 of Table 4. The assumed change in policy is in line with the magnitude of observed annual policy movements through the sample period. The standard deviation is calculated as the average of the country standard deviations (rather than the standard deviation of the pooled sample).

to reduce their tax bill (see Hall and Van Reenen, 2000). If such activity was pervasive, it is likely that – holding R&D constant – a decrease in the B-index (i.e. an increase in the generosity of R&D tax incentives) would be associated with a fall in patenting activity. Instead, these results seem to suggest that an increase in R&D tax generosity will make R&D more productive in terms of patent outcomes.

The results indicate that more protective domestic IPR policy is associated with a higher number of triadic patents. This link is somewhat surprising, as triadic patents are governed by IPR settings in the United States, Japan and the EU and not those in the country of the inventor. However, an indirect relationship will exist to the extent that the number of triadic patents filed by a country is positively related to the number of domestic patents and these respond to stronger domestic IPR protection. The coefficient in column 3 suggests that an increase in patent protection equivalent to Portugal moving IPR policy settings in line with Britain in 2008 (an increase in the index of 0.15) is associated with a 3.5% increase in the number of triadic patents.[31] A natural question is the extent to which an IPR-induced increase in patenting indicates more innovation or simply a higher propensity for firms to file patent applications. The fact that the IPR term was not statistically significant in the R&D regressions does not lend support to the idea that strengthening domestic patent rights will increase innovation intensity. That said, it is possible that strengthening IPR encourages greater innovation outside of formal R&D programmes.

As discussed earlier, a key aim of this paper is to investigate any policy complementarities or trade-offs that may exist. With this objective in mind, an interaction term between the IPR indicator and a dummy variable for those countries with above-average barriers to firm entry is included in the patent regressions. The estimated coefficient (column 6) is negative and statistically significant, which suggests that tightening IPR has a weaker impact on patenting in countries with high barriers to entry. Given that raising patent rights entails granting patent holders an (albeit temporary) increase in market power, such a change in policy may reinforce the high barriers to firm entry that exist in some countries.[32] Young firms are often an important source of new ideas (Akcigit and Kerr, 2010) and the inability of firms to gain access to markets in such countries may be reflected in a lower number of patents. Indeed, the size of the estimate on the interaction term implies that the positive effect on patenting of an increase in IPR is largely offset in those countries with above average barriers to firm entry.

5.2.3. *Framework policies*

Tighter anti-competitive regulation as represented by an increase in PMR is found to have a direct negative impact on patenting, suggesting a role for pro-competition policies

in encouraging innovation (column 3). A hypothetical reduction in regulation in Finland in 2008 equivalent to the PMR indicator falling to the sample average in that year (a decline of 0.35) is estimated to result in a 3% rise in patents per capita.[33]

The importance of international knowledge spillovers is highlighted by the finding that greater exposure to foreign R&D is associated with higher domestic patenting (column 4). This contrasts with the evidence from the R&D regressions and indicates that while greater access to foreign R&D may cause some firms to scale back domestic R&D expenditure, knowledge flows from abroad have a positive effect on innovation outputs given a constant stock of domestic R&D. The estimated coefficient implies that an increase in exposure to trading partner's R&D stocks from the average level in Spain (around the sample average in 2005) to the higher level in Canada (corresponding to the 75th percentile) would boost patents per capita by around 20% in the long run.[34]

The results also suggest that policies affecting competition and trade can be complementary in fostering innovation output. Column 5 presents the findings from a specification in which the high barriers to entry dummy is interacted with the measure of the foreign R&D stock. The estimated coefficient on the interaction term is negative, indicating that the positive effect on domestic patenting of higher exposure to foreign R&D is almost completely offset in those countries where high barriers to firm entry exist. As discussed, this may reflect incumbent firms with monopoly rights over existing technologies resisting the adoption of new innovations from abroad.

5.2.4. *Dynamic effects*

The estimated coefficient on the ECM term across the specifications in Table 4 suggests that 50% of the long-run effect of a change in a policy determinant will occur in the following year, with 50% of the remaining effect being felt in the year after and so on. This is a much faster response to movements in the long-run parameters than estimated in the R&D regression, suggesting markedly shorter lags of policy in affecting patenting activity than R&D. This is not entirely surprising given that once the R&D stock is determined it is likely to require less planning and resources for a firm to increase or decrease the volume of patent applications (and, consequently, grants) than to significantly adjust the volume of R&D expenditure.

There is very weak evidence that inflation has a short-run negative effect on patenting activity above and beyond its impact on R&D. However, the variable is only statistically significant in the specification in column 4. The lagged dependent variable is not found to be statistically significant and is subsequently omitted from all the specifications presented from columns 2 to 7.

5.2.5. *Unreported results*

The impact of absorptive capacity (proxied by average years of schooling or the ratio of business researchers to industry employment) on patenting was also investigated, but the estimated coefficients on the relevant variables were statistically insignificant. In common with the R&D regressions, there was no discernable effect of some framework policies such as those relating to EPL and bankruptcy regimes on patenting at the country-level.

The determinants of the stock (instead of the flow) of patents may also be of interest, particularly if productivity gains derive from the cumulated pool of publicly available patented technology rather than only new innovations (as in Figure 2). Reassuringly, re-

estimating the regressions with the patent stock as the dependent variable does not change any of the policy conclusions arising from the results in Table 4 (see Appendix D for results from the patent stock regressions).[35]

5.3. *Multifactor productivity*

As suggested in Figure 2, a key reason for encouraging innovation is the belief that such activity will result in productivity gains. Innovations developed either domestically or from abroad may be important in this process. The empirical framework in this paper examines both of these channels, investigating the link between indicators of domestic innovation (the stock of R&D and patents) and productivity, as well as the role of productivity developments in technological frontier countries that may diffuse to non-frontier countries. The transmission of innovations – from either domestic or international sources – into higher productivity can be influenced by government policy settings, as illustrated by the empirical results.

5.3.1. *The impact of R&D and patenting on productivity growth*

Despite its drawbacks as a proxy for innovation, the number of domestic patents are found to be positively associated with MFP growth (columns 1 to 3).[36] This is the case whether a stock or flow measure of patents is used as the regressor, but relies on the presence of Japan in the sample (for a robustness discussion see Section 5.4).[37] Consistent with the process outlined in Figure 2, the results in Table 6 relate to the patent stock (see Appendix E for results for patent flows) and reflect not only the direct productivity effect from patented innovations but also any effect of knowledge spillovers on follow-on innovators that can access relevant technical information once patents are granted. Nonetheless, the results of the specification in column 2 suggest that the positive effect of higher patenting on MFP is lessened when PMR is relatively high. This may reflect the fact that stricter regulation can impede the efficient allocation of resources (Andrews and Cingano, 2012a) and thus make it more difficult for firms to commercialise and maximise the returns from new technologies.

The role of absorptive capacity in explaining MFP growth is investigated in column 3 through the introduction of an interaction term between the number of patents and the share of business enterprise researchers in employment. The coefficient is estimated to be positive and statistically significant, suggesting that an increase in patenting has a larger effect on MFP growth when the proportion of researchers is relatively high. This is consistent with past work (Griffith et al., 2004) and highlights the importance of absorptive capacity for the proliferation of knowledge spillovers once a new idea is made public.

The stock of business R&D to GDP is estimated to have a positive effect on MFP growth (column 4).[38] Given that part of the private labour and capital resources devoted to R&D are included as production factors in the calculation of MFP, a positive statistical relationship suggests an excess return to business R&D beyond normal remuneration and the presence of positive externalities, or spillovers (Guellec and van Pottelsberghe, 2004). It must be acknowledged that, especially in the case of multinational enterprises, domestic R&D may benefit entities outside the domestic economy and that such activity is not explicitly captured in the empirical model. The estimated link between business R&D and domestic productivity is statistically significant, though the magnitude is in the lower range of estimates from the existing literature (for a review, see Parsons and Phillips, 2007).[39] There is no evidence of a positive impact of R&D performed outside the business sector on MFP:

Table 6. **The estimated effect of knowledge flows for MFP growth**

1986-2008	Dependent variable = d.log(MFP)																
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
MFP gap from frontier $_{t-1}$	-0.03**	-0.03**	-0.03**	-0.03**	-0.03**	-0.02	-0.02	-0.03**	-0.09***	-0.03**	-0.04**	-0.02	-0.02	-0.05*	-0.03**	-0.03**	-0.03**
	(2.47)	(2.24)	(2.40)	(2.34)	(2.15)	(1.46)	(1.33)	(2.11)	(3.69)	(2.11)	(2.29)	(1.49)	(1.44)	(1.75)	(1.88)	(1.80)	(1.92)
MFP growth of frontier country$_t$	0.31***	0.30***	0.30***	0.31***	0.34***	0.32***	0.31***	0.32***	0.31***	0.31***	0.31***	0.32***	0.32***	0.32***	0.32***	0.32***	0.32***
	(8.22)	(7.81)	(7.82)	(8.13)	(8.84)	(8.06)	(8.07)	(9.60)	(8.37)	(8.18)	(8.39)	(8.63)	(8.48)	(8.42)	(8.18)	(8.12)	(8.48)
Patent stock $_{t-1}$	0.005*	0.008***	0.008**														
	(1.65)	(2.68)	(2.55)														
PMR $_{t-1}$		-0.002	-0.002	-0.001	0.0004	-0.001	-0.001	0.001	0.003	-0.001	-0.0002	-0.001	-0.001	-0.002			
		(1.24)	(1.11)	(0.43)	(0.27)	(0.28)	(0.60)	(0.47)	(1.38)	(0.35)	(0.11)	(0.55)	(0.58)	(0.93)			
(PMR)$_{t-1}$ * (Patent stock)$_{t-1}$		-0.012***															
		(4.08)															
Hours worked$_t$	0.11***	0.10**	0.10***	0.09***	0.10***	0.11***	0.10***	0.09***	0.10**	0.09***	0.09***	0.07**	0.07**	0.07**	0.10***	0.10***	0.12***
	(2.93)	(2.87)	(2.85)	(3.03)	(3.26)	(3.38)	(3.01)	(2.82)	(2.79)	(3.02)	(2.83)	(2.57)	(2.49)	(2.45)	(3.69)	(3.42)	(4.18)
Business researchers/Total employment$_{t-1}$	0.002***	0.002***	0.002***														
	(2.86)	(2.86)	(2.93)														
(Business res./Total empl.)$_{t-1}$ * (Pat. stock)$_{t-1}$			0.008***														
			(3.39)														
Business R&D stock/GDP$_{t-1}$				0.33**				0.26	0.25*	0.33***	0.28**	0.32***	0.32***	0.26**			
				(2.54)				(1.51)	(1.86)	(2.64)	(2.36)	(2.86)	(2.87)	(2.45)			
Foreign R&D stock/GDP$_{t-1}$					0.007			0.005									
					(1.53)			(1.11)									
Government basic research stock/GDP$_{t-1}$						0.79		0.51									
						(1.38)		(0.74)									
Higher education research stock/GDP$_{t-1}$							0.77	0.47									
							(1.22)	(0.84)									
(PMR) $_{t-1}$* (MFP gap from frontier) $_{t-1}$									0.015***								
									(2.64)								
Trade openness $_{t-1}$										0.005	-0.02						
										(0.30)	(1.14)						
(Trade openness)$_{t-1}$ * (MFP gap from frontier)$_{t-1}$											-0.11***						
											(4.91)						
Exposure to the frontier/GDP$_{t-1}$												0.05*	0.07*	0.07*			
												(1.75)	(1.97)	(2.09)			
(Expo. to the frontier/GDP)$_{t-1}$ * High bar. dummy													-0.10**	-0.11***			
													(2.11)	(2.69)			
(Bankruptcy cost) * (MFP gap from frontier)$_{t-1}$														0.004*			
														(1.66)			
B-index $_{t-1}$															0.02		
															(1.19)		
Government R&D subsidies$_{t-1}$																0.0000	
																(0.88)	
IPR $_{t-1}$																	0.004
																	(0.67)
Constant	-0.19***	-0.17***	-0.17***	-0.18***	-0.19***	-0.20***	-0.18***	-0.20***	-0.21***	-0.18***	-0.16***	-0.14***	-0.13***	-0.13**	-0.20***	-0.18***	-0.22***
	(3.10)	(3.02)	(3.04)	(3.22)	(3.72)	(3.39)	(3.54)	(3.68)	(3.10)	(2.97)	(2.61)	(2.67)	(2.62)	(2.54)	(3.71)	(3.47)	(4.00)
Observations	428	428	428	428	428	428	428	428	428	428	428	428	428	428	428	428	428
Adjusted R^2	0.26	0.27	0.27	0.27	0.26	0.25	0.25	0.27	0.28	0.26	0.27	0.28	0.28	0.28	0.25	0.25	0.25

*** p < 0.01; ** p < 0.05; * p < 0.10.

Note: t-statistics in parentheses. The regressions include country and time fixed effects, and standard errors are clustered at the country level. *High Barriers dummy* is a dummy for those countries with average barriers to entry over the sample period in the top 50% of the distribution (Austria, Belgium, Switzerland, Germany, Spain, Finland, France, Italy and Japan). The variables used to construct the interaction terms for PMR/patents and business researchers/patents are deviations from the sample mean, so that the estimated coefficients relate to marginal effects evaluated at the mean value. Similarly, the trade openness variable in the Trade Openess/MFP gap interaction term is de-meaned.

the estimated coefficients on terms relating to the stock of government basic research (column 6) and higher education R&D (column 7) are not statistically significant. It is likely that these types of R&D affect productivity with long and variable lags since they are not directly performed by the business sector, although experimenting with longer lag structures did not uncover a stronger statistical relationship.[40] This is consistent with much of the existing empirical literature that finds the productivity return to many forms of publicly financed R&D to be near zero (Sveikauskas, 2007). While the estimation also found no direct impact of the foreign R&D stock on MFP growth (column 5), much of this effect may be captured in the terms relating to the MFP gap and MFP growth of the frontier country (see below).

5.3.2. *The effect of knowledge flows from technological leader countries*

Highlighting the importance of the diffusion of innovations from abroad, MFP growth in the frontier country is found to have a positive and significant effect on MFP in less productive countries.[41] Additionally, the estimated coefficient on the MFP gap term indicates that technological adoption is more beneficial for countries that are further behind the productivity frontier. While of similar magnitude to estimates in past work (Griffith et al., 2004; Nicoletti and Scarpetta, 2003) the coefficient on the gap term is not statistically significant in four of the 17 specifications, which may reflect a degree of multicollinearity with some of the other regressors.

The estimation results also provide some evidence that international knowledge spillovers can be influenced by government policy. The coefficient on the interaction term in column 9 indicates that the speed of convergence to the frontier will be slower for countries that have higher levels of product market regulation. This is consistent with the empirical evidence in Section 5.2.3 of this paper and past research (Conway et al., 2006) that suggests that incumbents with high market power may resist the adoption of new innovations from abroad.

The results also suggest that productivity convergence will be faster for countries that are more open to trade (column 11). Greater trade openness may assist the diffusion of foreign knowledge through the ideas embodied in imported goods and is often positively related to other factors that benefit productivity growth such as foreign direct investment (Seim, 2009) and international labour mobility (Ortega and Peri, 2012).[42]

A number of past studies have explored the importance of geographical proximity to technological leaders in explaining knowledge spillovers and the innovation intensity of non-frontier entities (e.g. Audretsch and Feldman, 2003). Here, this spatial dimension is represented by a variable reflecting geographical distance from a country's capital city to the capital city of the frontier country. A country is a very broad geographic unit for which to investigate spatial knowledge spillovers, with existing studies more inclined to focus on cities or even postal codes. Nevertheless, the results in column 12 indicate that countries with higher geographic exposure to the technological leader exhibit higher productivity. This effect is dependent on framework policies though, with the benefits of spatial knowledge flows completely offset in economies with high barriers to firm entry that restrict the ability of young firms to build on the ideas generated in neighbouring countries (column 13).

There is also weak support for the idea that government policies which increase the cost of firm exit may cause resources to be trapped in inefficient firms that would otherwise be released to businesses that are relatively adept at technological adoption

(Bartelsman et al., 2008).[43] This is illustrated by the positive coefficient on the "bankruptcy cost" interaction term in column 14 which suggests that technological adoption from abroad is slower in countries with more punishing bankruptcy regimes.

5.3.3. *Innovation policies and MFP growth*

Columns 5, 16 and 17 of Table 6 present the results when policy instruments that appear to be influential in encouraging private R&D and patenting activity are included in the MFP regression. Increases in R&D tax incentives, direct government support and patent rights do not appear to have a direct positive effect on MFP growth, despite the regressions not controlling for the stock of patents or domestic business R&D.[44] This is somewhat surprising given that these policy variables were found to affect the measures of innovative activity that are significant determinants of productivity growth (in Sections 5.1 and 5.2). Possible explanations for the lack of a statistical relationship between innovation policies and MFP may be that:

- The indirect nature of the relationship means that the estimation is more likely to be affected by measurement error.
- The types of R&D or patents that these policies encourage are not positively related to productivity. Fiscal incentives induce the marginal project that will have a relatively low value to businesses and perhaps also to society. This would be surprising in the case of direct support measures that are thought to be targeted at R&D activities with high social worth, but may highlight the information asymmetries that are often seen as a reason against such government intervention. In relation to IPR, the lack of a significant finding may suggest that tightening patent rights causes an increase in firm's propensity to patent without significantly increasing innovation intensity.
- Although the activity induced by innovation policies may benefit productivity, such policies may also have unintended detrimental effects for MFP. As discussed, tightening patent rights involves affording some market power to innovators which can have negative implications for competition. Similarly, recent work suggests that R&D tax incentives can protect incumbents to the detriment of potential entrants, slowing down the reallocation process (Bravo-Biosca et al., 2013).
- Fiscal incentives for R&D entail an opportunity cost in the form of the welfare-enhancing projects that could otherwise be funded by the government or the private sector activities suppressed by any additional taxation levied on firms to pay for incentives. If the foregone projects are as beneficial to productivity as the innovation encouraged by fiscal incentives, there should be no identifiable effect of these R&D policies on aggregate productivity growth.

To gain a better understanding of whether innovation-specific policies have an overall benefit to society, a full net welfare analysis of the policy measures is required. While such an exercise requires reliable cross-country estimates for a number of parameters that are outside the scope of this paper, Box 1 outlines the basic framework for such an analysis for fiscal support measures. This makes clear that innovation policies are not necessarily welfare-enhancing, and whether they accrue a net social benefit will depend to a large extent on the magnitude of the spillovers from induced innovation activity and the distortions that such policies and their financing create. The relative dimensions of these effects are likely to depend on a variety of policy and institutional factors, motivating further investigation at the individual country level. One example of this type of analysis

for the Canadian federal tax credit for scientific research and experimental development finds a net economic gain from the policy (Parsons and Phillips, 2007), though this conclusion is fairly sensitive to a number of assumptions.

Box 1. **A framework for calculating average welfare gain/loss from an R&D tax incentive/subsidy**

A positive relationship between an R&D tax incentive or subsidy and measures of innovation or MFP growth does not necessarily indicate that the policy provides a net welfare gain to society. Such policies raise costs as they must be financed through an increase in tax revenue that will impose an excess burden on the economy by creating distortions in the allocation of resources. Furthermore, the schemes result in administration costs for the government and compliance costs faced by participating firms. Given that firms would not undertake the induced R&D in the absence of subsidisation, it is very unlikely that the private rate of return to such projects will outweigh these manifold costs. Thus, for the policy to be welfare-enhancing overall, there must be significant positive externalities in the form of spillovers from the R&D activity. Parsons and Phillips (2007) incorporated each of these factors into a partial equilibrium model for evaluating the welfare gain or loss from an R&D policy. This is summarised by the following equation:

$$\Delta W = \Delta PS + E - T - (A - C)$$

where W is total welfare change per dollar of subsidy, PS is the net welfare change to producers (encapsulating the increase in producer surplus for subsidised firms less the cost to other firms that must pay for the subsidy through higher taxes), E is the externality resulting from the induced R&D, T is the marginal excess burden incurred from distortionary taxation and A and C are the administration and compliance costs of the scheme, respectively.

5.4. *Robustness*

Robustness tests were undertaken on the results for the R&D, patent and MFP regressions. In general, these reveal that the sign and magnitude of the estimated coefficients are not sensitive to slight modifications in the time period or the countries included in the sample, though such sample variations can cause some coefficients to become statistically insignificant at conventional levels. This should be kept in mind when interpreting the estimation results, especially given that such robustness tests are not highly demanding.

Estimating the specification for the R&D equation presented in column 4 of Table 1 but dropping one country at a time, the user cost becomes marginally insignificant in two (of 19) regressions, though in these instances the coefficient is significant at the 12-13% level. The estimated coefficient for the foreign R&D stock appears to be the least robust of the long-run parameters, with the estimate statistically insignificant in nine of the 19 regressions.[45] The coefficient on the direct government support variable is significant regardless of which country is omitted. However, as illustrated in Table 2 and discussed above, the coefficient on this variable is the only one that becomes statistically insignificant when the time period is shortened considerably.

In the patent regressions, small variations in the length of the sample have no impact on the statistical significance of the coefficients in the specification presented in column 3 of Table 4. While the terms relating to non-business sector R&D, the B-index and PMR

become statistically insignificant in a few instances when one country is dropped at a time, the business R&D term is statistically significant in all of the 18-country regressions.[46]

All of the variables in the MFP specifications presented in columns 2 and 4 of Table 6 remain statistically significant in regressions that exclude four years of observations from either the beginning or the end of the sample. The terms relating to the MFP gap, growth of the frontier country, business R&D and the various control variables remain statistically significant when any one country is excluded from the sample. However, the patents variable (along with the patents/PMR interaction term) becomes statistically insignificant when Japan is dropped. Applying the estimated coefficients from Table 6 to country-specific parameter values within-sample, the model cannot explain the extent of the recent MFP slowdown in a number of countries (highlighted in Figure 1, Panel A). This can be partly attributed to utilising coefficients that are the average effect across countries and any future analysis at the individual country level is likely to be helpful in better determining the magnitude of productivity effects by region.

6. Conclusion

This paper explores the policy determinants of business R&D and patenting activity and establishes an empirical link between such measures of innovation and productivity growth at the country level. The effect of various policies on the diffusion of knowledge has also been examined and where possible, any consequences stemming from the interaction of government policies highlighted. The results suggest that innovation specific policies such as R&D tax incentives, direct government support of business R&D and patent rights achieve their objective of encouraging private sector innovative activity (insofar as such activity is adequately proxied by business R&D and patent count measures). However, no evidence of a direct effect of such policies on aggregate productivity growth could be found. Given that the economic impact of innovation policies may be heterogeneous across countries, any future attempt to collect cross-country data on the design of these measures would be informative.

A recurring theme through the findings is the importance of well-designed framework policies relating to competition in incentivising innovation and the returns to new ideas. While there is evidence that lower PMR may have a positive effect on private innovation activity, the results also suggest that pro-competition policies are important for the effective diffusion of knowledge from both domestic and overseas sources.

Although using highly aggregated data is preferable for assessing the spillovers to innovation, the estimated effects presented in this paper are averages across countries, sectors and firms. As such, the findings will benefit from any further work using industry or firm level data that can explore the channels through which these aggregate effects become apparent.

Notes

1. Countries included are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and the United States.
2. A triadic patent family is a series of patents for the one invention filed at the European Patent Office and the Japan Patent Office, and granted at the US Patent and Trademark Office. These are generally high value patents less affected by patenting propensity than other measures (de Rasenfosse and van Pottelsberghe, 2008).

3. This is the case in 17 of the 19 countries even if the post-2007 financial crisis period is excluded from the calculation.

4. Innovation is the result of a range of activities including R&D, informal learning by doing, non-R&D knowledge-building expenditures as well as the registration and commercialisation of new ideas. While R&D expenditure and the number of new patents are the most suitable proxies for innovation when making comparisons across time and countries, they do not capture the full range of such activity and can be affected by factors outside the innovation process. For example, a patent does not reflect the quality of new ideas and can be lodged for strategic purposes such as earning licensing revenue, for increasing the chance of attracting capital or as a negotiation tool with competitors and collaborators (Danguy et al., 2009).

5. This seeming inconsistency may also have to do with the fact that the knowledge-based capital thought to underlie innovation comprises a broader range of assets than just R&D and patents.

6. The estimated relationship between R&D and patenting tends to be weaker in studies utilising within-firm or industry level data (Czarnitzki et al., 2008; Hausman et al., 1984) than in a number of cross-country panel studies (de Rassenfosse and van Pottelsberghe, 2009; Falk, 2004b; Jaumotte and Pain, 2005a; Sanyal and Jaffe, 2005). This may reflect the presence of spillovers to R&D that are better captured at higher levels of data aggregation.

7. The amount of R&D may also influence the ability of firms to absorb new knowledge (i.e. Griffith et al., 2004) and hence the productivity effects.

8. The tax policy reversals variable is calculated by tallying the number of times that a country implements more generous R&D tax policy immediately after implementing less generous policy or *vice versa* over the sample period.

9. This process is also followed for choosing the lag lengths in the patent and MFP models.

10. See Appendix A for a more detailed decomposition of the user cost term. Some recent work has suggested that R&D depreciation rates are higher than 15% (see Li, 2012). However, a sensitivity analysis indicates that the conclusions from the estimation results below are not highly sensitive to the chosen depreciation rate.

11. Patent data prior to 1986 from the European Patent Office are thought to be unreliable (for details, see Jaumotte and Pain, 2005a). The use of starting values in the calculation of the R&D stock (for details of the construction, see the Data Annex) allow the R&D regressions to be run over the 1983-2008 period regardless of whether the dependent variable is the change in the stock or flow of R&D.

12. The assumption of slope homogeneity can lead to the fixed effects estimator producing inconsistent and potentially misleading results. Alternatively, the mean-group (MG) estimator (Pesaran and Smith, 1995) allows the intercepts, slope coefficients and error variances to differ across countries or the pooled mean-group (PMG) estimator (Pesaran, Shin and Smith, 1997) allows heterogeneous short-run coefficients, intercepts and error variances but constrains the long-run coefficients to be equal. However, conducting Hausman tests on the coefficients of the base R&D and patent equations estimated with each of the dynamic fixed effects, MG and PMG estimators indicates that pooling across countries yields efficient and consistent estimates and hence, that fixed effects estimation is appropriate. Results are available from the author upon request.

13. For example, in calculating the foreign R&D stock of country *i* in time *t*, the R&D stock of trading partner *j* is weighted according to the trade inflows from country *j* to country *i* in period *t*.

14. The dummy variable is constructed from the barriers to entrepreneurship (overall) sub-index of the OECD indicator of economy-wide PMR. This measure reflects obstacles to easy access to information on existing regulation, general or sector-specific administrative burdens for business start-ups and other general or sector-specific regulations that hinder firm entry.

15. This data is only available for Austria from 1995, constraining the number of observations in the MFP regressions to 428 (rather than 437).

16. While the short-run effects can be obtained from the coefficients of the fixed effects estimator, the long-run effects depend on the speed of adjustment term and require additional computation. The method outlined by Bardsen (1989; see Appendix B for details) is used for calculating the long-run coefficients and the associated error variances.

17. Across the specifications presented in Table 1, Wald tests cannot reject that the long-run parameter equals one.

18. The dummy variable is for countries in the top 10% of the distribution of tax policy reversals (Australia and Belgium). Interaction terms between the user cost and a dummy variable for

countries in the top quartile and top half of the distribution were not statistically significant at conventional levels, though the estimated sign of the coefficients were also positive. The measure of changes in the B-index is dropped in the regressions that include the policy reversals interaction variable given the dummy is defined with regard to movements in the B-index.

19. A hypothesis test finds that the impact of a decline in the user cost on R&D expenditure in such countries is not statistically different from zero.

20. David et al. (2000) note that, when input supply is inelastic, studies at the aggregate level may report a positive relationship between publicly financed R&D and private R&D that stems from the positive effect of increased government demand on R&D input prices. To the extent that such input price movements are adequately reflected in the price deflators used to construct the constant price measures, such bias should be avoided. Optimally, a deflator reflecting upstream prices in the R&D sector would be used to completely allay such concerns, though no such reliable measure is currently available for the full range of countries in the sample.

21. Unreported regressions using the outlined base specification do not provide support for the idea that the effectiveness of government funding for stimulating R&D increases up to a particular threshold but decreases thereafter (Guellec and Van Pottelsberghe, 2003).

22. Running the estimation over two unique sample periods, 1981-1994 and 1996-2008, confirms the sensitivity of the direct government support coefficient to the time period, with the variable only statistically significant in the latter.

23. Although the private funds used to match a government grant could come at the expense of another of a firm's R&D projects. Indeed, the empirical evidence of the importance of matching grants for stimulating R&D spending is rather mixed (for a review, see Klette and Moen, 2011).

24. The measure of financial reform is an aggregate that summarises financial sector policy along seven dimensions: credit controls and reserve requirements, interest rate controls, entry barriers, state ownership in the banking sector, financial account restrictions, prudential regulations and supervision of the banking sector and securities market policy (for details, see Abiad et al., 2008).

25. Following Bloom and Griffith (2001), the variable was calculated with weights based on average foreign direct investment flows from the domestic to the foreign country over the sample period.

26. Previous work (Barlevy, 2007) suggests that while counter-cyclicality might make sense given lower opportunity costs for the resources employed in the R&D process, R&D activity is usually found to be pro-cyclical. This may be because firms are more inclined to introduce innovations during boom-times to extract the highest benefit from R&D that is often only partially excludable and will soon be adopted by other firms. In addition, in the presence of credit constraints, firms' R&D spending will fluctuate with company earnings (Andrews and de Serres, 2012b).

27. The estimation results are not materially different depending on whether contemporaneous or lagged inflation is included as a dynamic parameter. Theoretically, observable (lagged) inflation is likely to be more important for firm investment decisions and hence estimates for this variable are presented in Table 1.

28. FDI restrictions are proxied by the OECD FDI Restrictiveness Index that reflects foreign equity restrictions, screening and prior approval requirements, rules for key personnel and other restrictions on the operation of foreign enterprises (for details, see Golub, 2003).

29. Past work finds that the rank of countries by R&D intensity does not change substantially after controlling for differences in industrial structure (OECD, 2006).

30. However, the inclusion of an interaction term between the B-index and a dummy variable for such countries is not found to be statistically significant.

31. The fact that a component of the IPR index is the "patentability of software" and a large proportion of patents in some countries are software-related should be kept in mind when interpreting this result.

32. For further discussion, see OECD (2006).

33. Unreported regressions could not identify the inverted U relationship between product market competition and innovation intensity highlighted by Aghion et al. (2005). The hypothesis was tested in both the R&D and patent stock regressions by including an indicator of product market competition as well as its squared term.

34. Taken together, the estimated coefficients in Table 1 and Table 4 imply that an increase in foreign R&D will have a negative effect on patenting through the business R&D channel that will only be partially offset through the direct positive effect of greater access to the foreign knowledge stock on patenting. This must be interpreted with caution, however, as the coefficient on the foreign

R&D stock term in the R&D regressions does not appear to be very robust (discussed in Section 5.4).

35. Following past work (Griliches and Cockburn, 1988; Bloom and van Reenen, 2002) patent stocks are derived using the perpetual inventory method with an annual depreciation rate of 30%, though the results are robust to assuming other commonly-used depreciation rates for knowledge assets (e.g. 15% used by Hall et al., 2005). The method used to calculate starting values is analogous to that for the R&D stocks detailed in the Data Annex.

36. Including the control variable for business researchers as a share of total industrial employment reduces the chance that the estimated impact of an increase in patents is being overstated by reflecting an increase in the size of the research sector.

37. The independent variable is the stock of patents in units of 10 000. Unreported regressions (available from the author upon request) indicate that the results in columns 1-3 of Table 6 are robust to controlling for population size. The population size variable was dropped from the specifications presented here as it was not statistically significant at conventional levels.

38. Given the empirical link established in Section 5.2.1, R&D and patenting are not included as regressors at the same time in the MFP regression. Setting aside the issues with multicollinearity, if both variables are included in the same regression (along with frontier growth, MFP gap from the frontier and hours worked), the business R&D coefficient remains relatively unchanged from that presented in column 4 of Table 6 while the patents term becomes statistically insignificant at conventional levels.

39. Of course, past studies are mostly based on less recent data than the present analysis. A series of re-estimations of the coefficient on the business R&D term over time periods truncated before 2008 indicates that the size of the estimated positive effect of business R&D on MFP tends to fall as samples include more recent data. Policy conclusions from this exercise are limited, however, as the diminishing size of the R&D coefficient is not robust to dropping one country at a time from the sample.

40. When all of the R&D variables are included in the same specification (column 8), the business R&D term becomes statistically insignificant at conventional levels. This is partly due to an increase in the standard error that the pairwise correlation coefficients indicate may arise from multicollinearity with the other measures of domestic R&D.

41. The frontier country is controlled for, so that the coefficient on the frontier growth variable is the estimated effect for non-frontier countries.

42. This supports the conclusions of Giles and Stroomer (2006). In the presence of concerns about the potential endogeneity of the trade openness variable (Frankel and Romer, 1999), an avenue for further work may be to confirm the result using instrumental variable (IV) estimation (insofar as an appropriate time-varying instrument can be found).

43. This is consistent with the empirical findings of Andrews and Cingano (2012) that low exit costs make it less likely that resources are tied up in inefficient firms.

44. The finding that stricter patent rights are not associated with higher MFP growth contrasts with the results of Bouis et al. (2011).

45. The user cost term becomes insignificant when either Denmark or Portugal are excluded, while the foreign R&D variable is statistically insignificant without either Australia, Denmark, Germany, Ireland, Netherlands, Portugal, Spain, Switzerland or the United States in the sample.

46. The non-business sector R&D variable becomes insignificant when Australia, Ireland or Portugal are excluded. This is the case for the B-index term when Australia or Portugal are excluded and the PMR variable without Switzerland or Finland.

References

Abiad, A., E. Detragiache and T. Tressel (2008), "A New Database of Financial Reforms", *IMF Working Paper* WP/08/266, IMF, Washington, DC.

Acharya, V.V., R.P. Baghai and K.V. Subramanian (2010), "Labour Laws and Innovation", *NBER Working Paper Series,* No. 16484.

Acs, Z.J., D.B. Audretsch, P. Braunerhjelm and B. Carlsson (2009), "The Knowledge Spillover Theory of Entrepreneurship", *Small Business Economics,* January, Volume 32, Issue 1, pp. 15-30.

Akcigit, U. and W.R. Kerr (2010), "Growth Through Heterogeneous Innovations", *Penn Institute for Economic Research Working Paper*, 10-035.

Aghion, P. and P. Howitt (1997), *Endogenous Growth Theory*, The MIT Press.

Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt (2005), "Competition and Innovation: An inverted-U relationship", *Quarterly Journal of Economics 120*, No. 2, pp. 701-728.

Andrews, D. and C. Criscuolo (2013), "Knowledge-Based Capital, Innovation and Resource Allocation", *OECD Economics Department Working Paper*, No. 1046.

Andrews, D., C. Criscuolo and C. Menon (2013), "Do Resources Flow to Innovative Firms? Cross-Country Evidence from Firm-Level Data", *OECD Economics Department Working Paper*, forthcoming.

Andrews, D. and F. Cingano (2012a), "Regulation, Allocative Efficiency and Productivity: Evidence from Firms in OECD Countries", *OECD Economics Department Working Paper*, No. 996.

Andrews, D. and A. de Serres (2012b), "Intangible Assets, Resource Allocation and Growth", *OECD Economics Department Working Paper,* No. 989.

Arellano, M. and S. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies* 58, pp. 277-297.

Arrow, K.J. (1962), "Economic Welfare and the Allocation of Resources to Invention", *The Rate and Direction of Inventive Activity,* Princeton University Press.

Audretsch, D.B. and M.P. Feldman (2003), "Knowledge Spillovers and the Geography of Innovation", *The Handbook of Urban and Regional Economics*, Vol. 4.

Autor, D.H. (2003), "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing", *Journal of Labor Economics*, 21(1), pp. 1-42.

Bardsen, G., (1989), "Estimation of Long Run Coefficients in Error Correction Models", *Oxford Bulletin on Economics and Statistics*, 51, 2, pp. 345-350.

Barlevy, G. (2007), "On the Cyclicality of Research and Development", *American Economic Review,* 97(4), pp. 1131-1164.

Bartelsman, E.J., P.A. Gautier and J. De Wind (2010), "Employment Protection, Technology Choice, and Worker Allocation", *IZA Discussion Paper,* No. 4895.

Bartelsman, E.J., E. Perotti and S. Scarpetta (2008), "Barriers to Exit, Experimentation and Comparative Advantage", *RICAFE2 Working Paper,* No. 056, London School of Economics.

Bassanini, A., L. Nunziata and D. Venn (2009), "Job Protection Legislation and Productivity Growth in OECD Countries", *Economic Policy,* Vol. 24, Issue 58, pp. 349-402.

Bas, M. (2013), "The Effects of Labor Market Regulations on R&D Growth: Panel Data Evidence for OECD Countries", mimeo.

Blanco Armas, E., I. Goldberg, A. Jaffe, T. Muller, J. Sunderland and M. Trajtenberg (2006), "Public Financial Support for Commercial Innovation: Europe and Central Asia Knowledge Economy Study Part 1", *World Bank Chief Economist's Regional Working Paper Series*, Vol. 1, No. 1.

Bloch, C. and E.K. Graversen (2008), "Additionality of Public R&D Funding in Business R&D", *The Danish Centre for Studies in Research and Research Policy Working Papers*, 2008/5.

Bloom, N. and R. Griffith (2001), "The Internationalisation of UK R&D", *Fiscal Studies*, Vol. 33, No. 3, pp. 337-355.

Bloom, N., R. Griffith and J. van Reenen (2002), "Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979-1997", *Journal of Public Economics*, Vol. 85, pp. 1-31.

Bouis, R., R. Duval and F. Murtin (2011), "The Policy and Institutional Drivers of Economic Growth Across OECD and Non-OECD Economies: New Evidence from Growth Regressions", *OECD Economics Department Working Papers,* No. 843.

Bravo Biosca, A., C. Criscuolo and C. Menon (2013), "What Drives the Dynamics of Business Growth?", OECD *STI Policy Papers*, No. 1, OECD Publishing, Paris.

Caselli, F. (2005), "Accounting for Cross-Country Income Differences", *Handbook of Economic Growth,* Edition 1, Vol. 1, Chapter 9, pp. 679-741.

Cockburn, I. and Z. Griliches (1988), "Industry Effects and Appropriability Measures in the Stock Markets Valuation of R&D and Patents", *American Economic Review*, Vol. 78, pp. 419-23.

Coe, D.T. and E. Helpman (1995), "International R&D Spillovers", *European Economic Review*, Vol. 39, pp. 859-887.

Conway, P., D. de Rosa, G. Nicoletti and F. Steiner (2006), "Regulation, Competition and Productivity Convergence", *OECD Economics Department Working Papers*, No. 509.

Czarnitzki, D., K. Kraft and S. Thorwarth (2008), "The Knowledge Production of 'R' and 'D'", *The Centre for European Economic Research Discussion Papers*, No. 08-046.

Danguy, J., G. de Rassenfosse and B. Van Pottelsberghe de la Potterie (2009), "The R&D-Patent Relationship: An Industry Perspective", *European Investment Bank Papers*, Vol. 14, No. 1, pp. 170-195.

David, P.A., B.H. Hall and A.A. Toole (2000), "Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence", *Research Policy*, 29, pp. 497-529.

De Boef, S. and J. Granato (2000), "Testing for Co-integrating Relationships with Near-integrated Data", *Political Analysis,* 8, pp. 99-117.

De Rassenfosse, G. and B. van Pottelsberghe (2009), "A Policy Insight Into the R&D-Patent Relationship", *Research Policy*, Elsevier, Vol. 38(5), pp. 779-792.

De Serres, A., S. Kobayakawa, T. Slok and L. Vartia (2006), "Regulation of Financial Systems and Economic Growth in OECD Countries: An Empirical Analysis", *OECD Economic Studies*, No. 43, pp. 77-113.

Falk, M. (2004a), "What Drives Business R&D Intensity Across OECD Countries?", *WIFO Working Papers*, 236, WIFO.

Falk, M. (2004b), "What Determines Patents Per Capita in OECD Countries", *WIFO Working Papers*, 242, WIFO.

Frankel, J.A. and D. Romer (1999), "Does Trade Cause Growth?", *American Economic Review*, Vol. 89, No. 3, pp. 379-99.

Giles, D.E.A. and C.N. Stroomer (2006), "Does Trade Openness Affect the Speed of Output Convergence? Some Empirical Evidence", *Empirical Economics*, 31, pp. 883-903.

Ginarte, J.C. and W.G. Park (1997), "Determinants of Patent Rights: A Cross-National Study", *Research Policy,* 26, pp. 283-301.

Golub, S (2003), "Measures of Restrictions on Inward Foreign Direct Investment for OECD Countries", *OECD Economic Studies*, No. 36, 2003/1.

Goolsbee, A. (1998), "Does Government R&D Policy Mainly Benefit Scientists and Engineers?", *National Bureau of Economic Research Working Papers,* 6532.

Griffith, R., S. Redding and J. Van Reenan (2004), "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries", *The Review of Economics and Statistics*, 86(4), pp. 883-895.

Griffith, R. and G. MacArtney (2010), "Employment Protection Legislation, Multinational Firms and Innovation", *CEPR Discussion Papers*, No. 7628.

Guellec, D. and B. Van Pottelsberghe de la Potterie (2003), "The Impact of Public R&D Expenditure on Business R&D", *Economics of Innovation and New Technology*, Vol. 12(3), pp. 225-243.

Guellec, D. and B. Van Pottelsberghe de la Potterie (2004), "From R&D to Productivity Growth: Do the Institutional Settings and the Source of Funds of R&D Matter?", *Oxford Bulletin of Economics and Statistics*, Vol. 66, No. 3(07), pp. 353-378.

Hægeland, T. and Jarle Møen (2007), "The Relationship Between the Norwegian R&D Tax Credit Scheme and Other Innovation Policy Instruments", Reports 2007/45, Statistisk sentralbyrå/ Statistics Norway, Oslo-Kongsvinger.

Hall, R. and D. Jorgenson (1967), "Tax Policy and Investment Behaviour," *American Economic Review*, pp. 1034-1076.

Hall, B.H. and J. Mairesse (1995), "Exploring the Relationship Between R&D and Productivity in French Manufacturing Firms", *Journal of Econometrics,* 65, pp. 263-293.

Hall, B.H. and J. Van Reenen (2000), "How Effective Are Fiscal Incentives for R&D? A Review of the Evidence", *Research Policy 29*, pp. 449-469.

Hall, B.H., A.B. Jaffe and M. Trajtenberg (2005), "Market Value and Patent Citations", *The RAND Journal of Economics*, 36, pp. 16-38.

Hall, B.H. and A. Maffioli (2008), "Evaluating the Impact of Technology Development Funds in Emerging Economies: Evidence from Latin America", *NBER Working Paper*, No. 13835.

Hall, B.H., J. Mairesse and P. Mohnen (2009), "Measuring the Returns to R&D", *Handbook of the Economics of Innovation,* pp. 1034-1076.

Hausman, J., B. Hall and Z. Griliches (1984), "Econometric Models for Count Data with an Application to the Patents-R&D Relationship", *Econometrica*, Vol. 52, No. 4, pp. 909-938.

Jaumotte, F. and N. Pain (2005a), "From Ideas to Development: The Determinants of R&D and Patenting", *OECD Economics Department Working Papers*, No. 457, Paris.

Jaumotte, F. and N. Pain (2005b), "An Overview of Public Policies to Support Innovation", *OECD Economics Department Working Papers*, No. 456, Paris.

Jaumotte, F. and N. Pain (2005c), "Innovation in the Business Sector", *OECD Economics Department Working Papers*, No. 459.

Johansson, Å., Y. Guillemette, F. Murtin, D. Turner, G. Nicoletti, C. de la Maisonneuve, P. Bagnoli, G. Bousquet and F. Spinelli (2012), "Long-term Growth Scenarios", *OECD Economics Department Working Papers*, No. 1000.

Kleinknecht, A., K. Van Montfort and E. Brouwer (2002), "The Non-Trivial Choice Between Innovation Indicators", *Economics of Innovation and New Technology*, Vol. 11(2), pp. 109-121.

Klette, T.J. and J. Møen (2011), "R&D Investment Responses to R&D Subsidies: A Theoretical Analysis and Microeconometric Study", *Discussion Paper,* No. 15, Department of Finance and Management Science, Norwegian School of Economics.

Kuemmerle, W. (1999), "The Drivers of Foreign Direct Investment Into Research and Development: an Empirical Investigation", *Journal of International Business Studies*, Vol. 30, pp. 1-24.

Lichtenberg, F.R. and B. van Pottelsberghe de la Potterie (1998), "International R&D Spillovers: A Comment", *European Economic Review*, 42, pp. 1483-1491.

Li, W.C.Y (2012), "Depreciation of Business R&D Capital", paper prepared for the 32nd General Conference of The International Association for Research in Income and Wealth.

Lokshin, B. and P. Mohnen (2009), "What Does it Take for an R&D Tax Incentive Policy to Be Effective", *Updated UNU-MERIT Working Paper,* No. 29, University of Maastricht.

Montmartin, B. (2013), "Business-Funded R&D Intensity: Impact and Complementarity of Public Financial Support", *Universitat de Barcelona Working Paper*, WP4/06.

Nadiri, I. (1993), "Innovations and Technological Spillovers", *NBER Working Paper Series*, No. 4423, Cambridge, Ma.

Nickell, S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica*, Vol. 49, No. 6, pp. 1417-1426.

Nicoletti, G., S. Scarpetta and O. Boylaud (2000), "Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation", *OECD Economics Department Working Papers*, No. 226.

Nicoletti, G. and S. Scarpetta (2003), "Regulation, Productivity, and Growth: OECD Evidence", *Policy Research Working Paper*, No. 2944.

OECD (2006), "Encouraging Innovation: An Overview of Performance and Policies", *Economic Policy Reforms: Going For Growth,* Chapter 3, pp. 55-80, OECD Publishing.

Ortega, F. and G. Peri (2012), "The Effect of Trade and Migration on Income", *NBER Working Papers*, No. 18193.

Park, W.G. (2008), "International Patent Protection: 1960-2005", *Research Policy*, 37, pp. 761-766.

Parente, S.L. and E.C. Prescott (2000), *Barriers to Riches*, MIT Press: Cambridge, Ma.

Parsons, M. and N. Phillips (2007), "An Evaluation of the Federal Tax Credit for Scientific Research and Experimental Development", *Department of Finance Canada Working Papers*, 2007-08.

Pesaran, M.H., Y. Shin and R.P. Smith (1997), "Estimating Long-Run Relationships in Dynamic Heterogeneous Panels", *DAE Working Papers Amalgamated Series*, 9721.

Pesaran, M.H. and R.P. Smith (1995), "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics* 68, pp. 79-113.

Pindyck, R.S (1992), "Investment of Uncertain Cost", *NBER Working Papers*, No. 4175.

Primo, D.M., M.L. Jacobsmeier and J. Milyo (2007), "Estimating the Impact of State Policies and Institutions with Mixed-Level Data", *State Politics and Policy Quarterly*, Vol. 7, No. 4 (Winter), pp. 446-459.

Reinganum, J.F (1983), "Uncertain Innovation and the Persistence of Monopoly", *The American Economic Review*, Vol. 73, No. 4 (September), pp. 741-748.

Sanyal, P. and A. Jaffe (2005), "Peanut Butter Patents Versus the New Economy: Does the Increased Rate of Patenting Signal More Invention or Just Lower Standards?", *Annals of Economics and Statistics*, No. 79/80, pp. 211-240.

Saint-Paul, G. (2002), "Employment Protection, International Specialization, and Innovation", *European Economic Review*, Vol. 46, pp. 375-395.

Saint-Paul, G. (1997), "Is Labour Rigidity Harming Europe's Competitiveness? The Effect of Job Protection on the Pattern of Trade and Welfare", *European Economic Review*, Vol. 41, pp. 499-506.

Seim, L.T. (2009), "FDI and Openness: Differences in Response Across Countries", *Chr. Michelsen Institute.*

Sheehan, J. and A. Wyckoff (2003), "Targeting R&D: Economic and Policy Implications of Increasing R&D Spending", *OECD Science, Technology and Industry Working Papers*, 2003/08.

Solow, R. (1956), "A Contribution to the Theory of Economic Growth", *The Quarterly Journal of Economics*, Vol. 70, No. 1 (February), pp. 65-94.

Sveikauskas, L. (2007), "R&D and Productivity Growth: A Review of the Literature", *Bureau of Labor Statistics Working Papers*, *408*.

Uppenberg, K. (2009), "Innovation and Economic Growth", *European Investment Bank Papers,* Vol. 14, No. 1, pp. 10-35.

Varsakelis, N.C. (2001), "The Impact of Patent Protection, Economy Openness and National Culture on R&D Investment: A Cross-Country Empirical Investigation", *Research Policy*, 30, pp. 1059-1068.

Warda, J. (2006), "Tax Treatment of Business Investment in Intellectual Assets: An International Comparison", *Directorate for Science, Technology and Industry Working Papers*, No. 4.

Warda, J. (2001), "Measuring the Value of R&D Tax Treatment in OECD Countries", OECD *STI Review*, No. 27, pp. 185-211.

Wasmer, E. (2006), "Interpreting Europe-US Labor Market Differences: The Specificity of Human Capital Investments", *American Economic Review*, 96(3), pp. 811-31.

Wölfl, A., I. Wanner, T. Kozluk and G. Nicoletti (2009), "Ten Years of Product Market Reform in OECD Countries 1998-2008: Insights from a Revised PMR Indicator", *OECD Economics Department Working Papers*, No. 695.

Data Annex

R&D stocks

R&D stocks are calculated using the perpetual inventory method followed by Jaumotte and Pain (2005a). Specifically, the starting value is calculated according to the formula $S_{i,\,t=0} = R_{i,\,t+1}/(g_i + \delta)$ where g is the sample mean growth rate of R&D expenditures over the estimation period and δ is the constant depreciation rate of R&D capital. In subsequent time periods, the R&D stock evolves according to the accumulation equation:

$$S_{i,\,t+1} = R_{i,\,t+1} + (1 - \delta)S_t$$

For the R&D regressions, the business sector stock is calculated with a starting value at 1980. This is not the case for Portugal which, due to missing data, has a starting value at 1981. For the patent regressions, the non-business sector R&D stock is taken as the difference between the whole-economy and business sector stocks. As whole-economy R&D expenditure data for Belgium are only available from 1983, the R&D stocks for all countries used in these regressions are calculated with a starting value at 1982. The non-business sector stock will include higher education expenditure on R&D, government intramural expenditure on R&D and direct government support for private R&D.

Multifactor productivity

MFP is calculated in the manner outlined by Johansson et al. (2012). Specifically, the supply side of the economy is modelled according to a Cobb-Douglas production function with constant returns to scale:

$$Y_{it} = K_{it}^{\alpha}(A_{it}hc_{it}L_{it})^{1-\alpha}$$

where Y, K, A, hc and L denote output, physical capital, technical progress (MFP), human capital per worker, total employment and subscript t and i denote year and country. The share of capital is set equal to 1/3. After some manipulations, MFP is derived as:

$$A_{it} = Y_{it}/(\frac{K_{it}}{Y_{it}})^{\alpha/(1-\alpha)}hc_{it}L_{it}$$

GDP is expressed in constant 2005 PPP USD terms, capital stocks are taken from the OECD *Productivity Database* or calculated from investment series (excluding residential housing investment) based on the perpetual inventory method assuming a 4% annual depreciation rate. Human capital stocks are derived from data relating to the average number of years of schooling across the population aged 25-64 and an assumption for the returns to education (see Johansson et al., 2012, for further details).

Geographical exposure to the frontier

The variable in the MFP regressions representing geographical exposure to the MFP leader for a country *i* is calculated by weighting the frontier level of MFP by the distance between the capital city of country *i* and the capital city of the frontier country in kilometres. Distances are obtained from the web page of Jon Haveman at Macalester College.* In the regressions, the variable enters as weighted units of frontier MFP per unit of domestic GDP.

Descriptive statistics

R&D regression (1982-2008)

Variable	Obs.	Mean	Std. dev.	Min.	Max.
R&D stock	512	94 999	199 225	475	1 302 874
R&D flow	512	17 611	36 949	107	234 252
B-index	512	0.94	0.11	0.56	1.05
Inflation	512	0.04	0.03	-0.01	0.27
GDP growth	512	0.03	0.02	-0.06	0.11
User cost	512	0.18	0.03	0.08	0.26
Government financed private R&D	512	2 250	6 644	2	44 375
PMR	512	3.5	1.4	0.9	6.0
Foreign R&D	512	8 840	10 524	225	71 473

Patent regression (1986-2008)

Variable	Obs.	Mean	Std. dev.	Min.	Max.
Patents per capita	418	59	47	0	188
B-index	418	0.94	0.11	0.56	1.05
Inflation	418	0.03	0.02	-0.01	0.130
Business R&D/GDP	418	0.06	0.03	0.00	0.15
(1 + Non-business R&D/Business R&D)	418	1.94	0.59	1.34	4.782
Total R&D	418	176 809	352 537	2 029	2 066 418
IPR	418	4.22	0.55	1.67	4.880
PMR	418	3.23	1.24	0.94	5.83
Foreign R&D stock	418	10 263	11 110	534	71 473

MFP regression (1986-2008)

Variable	Obs.	Mean	Std. dev.	Min.	Max.
MFP gap from frontier	428	-0.31	0.19	-0.78	0.000
MFP growth of frontier country	428	0.007	0.025	-0.081	0.058
Patent stock	428	0.63	1.17	0.000	4.99
PMR	428	3.28	1.27	0.94	5.83
Hours worked	428	1.69	0.15	1.35	2.10
Business researchers/Total emp.	428	4.60	2.84	0.23	14.35
Business R&D stock/GDP	428	0.062	0.034	0.004	0.149
Foreign R&D stock/GDP	428	1.44	1.01	0.12	7.54
Government basic research stock/GDP	428	0.014	0.007	0.002	0.034
Higher education research stock/GDP	428	0.023	0.009	0.004	0.049
Trade openness	428	0.69	0.35	0.16	1.83
Exposure to the frontier/GDP	428	0.047	0.064	0.000	0.371
Bankruptcy cost	428	7.41	4.81	1.00	22.00
B-index	428	0.94	0.11	0.56	1.05
Government R&D subsidies	428	2 213	6 456	3	44 375
IPR	428	4.20	0.578	1.670	4.880

* *www.macalester.edu/research/economics/page/haveman/Trade.Resources/TradeData.html#Gravity*.

APPENDIX A

Calculating the user cost of R&D

The user cost of R&D can be written as:

$$UC_{it} = \left[\frac{1 - A_{it}^{d} - A_{it}^{c}}{1 - v_{it}}\right](r_{it} + \delta_{it})\left(\frac{PRD_{it}}{P_{it}}\right)$$

where A^d and A^c denote the present value of depreciation allowances and tax credits and v the corporate tax rate. The first term in the user cost equation represents the B-index (described in Warda, 2001) and the final term is the price deflator for R&D relative to the price deflator for the goods and services produced. Consistent with the approach taken by Jaumotte and Pain (2005a), the price deflator term is assumed equal to unity in the empirical work.

APPENDIX B

Calculating standard errors for the long-run parameters

The long-run parameters from the ECM model are calculated by employing the transformation detailed by Bardsen (1989). This gives the procedure for a general ECM model such as:

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_1^* \Delta X_t + \beta_2^* X_{t-1} + \varepsilon_t$$

where the long-run multiplier for variable X_t is calculated as $\widehat{\theta_X} = -\widehat{\beta_2^*}/\widehat{\alpha_1^*}$. A disadvantage of following this procedure is that the standard error from the long-run multiplier is not directly estimated. Instead, the variance-covariance matrix for the estimated equation must be obtained and the variance of θ_X calculated as:

$$\widehat{var}(\hat{\theta}_X) = (\hat{\alpha}_1^*)^{-2}[\widehat{var}(\hat{\beta}_2^*) + (\hat{\theta}_X)^2 \widehat{var}(\hat{\alpha}_1^*) + 2\hat{\theta}_X \widehat{cov}(\hat{\beta}_2^*, \hat{\alpha}_1^*)]$$

APPENDIX C

Unreported results from the R&D regression

	Dependent variable = Δln(R&D stock)			
	(1) base	(2)	(3)	(4)
Short-run dynamics				
$\Delta\ln(\text{R\&D stock})_{t-1}$	0.842***	0.843***	0.843***	0.842***
	(28.28)	(28.41)	(28.80)	(27.99)
ECM_{t-1}	-0.013***	-0.012***	-0.011***	-0.013***
	(2.84)	(2.77)	(2.93)	(2.80)
$\Delta\ln(\text{B-index})_t$	-0.035**	-0.036**	-0.035*	-0.035*
	(1.96)	(2.00)	(1.94)	(1.95)
Inflation_{t-1}	-0.133***	-0.143***	-0.139***	-0.129***
	(3.47)	(2.78)	(3.28)	(3.33)
GDP growth_t	0.105**	0.104**	0.103**	0.105**
	(2.36)	(2.34)	(2.33)	(2.38)
Long-run parameters				
$\ln(\text{Government financed private R\&D})_{t-1}$	0.359**	0.412**	0.356**	0.367**
	(2.02)	(2.07)	(2.09)	(2.00)
$\ln(\text{User cost})_{t-1}$	-0.98*	-1.210	-1.007*	-0.969
	(1.65)	(1.63)	(1.65)	(1.56)
$\ln(\text{IPR})_{t-1}$		-0.479		
		(0.42)		
$\ln(\text{Foreign user cost})_{t-1}$			0.83	
			(0.76)	
EPL_{t-1}				-0.094
				(0.65)
Constant	-0.070***	-0.061**	-0.052*	-0.065**
	(2.84)	(2.06)	(1.76)	(2.46)
Observations	512	512	512	512
Adjusted R^2	0.759	0.759	0.759	0.759

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Note: t-statistics in parentheses. The regressions include country and time fixed effects, standard errors are clustered at the country level.

APPENDIX D

The determinants of growth in the patent stock

1986-2008	Dependent variable = Δln(Patent stock per capita)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Short-run dynamics							
$\Delta\ln(\text{Patent stock per capita})_{t-1}$	0.296***	0.285***	0.286***	0.296***	0.294***	0.278***	0.281***
	(5.58)	(6.08)	(6.18)	(5.32)	(5.90)	(6.75)	(5.78)
ECM_{t-1}	-0.16***	-0.16***	-0.16***	-0.14***	-0.15***	-0.17***	-0.14***
	(4.11)	(3.84)	(4.26)	(4.21)	(4.48)	(3.87)	(4.92)
$\Delta\ln(\text{B-index})_t$	-0.20**	-0.19**				-0.18**	-0.16**
	(2.26)	(2.20)				(2.16)	(2.24)
Inflation_{t-1}				-0.46			
				(1.61)			
Long-run parameters							
$\ln(\text{Business R\&D/GDP})_{t-1}$	1.28***	0.97***	0.90***	1.18***	1.41***	1.10***	
	(3.22)	(2.83)	(2.62)	(2.93)	(3.53)	(3.14)	
$\ln(1 + \text{Non-business R\&D/Business R\&D})_{t-1}$	0.90***	0.67*	0.66*	1.03***	1.80***	0.95**	
	(2.68)	(1.76)	(1.74)	(3.00)	(6.88)	(2.43)	
$\ln(\text{Total R\&D})_{t-1}$							0.62**
							(1.98)
$\ln(\text{B-index})_{t-1}$	-0.96***	-0.85***	-0.44**			-0.71***	-0.68***
	(3.24)	(3.38)	(2.25)			(2.96)	(3.41)
$\ln(\text{IPR})_{t-1}$		0.63***	0.79***			0.54	0.90***
		(2.46)	(3.07)			(1.62)	(3.37)
PMR_{t-1}			-0.10**			-0.09**	-0.12**
			(2.33)			(1.99)	(2.59)
$\ln(\text{Foreign R\&D stock})_{t-1}$				0.33*	0.37**		
				(1.94)	(2.43)		
$\ln(\text{Foreign R\&D stock})_{t-1}$ * (High barriers dummy)					-0.28***		
					(2.69)		
$\ln(\text{IPR})_{t-1}$ * (High barriers dummy)						-0.70**	
						(1.97)	
Constant	1.23***	1.01**	0.98**	0.65	0.83	1.19***	-0.45
	(2.79)	(2.41)	(2.56)	(1.47)	(1.63)	(2.44)	(0.82)
Observations	418	418	418	418	418	418	418
Adjusted R^2	0.19	0.20	0.20	0.18	0.19	0.21	0.19

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Note: t-statistics in parentheses. The regressions include country and time fixed effects, standard errors are clustered at the country level and calculated according to the technique outlined by Bardsen (1989) for the long-run parameters. *High barriers dummy* is a dummy for those countries with average barriers to entry over the sample period in the top 50% of the distribution. These countries are Austria, Belgium, Switzerland, Germany, Spain, Finland, France, Italy and Japan.

APPENDIX E

MFP and the flow of new patents

1986-2008	Dependent variable = d.log(MFP)		
	(1)	(2)	(3)
MFP gap from frontier$_{t-1}$	-0.03**	-0.03**	-0.03**
	(2.40)	(2.01)	(2.14)
MFP growth of frontier country$_t$	0.31***	0.30***	0.30***
	(8.22)	(7.76)	(7.68)
Patent flow$_{t-1}$	0.0018*	0.0031***	0.0031**
	(1.86)	(2.72)	(2.34)
PMR$_{t-1}$		-0.003	-0.002
		(1.32)	(1.19)
(PMR)$_{t-1}$ * (Patent flow)$_{t-1}$		-0.0036***	
		(4.17)	
Hours worked$_t$	0.11***	0.10***	0.10***
	(2.93)	(2.93)	(2.95)
Business researchers/Total employment$_{t-1}$	0.002***	0.002***	0.002***
	(2.84)	(2.80)	(2.88)
(Business researchers/Total employment)$_{t-1}$ * (Patent flow)$_{t-1}$			0.002***
			(2.71)
Constant	-0.19***	-0.17***	-0.18***
	(3.11)	(3.05)	(3.10)
Observations	428	428	428
Adjusted R^2	0.27	0.27	0.27

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Note: t-statistics in parentheses. The regressions include country and time fixed effects, standard errors are clustered at the country level.

OECD Journal: Economic Studies
Volume 2013
© OECD 2013

Technological effects of intra-OECD trade in manufacturing: A panel data analysis over the period 1988-2008

by
Javad Abedini*

This article seeks to study how intra-OECD trade in manufacturing goods has affected technological heterogeneity across member states during 1988-2008. To this aim, we derive a panel data version of the Eaton and Kortum (2002) normalised trade model to estimate, annually, the technological heterogeneity of OECD countries. We find a gradual technological convergence across the group as the sensitivity of intra-group trade to price factors increases over time. However, the results diverge when considering European and non-European OECD sub-samples, separately. We find that technological convergence is not an automatic result of intra-group trade but, for that, a more general programme of economic liberalisation, including free movement of capital and labour, is also required.

JEL classification: F13, F15, F33.

Keywords: Regional trade, comparative advantage, technological convergence

* Javad Abedini (abedini@sharif.edu), Graduate School of Management and Economics (GSME), Sharif University of Technology, Teheran, Iran. The views expressed are those of the author and should not be interpreted as representing those of the OECD or its member governments. The author wishes to thank László Halpern and Sébastien Miroudot for helpful comments.

Traditional trade theory treats the state of technology as exogenous. In its most known form, the Ricardian model shows how technological differences across countries give rise to international specialisation and trade. The more countries are (initially) productive in certain industries, the more they should specialise in these. This approach directed many studies which recognise an economic causation from technology to trade (e.g. Markusen and Svenson, 1985; Davis, 1995; Harrigan, 1997).

Since the 1990s, the view has substantially changed. Now theoreticians who endogenise technological heterogeneity across countries are more prominent.

> "Models with exogenous differences in technological capabilities have much to offer trade theory. … Still, they are rather limited in what they can teach us, because they fail to identify the primitive sources of national competitiveness. We turn now to recent developments in the theory that allow us to address issues having to do with the endogenous creation of comparative advantage." (Grossman and Helpman, 1995.)

Trade would affect the state of technology through multiple channels. These channels may, however, cause technological convergence or divergence across countries.

The literature represents two channels for technological convergence through trade. First, a technological upgrading in the production process would be achieved by the import of intermediate and final goods endowed with high technology and/or human capital (Romer, 1990; Coe and Helpman, 1995; Coe et al., 1997; Jones, 2011). Such imports also stimulate technological and organisational learning (Gereffi, 1999; Acharya, and Keller, 2008). Using a rich database, Madsen (2007) shows that 93% of the worldwide increase in TFP over the past century has been solely due to imports of knowledge.

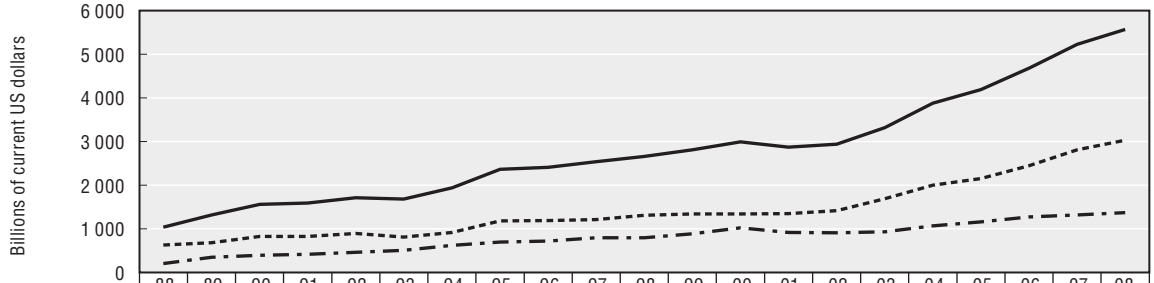
Second, trade in capital goods, even between similarly-endowed countries in technology and capital, motivates international technology diffusion (Xu and Wang, 2000). Examining a sample of 48 countries, Xu and Chiang (2005) indicate that technology spillover, from this last channel, takes place more easily when trade partners belong to high- or middle-income countries.

On the other hand, trade may cause (or reinforce) technological divergence between countries. If countries have, initially, specialised differently in skilled- and unskilled-intensive goods, trade would result in technological divergence by reinforcing the initial relative specialisation in each country. So, trade may strengthen "bad" specialisation in one country and "good" specialisation in another. Baldwin et al. (2001) and O'Rourke et al. (2011) refer the great divergence between the industrialising north and the lagging south during the 18th and 19th centuries to such a phenomenon. Zeira (2010) indicates that the effects may even persist in the long-term if the factors of production (especially, labour force) are immobile across countries.

Including 34 member states, OECD is a forum of countries for international economic co-operation across the globe. It includes many developed countries along with some developing and emerging ones, namely, Chile, Czech Republic, Estonia, Hungary, Mexico,

Poland, Slovakia, Slovenia and Turkey. In addition, a large number of OECD countries are also involved in the European Union or NAFTA which are among the most integrated trade arrangements in the world. The rest of the members are, in part, involved in some multilateral and bilateral trade agreements between themselves such as Japan-Mexico FTA (2005), USA-Chile FTA (2004), PACER[1] (2011), TPP[2] (2006), EURO-MED[3] (1995) and USA-Australia FTA (2005). Indeed, a major part of current OECD countries has been included in free trade zones with other members since the mid-1990s (bilateral.org). Although, these agreements are usually independent of the OECD context of members, they generate a considerable trade dynamic within the sample.

Figure 1. **Total intra-group manufacturing exports, OECD sub-samples**



	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08
OECD	1 040	1 320	1 560	1 590	1 710	1 680	1 940	2 360	2 410	2 540	2 660	2 810	2 990	2 870	2 940	3 320	3 880	4 190	4 680	5 230	5 570
EU OECD	629	679	825	827	892	813	918	1 180	1 190	1 210	1 310	1 340	1 340	1 350	1 420	1 690	2 000	2 150	2 450	2 810	3 030
Non-EU OECD	206	350	395	417	459	504	618	700	721	792	792	884	1 020	913	907	935	1 070	1 160	1 270	1 320	1 370

Note: Author's calculation based on the *OECD-STAN Database*.

Figure 1 shows the huge increase in intra-OECD manufacturing trade during the 1988-2008 period. The same upward trends are also found for European and non-European sub-samples, separately. On average, the intra-group trade, i.e. trade within each OECD sub-sample, has multiplied by more than six compared with that at the end of the 1980s.

Such an expansion in trade exposes OECD countries to certain sources of technological convergence or divergence. While technology and skill exchanges, trade in capital goods, and intra-FDI flows might facilitate technological convergence, a pattern of trade specialisation based on the capital/labour ratio and economies of scale might cause divergence. Taking into account significant heterogeneity between developed and developing OECD member countries, in terms of technology, the scale of changes may be considerable too.

This article seeks to investigate how intra-OECD trade in manufacturing has affected technological heterogeneity of member states during the period 1988-2008.[4] Based on the Eaton and Kortum (2002) theoretical trade model, we develop an augmented panel-data version which allows obtaining annual estimates of technological heterogeneity across the group as well as by the European and non-European sub-samples, separately. This method also requires computing, in a more direct way, bilateral iceberg trade costs in the model.

Depicting the series over time shows that OECD countries achieve a slight technological convergence over 1988-2008. The findings are, however, heterogeneous when considering the European and non-European sub-samples, separately. While a significant technological convergence accompanies European OECD intra-group trade, no sign of technological convergence is found following non-European OECD intra-group trade.

Further analyses provide some explanation for these conflicting results. In particular, we find that technological convergence is not an automatic result of intra-group trade but the latter must also be supported by a general programme of economic liberalisation, including free movement of capital and labour across borders. These policies are those which largely distinguish European members from other OECD countries.

The rest of this article is structured as follows: Section 1 introduces the theoretical model and methodology. We also represent an alternative data generating process which replaces bilateral iceberg trade costs in our Eaton and Kortum-based trade model. Section 2 derives the empirical model with corresponding econometric considerations. The estimates are also provided in the same section. Finally, Section 3 concludes.

1. Theoretical model and methodology

In this section we represent the Eaton and Kortum model which allows us to obtain, in a direct relationship with trade, technological heterogeneity measures across countries over time. To this aim, however, we need to augment the model with panel dimensions and alternatively replace bilateral trade cost data.

Eaton and Kortum (2002) built their model on Dornbusch, Fischer, and Samuelson's (1977) model of Ricardian trade with a continuum of goods. They assume that countries have differential access to technology in the way that efficiency varies across commodities and countries. According to Eaton and Kortum (2002), the state of technology within either country is described by a probability distribution including a parameter which specifically stands for the absolute advantage of each country across the continuum of goods, and another parameter which commonly describes the comparative advantage of countries to trade with each other. They assume that the distributions are independent across countries.

For the sake of brevity, we only introduce their final equation here and ignore the intermediate equations which are already developed by Eaton and Kortum (2002). So, our starting point here is an augmented version of the Eaton and Kortum model which relates trade flows to price factors and technological heterogeneity:

$$\frac{X_{nit} / X_{nt}}{X_{iit} / X_{it}} = \left(\frac{P_{it} d_{nit}}{P_{nt}} \right)^{-\theta_t} . \qquad (1)$$

In equation (1), X_{nt} is country n's total manufacturing spending in year t, of which X_{nit} is spent (c.i.f.) on goods from country i. P_{nt} is the price index for the CES objective function in country n at time t. d_{nit} stands for trade costs between n and i at time t, assumed as having the iceberg form (i.e. $d_{nit} > 1 \quad \forall n \neq i$, $d_{nit} = 1 \quad \forall n = i$) and respecting the triangle inequality rule for trade costs (i.e. for any three countries i, k, and n, $d_{nit} \leq d_{kit} d_{nkt}$). θ_t measures to which degree typical countries i and n differ from each other at year t in terms of technology across manufacturing goods. It measures the sensitivity of normalised bilateral trade (the left-hand side of equation [1]) to price factors (including trade costs) due to technological heterogeneity across countries. Smaller (greater) θ_t implies more (less) technological heterogeneity between countries. The more (less) two countries are different in technology, the more (less) they benefit from the comparative advantage of trade with each other. So, θ_t reflects the extent of (technological) comparative advantage across countries to trade together.

As noted by Eaton and Kortum, this Ricardian-based model bears a resemblance to the standard gravity equation, according to which bilateral trade is related to importers' and

exporters' total expenditures and to their bilateral geographic barriers. In particular, it is compatible with the Anderson and van Wincoop (2004) model in which bilateral trade is also a function of relative trade costs. Similarly, in equation (1), the normalised share of country *i* in country *n*'s national spending is determined by the relative effective country *i*'s costs for supply to country *n*. The triangle inequality implies that the normalised share never exceeds one.

It is important to note that the original model of Eaton and Kortum does not contain the time index *t*. They only rely on cross-sectional data for their empirical investigation. It is, however, possible to add a time index to the model while keeping the same theoretical framework. An additional assumption indicates that the equation is independently extended to either year *t*. Similarly to Eaton and Kortum, we hold the independency of technology distributions across countries but allow that the common θ changes over time. This interpretation of θ is similar to that in Eaton and Kortum (2002), except for the time variation of the parameter. It is used as a constant in the cross sectional study of Eaton and Kortum, but as a time variable in our panel data context.

As a second contribution to the model, we suggest a different data generating process to replace d_{nit}. The fact that markets are competitive and that any pair of countries has had a non-zero trade flow in any product category *j* (*j* = 1, …, 50) leads Eaton and Kortum to assume that $r_{ni}(j) \equiv \ln P_n(j) - \ln P_i(j)$ is bounded above by $\ln d_{ni}$. That is, the difference between the price index in importing and exporting countries should be sufficiently large to cover bilateral trade costs. Based on this assumption, they later calculate the logarithm of P_i / P_n as the mean across *j* of $-r_{ni}(j)$, and measure the entire $\ln(P_i d_{ni} / P_n)$ by the term D_{ni} defined as:

$$D_{ni} = \frac{max2_j\left\{r_{ni}(j)\right\}}{\sum_{j=1}^{50}\left[r_{ni}(j)\right]/50}. \qquad (2)$$

where $max2_j$ means second highest of $r_{ni}(j)$.

There are, however, some shortcomings as regards the Eaton and Kortum method of calculating d_{ni}. First, the single measure $max2_j\left\{r_{ni}(j)\right\}$ does not take into account the distribution of $r_{ni}(j)$ around D_{ni}. Second, data on the consumer price index are not readily available for a large number of countries, especially, at sector level. Finally, this method would not provide any reliable measure for D_{ni} when *n* = *i* resulting in $r_{ni}(j) = 0$. In this last case, domestic trade within each country is treated with the false assumption of zero domestic trade costs.

Instead, we replace d_{nit} using a standard method developed by Anderson and van Wincoop (2004) to directly generate bilateral iceberg trade costs. The method consists of two main steps. First, we specify an empirical gravity model including exporter and importer income variables, as well as a rich trade cost function which represents geographical, official, institutional, and cultural impediments to bilateral trade.[5] The equation also includes a time trend and fixed-country effects standing, respectively, for non-stationarity and any missing effect in the model. This model is then estimated using OECD bilateral trade data over the period 1988-2008.[6] Second, we introduce the fitted values of the same trade cost function in an exponential form to generate the corresponding bilateral iceberg trade costs of countries, as methodologically explained by Anderson and van Wincoop (2004). These bilateral measures are directly used to replace

d_{nit} in our study. This method allows us to estimate bilateral iceberg trade costs (including those belonging to domestic markets) in a continuous way and for any country-pair over time.

It is important to note that our method offers several advantages compared with that used in Eaton and Kortum (2002) to proxy bilateral iceberg trade costs. First, it is based on the gravity approach which is now a standard (and empirically successful) way in the literature to estimate trade costs (see for example, Péridy, 2005; Blum and Goldfarb, 2006; Huang, 2007; Ullah Khan and Kalirajan, 2011; Bergstrand et al., 2013; Abedini, 2013). Second, our trade-cost function controls for various geographical, official, institutional, and cultural trade-cost components, which increases in turn the explanatory power of the model. Third, the method allows us to estimate bilateral trade costs in their iceberg form which would conform more to the theoretical notation of d_{nit} in equation (1). Finally, using this approach, we should be able to estimate bilateral iceberg trade costs in a continuous way and for any country-pair (including those belonging to domestic markets) over time.

The data for the left-hand side of equation (1) has been entirely collected from the STAN database. The missing observations, for less than three consecutive years in some series, have been replaced by their trend estimates. X_{iit}, consumption by i's consumers from domestic production, has been obtained by subtracting manufacturing exports of country i from its manufacturing production at year t. X_{it}, total spending on manufactured goods by country i, is simply the sum of X_{iit} and the import of manufactured goods by country i. The figures are in current US dollars.

Finally, we generate data for the logarithm of the relative price between n and i ($r_{nit} \equiv \ln P_{nt} - \ln P_{it}$) using the EuroStat consumer price index, all items non-food and non-energy 2005 = 100.

2. Empirical model and results

Now, we have all elements in hand to estimate equation (1). Taking a simple logarithm of the latter delivers a regression model with θ_t as a coefficient to be estimated, annually:

$$\ln\left(\frac{X_{ijt}/X_{jt}}{X_{iit}/X_{it}}\right) = -\theta_t \ln\left(\frac{P_{it}d_{ijt}}{P_{jt}}\right). \tag{3}$$

However, the theoretical equation (3) might return inconsistent estimates because of missing variables. A practical solution to this, in the case of panel data, is to control individual (country) effects in the model. These effects stand for any unobserved factor and ensure the good behaviour of the residuals (Baltagi, 2005). As a result, the final empirical model is expressed as:

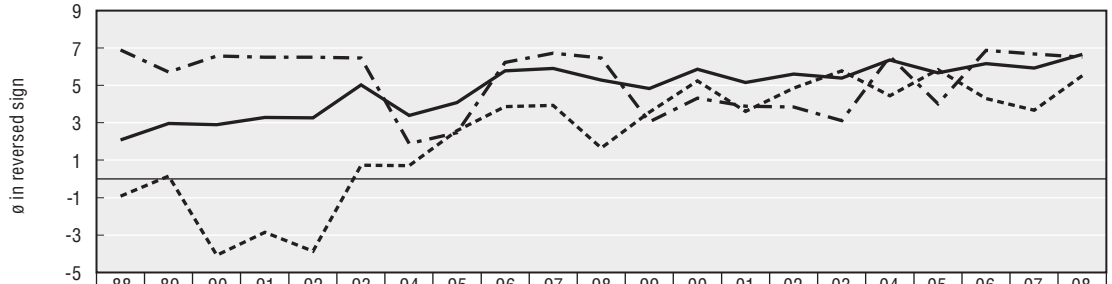
$$\ln\left(\frac{X_{ijt}/X_{jt}}{X_{iit}/X_{it}}\right) = \beta_0 + \theta_t \ln\left(\frac{P_{it}d_{ijt}}{P_{jt}}\right) + \gamma_i + \delta_j + \varepsilon_{ijt}, \tag{4}$$

where β_0 is a constant term ensuring zero-mean for residuals ε_{ijt}, and γ_i and δ_j are, respectively, exporter and importer fixed effects representing unobserved factors at those dimensions.[7] This is indeed an LSDV model which could be straightforwardly estimated using OLS. Due to the presence of exporter and importer fixed effects, the LSDV model (4) generates the same results as a fixed-effect model (FEM) does. In addition, the Hausman-Taylor test indicates the superiority of the FEM approach versus the alternative REM. We also checked for any collinearity between country-fixed effects and the explanatory

variable of the model. The variance inflation factor for each of them (with an average equal to 1.85) remains largely below the threshold of 5 which indicates no severe problem of multicollinearity.

With a negative expected value, θ_t indicates the sensitivity of *i*'s normalised import share in country *j* to the relative price between *i* and *j* at *t*. It corresponds to the relative comparative advantage of the two countries, in terms of technology, to trade with each other. So, any increase (decrease) in θ_t, in absolute value, would reflect a higher (lower) elasticity of trade to price factors due to less (more) technological heterogeneity among countries. In other words, the rise of $|\theta_t|$ implies technological convergence while its fall technological divergence among trading partners over time.

Figure 2. θ **variation over time, OECD sub-samples**



	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08
All OECDs	2.08	2.95	2.89	3.29	3.26	5.02	3.40	4.09	5.78	5.91	5.27	4.82	5.86	5.16	5.60	5.39	6.36	5.67	6.15	5.93	6.66
EU members	-0.9	0.15	-4.0	-2.8	-3.8	0.72	0.71	2.58	3.86	3.92	1.65	3.56	5.24	3.61	4.86	5.77	4.44	5.84	4.30	3.66	5.52
Non-EU members	6.89	5.72	6.57	6.51	6.5	6.46	1.88	2.46	6.22	6.71	6.47	3.06	4.32	3.88	3.85	3.11	6.54	4.01	6.87	6.67	6.51

Note: All estimates are significant at the 1% level (insignificant for EU OECD during 1988-89 and 1993-94).

Figure 2 represents the annual estimates of $-\theta_t$ over the period 1988-2008 using the full OECD dataset and two sub-samples.[8] The solid line shows the time trend for the OECD group, as a whole. The results have also been provided by European (dotted line) and non-European (dashed line) sub-samples to study any possible heterogeneity across them as regards the impact of their own intra-group trade on intra-technological heterogeneity.

The series for the entire OECD group indicates a general upward trend during the period of study. That is, OECD countries have experienced a significant technological convergence throughout 1988-2008, as long as θ_t increased by three, in absolute value, from 2.08 in 1988 to 6.66 in 2008. [9]

In particular, our estimate for θ_{1999} (= 4.82) is well below what Eaton and Kortum (2002) obtained for a sample of 19 OECD countries in 1999 (= 8.28). The gap can be explained by the fact that we have replaced, in a different way, bilateral trade costs in the model. As described in the previous section, our trade cost function includes a larger variety of trade costs and controls for the non-stationarity of variables. This substantially increases our estimates of trade costs and then reduces the amount estimated for θ_t, *ceteris paribus*.

Figure 2 also shows how technological heterogeneity changes for European and non-European members, separately. EU OECD countries differ by the fact that a larger number of them constitute a much deeper economic integration among themselves which liberalises the flows of goods and services but also those of labour and capital within an economic and monetary union.[10] The EU also predicts a supranational institution to harmonise diverse economic, legal, and even political decisions among member states.

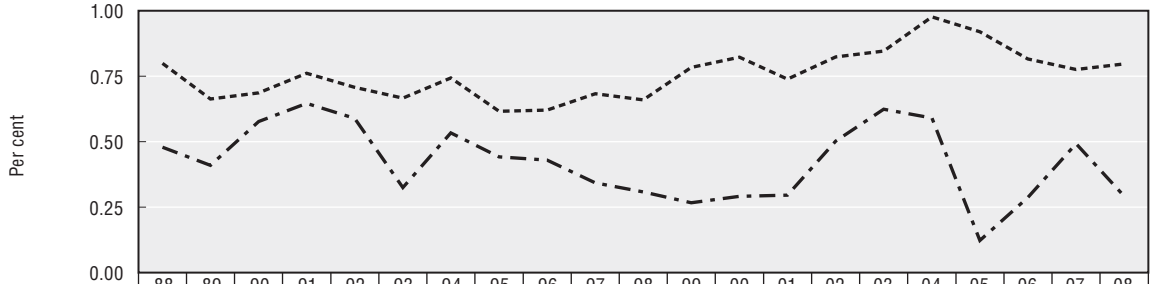
Our findings show that the pattern of technological dynamism for the Europeans substantially differs from that of the non-European members. While EU OECD countries enjoy rapid technological convergence, no evidence of this is found in the case of non-European members. The series of $-\theta_t$ is rapidly increasing for the European set whereas it is rather flat as regards the non-European one. Taking our theoretical background into account, this indicates that intra-group trade has successfully contributed to technological convergence within the European set but not within the rest of the OECD.

This heterogeneity is contrary to the fact that intra-European and intra-non-European OECD trade have increased at nearly the same speed during the period 1988-2008 (Figure 1). As a result, one should look for other sources which might motivate technological convergence within one group while impeding it in another.

According to the literature, FDI and migration (labour) flows are considered as alternative channels for the exchange of technology and expertise across countries (see for example, Borensztein, De Gregorio and Lee, 1998; Sinani and Meyer, 2004; Giroud, 2012; Iranzo and Peri, 2009; Larramona and Sanso, 2006). In addition, the fact that a fraction of trade is carried out outside the group, and is of intra-industry nature, would largely affect the effectiveness of current intra-group trade in causing technological convergence or divergence. As demonstrated below, we find that European and non-European groups of OECD differ significantly from each other as regards these last issues.

Figures 3 and 4 show that EU OECD countries are better positioned in both intra-FDI and intra-migration flows, normalised, respectively, by total FDI and the population of each group. While the trends are not generally increasing in the case of non-EU OECD countries, European members show their growing preference for EU OECD investments and immigrants. These sources contribute to the sharing of technology and expertise within the European set, while the effects are much more limited in the case of non-EU OECD countries.

Figure 3. **Intra-group share in inward FDI, OECD sub-samples**

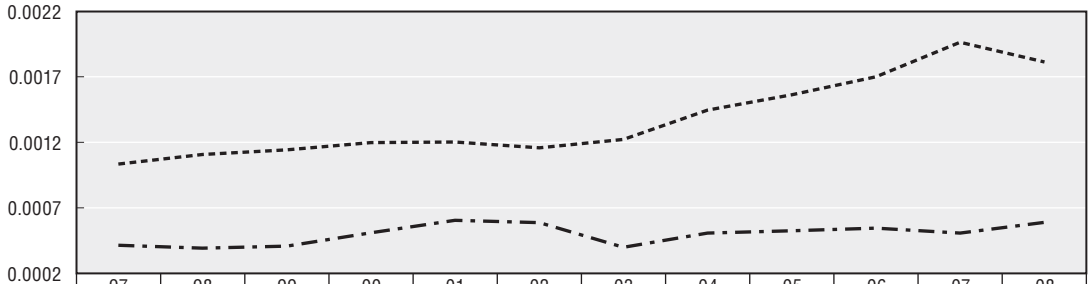


	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08
EU OECD	0.80	0.66	0.69	0.76	0.71	0.67	0.74	0.62	0.62	0.68	0.66	0.78	0.82	0.74	0.82	0.85	0.98	0.92	0.82	0.78	0.80
Non-EU OECD	0.48	0.41	0.58	0.65	0.59	0.32	0.53	0.44	0.43	0.34	0.31	0.27	0.29	0.30	0.50	0.62	0.59	0.12	0.29	0.49	0.29

Note: Author's calculations based on the *OECD-STAN Database*.

Another factor which influences the pattern of technological convergence (or divergence) among a set of countries is the degree to which they trade with countries outside the group. An increasing extra-trade (trade outside the group) exposes the member states to certain sources of technological divergence due to trade with countries differently endowed in technology and expertise.

Figure 4. **Intra-group migration to total population, OECD sub-samples**

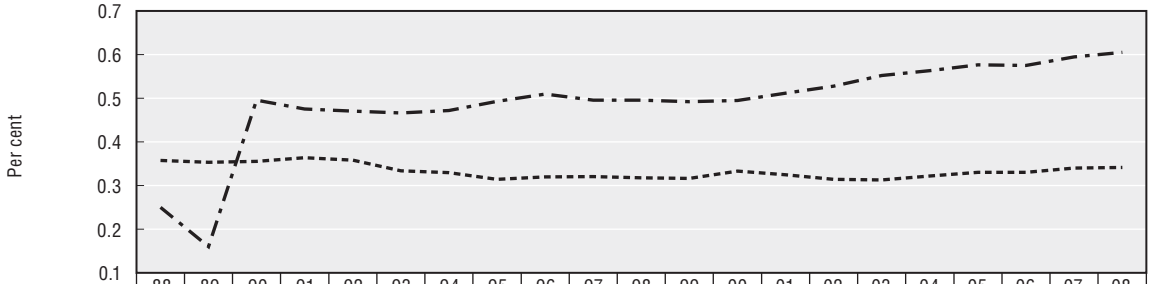


	97	98	99	00	01	02	03	04	05	06	07	08
EU OECD	0.00104	0.00111	0.00114	0.00120	0.00120	0.00116	0.00122	0.00145	0.00156	0.00170	0.00196	0.00181
Non-EU OECD	0.00041	0.00039	0.00041	0.00051	0.00060	0.00059	0.00040	0.00051	0.00052	0.00055	0.00051	0.00059

Note: Author's calculations based on the *OECD-STAN Database*.

Figure 5 shows how EU and non-EU OECD countries differ in this respect. The extra-trade share is lower and slightly decreasing in the case of European members (indicating a larger and increasing intra-group trade share). In contrast, non-EU OECD countries demonstrate a more important and increasing tendency to trade with countries outside their group. This pattern favours technological convergence in the first group while impeding it in the second group.

Figure 5. **Share of extra-group imports in manufacturing, OECD sub-samples**



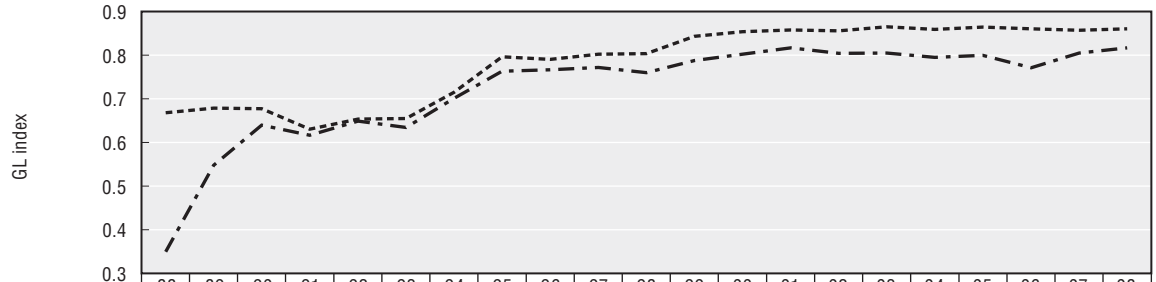
	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08
EU OECD	0.358	0.353	0.355	0.364	0.358	0.334	0.330	0.314	0.320	0.320	0.318	0.316	0.333	0.325	0.314	0.313	0.322	0.330	0.331	0.340	0.341
Non-EU OECD	0.250	0.160	0.496	0.475	0.471	0.466	0.472	0.493	0.510	0.496	0.496	0.492	0.495	0.512	0.528	0.553	0.563	0.577	0.575	0.595	0.605

Note: Author's calculations based on the *OECD-STAN Database*.

Having a large volume of intra-group trade does not suffice, however, to ensure technological convergence. The fact that this trade is of an inter-industry or intra-industry nature also matters. While inter-industry trade reinforces initial specialisation in each country and then results in a sort of technological divergence, intra-industry trade shares knowledge across partners in specific industries and so motivates technological convergence.

As evident from Figure 6, the Grubel-Lloyd index of manufacturing for EU OECD countries is always higher than that for non-European members. That is, intra-industry trade is traditionally more important in the first group than in the second. At other times, this motivates technological convergence within the European sample compared with the non-European one.

Figure 6. **Grubel-Lloyd index of manufacturing, OECD sub-samples**



	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08
EU OECD	0.67	0.68	0.68	0.63	0.65	0.65	0.71	0.80	0.79	0.80	0.80	0.84	0.85	0.86	0.86	0.87	0.86	0.86	0.86	0.86	0.86
Non-EU OECD	0.35	0.55	0.64	0.62	0.65	0.63	0.70	0.76	0.77	0.77	0.76	0.79	0.80	0.82	0.80	0.80	0.79	0.80	0.77	0.80	0.82

Note: Author's calculations based on the *OECD-STAN Database*.

3. Conclusion

Although the Ricardian approach of international trade predicts a causal direction from technology to trade, new approaches treat the state of technology as endogenous. In this article, we seek to study the impact of intra-OECD trade expansion in manufacturing during the period 1988-2008 on technological heterogeneity of member states.

To this aim, we employ the international trade model of Eaton and Kortum which relates bilateral normalised trade to price factors in exporting and importing countries and technological heterogeneity between them. However, we augment the model with panel dimensions in order to obtain annual estimates of technological heterogeneity across countries. We also use an alternative process to generate bilateral iceberg trade cost measures, directly from a trade-cost function.

Our annual estimates of technological heterogeneity indicate that OECD countries have experienced a technological convergence during the period 1988-2008 as the sensitivity of their intra-group trade gradually increases to price factors. In addition, we divide the database into two separate groups of European and non-European members. Our findings show that intra-group trade has affected the distribution of technology differently in each group. While intra-European OECD trade has resulted in a high degree of technological convergence within this group, no sign of convergence is observed following non-European OECD intra-group trade.

European members are distinct in the fact that they share a deeper economic integration which, in parallel, liberalises free movement of capital and labour across borders. Statistical evidence shows that intra-exchanges of goods, capital and labour have been more important and increasing in share for European members while less important and decreasing in share for non-European members. In addition, EU OECD countries benefit from higher intra-industry trade compared with non-European members. These features give EU OECD countries a better possibility of sharing knowledge and technology and so achieving a technological convergence in the sector. In contrast, non-EU OECD countries are far from this pattern as they generally prefer to exchange outside the group. In short, our results indicate that in order for technological convergence to take place, intra-group trade must be supported by a more general programme of economic liberalisation, including free movement of capital and labour. The role of intra-group trade, especially, of an intra-industry nature, is also emphasised.

Notes

1. Pacific Agreement on Closer Economic Relations.
2. Trans-Pacific Strategic Economic Partnership Agreement.
3. European Union-Mediterranean Free Trade Area.
4. Due to the recent world crisis, data for the years after 2008 may not reflect reliable trends. As a result, our analysis in this article is limited to data up to 2008.
5. We employ the same trade cost function used in Abedini (2013) to estimate bilateral iceberg trade costs across the same set of OECD countries. For the sake of brevity, the details are avoided here.
6. The full specification and corresponding estimates of this intermediate model are available upon request.
7. Baldwin and Taglioni (2006) and Magee (2008) suggest using other interaction effects such as country-pair, exporter-year and importer-year effects, instead. However, this method is not practical in our case as it would considerably reduce the degree of freedom in the model due to the increased number of parameters used.
8. The table of estimation with the corresponding standard errors is shown in Appendix A.
9. Please, note that the estimated amounts would vary by using alternative specifications, but the time trend of changes remains stable which indicates the credibility of our results in this sense.
10. Although some European OECD countries do not belong to the EURO zone, they share some of these main features with the EURO members. For example, there is a free circulation of labour and capital between EURO countries and non-EURO EU states such as Poland, United Kingdom, and Switzerland.

References

Abedini, J. (2013), "Estimating Bilateral Iceberg Trade Costs: A Panel Data Study of OECD Countries", forthcoming.

Acharya, R.C. and W. Keller (2008), "Estimating the Productivity Selection and Technology Spillover Effects of Imports", *NBER Working Papers*, 14079.

Anderson, J. and E. van Wincoop (2004), "Trade Costs", *Journal of Economic Literature* 42(3), pp. 691-751.

Baldwin, R. and D. Taglioni (2006), "Gravity for Dummies and Dummies for Gravity Equations", *NBER Working Papers* 12516, National Bureau of Economic Research.

Baldwin, R.E., P. Martin and G.I.P. Ottaviano (2001), "Global Income Divergence, Trade, and Industrialization: The Geography of Growth Take-Offs", *Journal of Economic Growth* 6(1), pp. 5-37.

Baltagi, B.H. (2005), *Panel Data: Theory and Applications*, 1st edition, Physica-Verlag.

Bergstrand, J.H., P. Egger and M. Larch (2013), "Gravity *Redux*: Estimation of Gravity-Equation Coefficients, Elasticities of Substitution, and General Equilibrium Comparative Statistics under Asymmetric Bilateral Trade Costs", *Journal of International Economics* 89(1), pp. 110-121.

Blum, B.S. and A. Goldfarb (2006), "Does the Internet Defy the Law of Gravity?", *Journal of International Economics* 70(2), pp. 384-405.

Coe, D.T. and E. Helpman (1995), "International R&D Spillovers", *European Economic Review* 39(5), pp. 859-887.

Coe, D.T., E. Helpman and A.W. Hoffmaister (1997), "North-South R&D Spillovers", *The Economic Journal* 107(440), pp. 134-149.

Davis, D.R. (1995), "Intra-industry Trade: A Heckscher-Ohlin-Ricardo Approach", *Journal of International Economics* 39(3-4), pp. 201-226.

Dornbusch, R., S. Fischer and P.A. Samuelson (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods", *American Economic Review* 67(5), pp. 823-839.

Eaton, J. and S. Kortum (2002), "Technology, Geography, and Trade", *Econometrica* 70(5), pp. 1741-1779.

Gereffi, G. (1999), "International Trade and Industrial Upgrading in the Apparel Commodity Chain", *Journal of International Economics* 48(2), pp. 37-70.

Giroud, A. (2012), "Heterogeneous FDI in Transition Economies – A Novel Approach to Assess the Developmental Impact of Backward Linkages", *World Development* 40(11), pp. 2206-2220.

Grossman, G.M. and E. Helpman (1995), "Technology and Trade", in *Handbook of International Economics*, Volume 3, pp. 1279-1337.

Harrigan, J. (1997), "Technology, Factor Supplies and International Specialization: Estimating the Neoclassical Model", *American Economic Review* 87(4), pp. 475-494.

Huang, R.R. (2007), "Distance and Trade: Disentangling Unfamiliarity Effects and Transport Cost Effects", *European Economic Review* 51(1), pp. 161-181.

Iranzo, S. and G. Peri (2009), "Migration and Trade: Theory with an Application to the Eastern-Western European Integration", *Journal of International Economics* 79(1), pp. 1-19.

Jones, Ch. (2011), "Intermediate Goods and Weak Links in the Theory of Economic Development", *American Economic Journal: Macroeconomics* 3(2), pp. 1-28.

Larramona, G. and M. Sanso (2006), "Migration Dynamics, Growth and Convergence", *Journal of Economic Dynamics and Control* 30(11), pp. 2261-2279.

Madsen, J.B. (2007), "Technology Spillover through Trade and TFP Convergence: 135 Years of Evidence for the OECD Countries", *Journal of International Economics* 72(2), pp. 464-480.

Magee, C.S.P. (2008), "New Measures of Trade Creation and Trade Diversion", *Journal of International Economics* 75(2), pp. 349-362.

Markusen, J.R. and L.E.O. Svenson (1985), "Trade in Goods and Factors with International Differences in Technology", *International Economic Review* 26(1), pp. 175-192.

O'Rourke, K.H., A.S. Rahman and A.M. Taylor (2011), "Trade, Technology and the Great Divergence", *United States Naval Academy, Working Paper*.

Péridy, N. (2005), "The Trade Effects of The Europe-Mediterranean Partnership: What Are The Lessons for ASEAN Countries?", *Journal of Asian Economics* 16(1), pp. 125-139.

Romer, P. (1990), "Endogenous Technological Change", *Journal of Political Economy* 98(5), pp. 71-102.

Sinani, E. and K.E. Meyer (2004), "Spillovers of Technology Transfer from FDI: the Case of Estonia", *Journal of Comparative Economics* 32(3), pp. 445-466.

Ullah Khan, I. and K. Kalirajan (2011), "The Impact of Trade Costs on Exports and Empirical Modelling", *Economic Modelling* 28(3), pp. 1341-1347.

Xu, B. and E.P. Chiang (2005), "Trade, Patents and International Technology Diffusion", *The Journal of International Trade and Economic Development* 14(1), pp. 115-135.

Xu, B. and J. Wang (2000), "Trade, FDI, and International Technology Diffusion", *Journal of Economic Integration* 15(4), pp. 585-601.

Zeira, J. (2010), "Globalization and Divergence", Centre for Economic Policy Research.

Appendix A

Table of estimation for $-\theta_t$, OECD sub-samples

		88	89	90	91	92	93	94	95	96	97
OECD	$\hat{\theta}$	2.08 (0.70)	2.95 (0.53)	2.89 (0.49)	3.29 (0.45)	3.26 (0.50)	5.02 (0.39)	3.40 (0.38)	4.09 (0.32)	5.78 (0.24)	5.91 (0.23)
	R2	0.69	0.77	0.70	0.75	0.70	0.83	0.76	0.78	0.84	0.83
	F	14.03	18.74	19.87	22.94	19.81	39.42	27.30	36.99	66.60	66.31
EU OECD	$\hat{\theta}$	-0.9 (1.03)	0.15 (0.80)	-4.0 (1.01)	-2.8 (0.92)	-3.8 (1.04)	0.72 (0.78)	0.71 (0.78)	2.58 (0.64)	3.86 (0.44)	3.92 (0.42)
	R2	0.52	0.61	0.62	0.60	0.57	0.75	0.74	0.75	0.80	0.76
	F	3.46	6.03	6.26	6.67	6.41	14.51	13.76	18.02	31.18	25.89
Non-EU OECD	$\hat{\theta}$	6.59 (1.61)	5.72 (1.08)	6.57 (0.76)	6.51 (0.69)	6.5 (0.73)	6.46 (0.75)	1.88 (0.64)	2.46 (0.53)	6.22 (0.46)	6.71 (0.42)
	R2	0.87	0.85	0.87	0.88	0.87	0.88	0.78	0.79	0.85	0.87
	F	12.02	12.88	28.20	31.69	31.10	32.85	18.52	23.14	39.20	46.09

		98	99	00	01	02	03	04	05	06	07	08
OECD	$\hat{\theta}$	5.27 (0.29)	4.82 (0.28)	5.86 (0.23)	5.16 (0.25)	5.60 (0.24)	5.39 (0.25)	6.36 (0.23)	5.67 (0.25)	6.15 (0.23)	5.93 (0.23)	6.66 (0.21)
	R2	0.76	0.77	0.83	0.80	0.80	0.82	0.84	0.81	0.84	0.83	0.87
	F	43.37	46.11	73.39	57.57	59.55	68.87	79.24	62.67	80.48	73.86	97.80
EU OECD	$\hat{\theta}$	1.65 (0.56)	3.56 (0.42)	5.24 (0.36)	3.61 (0.43)	4.86 (0.36)	5.77 (0.33)	4.44 (0.39)	5.84 (0.35)	4.30 (0.39)	3.66 (0.40)	5.52 (0.32)
	R2	0.62	0.77	0.84	0.77	0.83	0.88	0.81	0.86	0.81	0.80	0.88
	F	14.69	31.09	50.63	31.98	45.71	72.96	40.12	60.48	41.68	38.83	68.44
Non-EU OECD	$\hat{\theta}$	6.47 (0.45)	3.06 (0.56)	4.32 (0.44)	3.88 (0.46)	3.85 (0.48)	3.11 (0.52)	6.54 (0.43)	4.01 (0.53)	6.87 (0.43)	6.67 (0.46)	6.51 (0.44)
	R2	0.87	0.77	0.84	0.81	0.80	0.79	0.87	0.77	0.86	0.85	0.86
	F	48.80	24.11	38.92	32.33	29.35	27.28	49.63	24.79	47.12	41.48	44.65

Note: OECD = All OECD countries, EU OECD = European OECD countries, Non-EU OECD = Non-European OECD countries. The numbers in parentheses represent the standard errors of the estimates. All estimates of θ are significant at the 1% level, except in the EU OECD group over 1988-89 and 1993-1994 which are not significant at all. R2 = R-squared. The F-statistics indicate that all LSDV models are significant at the 1% level.

OECD Journal: Economic Studies
Volume 2013
© OECD 2013

Demographic or labour market trends: What determines the distribution of household earnings in OECD countries?

by
Wen-Hao Chen, Michael Förster *and* Ana Llena-Nozal*

This article assesses various underlying driving factors for the evolution of household earnings inequality for 23 OECD countries from the mid-1980s to the mid-2000s. There are a number of factors at play. Some are related to labour market trends – increasing dispersion of individual wages and changes in men's and women's employment rates. Others relate to shifts in household structures and family formation – more single-headed households and increased earnings correlation among partners in couples. The contribution of each of these factors is estimated using a semi parametric decomposition technique. The results reveal that marital sorting and household structure changes contributed, albeit moderately, to increasing household earnings inequality, while rising women's employment exerted a sizable equalising effect. However, changes in labour market factors, in particular increases in men's earnings disparities, were identified as the main driver of household earnings inequality, contributing between one-third and one-half to the overall increase in most countries. Sensitivity analysis applying a reversed-order decomposition suggests that these results are robust.

JEL classification: D31, J12, J22, I30

Keywords: Earnings inequality, assortative mating, female labour supply, decomposition

* Wen-Hao Chen (*wen-hao.chen@oecd.org*), Michael Förster (*michael.forster@oecd.org*) and Ana Llena-Nozal (*ana.llenanozal@oecd.org*) all work in the Social Policy Division of the OECD Directorate for Employment, Labour and Social Affairs. The opinions expressed are those of the authors and do not engage the OECD or its member countries. All errors are the sole responsibility of the authors. We would like to thank Brian Nolan and Mauro Pisu for their most useful comments.
The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Rising wage disparities are often seen as the main culprit behind the growth of household earnings inequality observed in most OECD countries over the past decades. However, the links between individual and household earnings distributions are rather complex (Gottschalk and Danziger, 2005). Individuals usually pool and share their earnings (and other income sources) with other household members, and the distribution of household earnings therefore depends on a number of other factors such as household composition, how earners are clustered within households, and how jobs are distributed among family members. While some of these factors partly offset each other, existing evidence shows that level of trends of household earnings inequality do not necessarily mirror those of individual wage inequality (e.g. Parker, 1995 for the United Kingdom; Saunders, 2005 for Australia; OECD, 2008 for a sample of 19 OECD countries; OECD, 2011 for 23 OECD countries).

There are a number of reasons why the trend in household earnings inequality may differ from that of individual earnings (see McCall and Percheski, 2010 and Burtless, 2011 for a review of the literature). For instance, demographic shifts and societal changes, in particular changes in household structure and family formation, are likely to influence household earnings. The steady increase in the share of single-parent families combined with a tendency for individuals to choose their spouses in groups of similar earnings or educational levels (so-called marital sorting or "assortative mating") may have driven inequality up. Conversely, the substantial increase in women's employment rates may have helped reduce household earnings inequality.

This paper provides a comparative analysis of the importance of both labour market-related and demographic/societal shifts for the evolution of household earnings in 23 OECD countries. Of particular interest are the effects of changing family formation practices. The latter effects have been identified in some case studies as a main factor for increased household earnings and income inequality (e.g. Myles, 2010 for Canada), while other studies either attribute marginal explanatory power to this phenomenon (e.g. Worner, 2006 for Australia) or a sizeable but modest effect (Schwartz, 2010 for the United States).

In the analytical framework used in this paper, inequality of household earnings is determined by two broad sets of factors, referred to as "labour market" factors (earnings dispersion and employment rates) and "family formation" factors (assortative mating and household structure). The aim is to assess their relative influences on changes in household earnings inequality.

Our results below yield the following key findings:

- Between the mid-1980s and mid-2000s, household earnings inequality increased in 21 of the 23 OECD countries studied.
- There was a trend towards more single-headed households, higher female employment, and greater earnings correlation among partners in couples.

- Marital sorting and household structure changes contributed, albeit moderately, to increasing inequality.
- By contrast, rising women's employment exerted a sizable equalising effect.
- Changes in labour market factors, in particular increases in men's earnings disparities, remain the main driver of household earnings inequality, contributing between one-third and one-half to the overall increase in most countries.

The article is organised as follows. Section 1 discusses the trend in the distribution of household earnings and its potential determinants, i.e. the polarisation of men's earnings, changes in the employment rates, and shifts in family formation practices. Section 2 presents an empirical model to quantify the inequality impact of labour market and demographic developments. This applies a decomposition method which relies on the calculation of specific counterfactuals such as "what level of earnings inequality would prevail in the most recent year if all factors but family formation (or other factors) were held constant over time?". The difference between this counterfactual inequality and actual inequality represents the starting point for understanding the role of family formation (and the other factors). Section 3 provides a sensitivity analysis applying a reversed-order decomposition and the final section summarises and concludes.

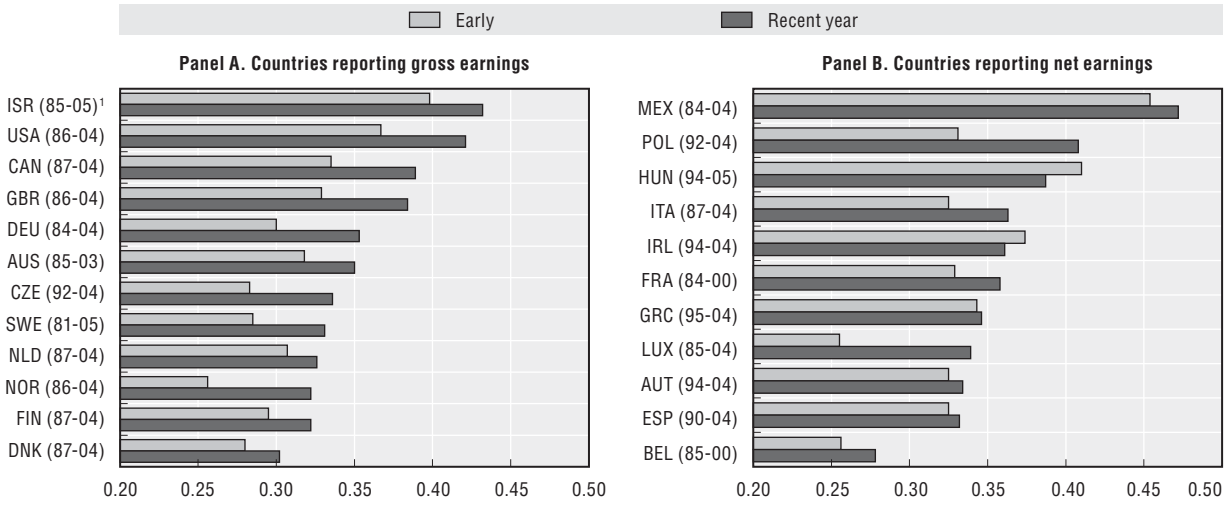
1. Trends in household earnings inequality and their determinants

The analysis in this paper draws on household micro data from the Luxembourg Income Study (LIS) for a period between the mid-1980s and the mid-2000s, covering 23 OECD countries. Samples are restricted to adults aged 25-64 living in a household with a working-age head.[1] Household earnings are calculated as the sum of annual wages and self-employment income from all household members, and are equivalised to account for the economies of scale associated with larger households.[2] For 11 of the 23 countries included in the analysis, only net (of taxes and social contributions) rather than gross earnings were available. It means that in these countries, the changes in any state redistribution mechanisms would also be captured here. The paper does not, however, analyse the role and contribution of tax/transfer policies explicitly. As levels and trends in the distribution of earnings, as well as the contributions of driving factors, will be different for gross than for net earnings, the two groups of countries are discussed separately below.

It is also important to note that this paper focuses on the developments of *annual* earnings rather than *hourly* wages as in most standard analysis of earnings dispersion literature. One implication of the adoption of the annual accounting period is that it captures both dispersion in wage rates (i.e. price effects) and labour force participation (i.e. employment effects). In other words, the observed upward trends in individual earnings inequality can reflect either the widening wage dispersion or fluctuations in the proportion of earners with earnings for part of the year, or both. The extent to which component contributed a larger portion to annual earnings inequality however varies greatly across nations. A simple decomposition analysis in Annex Table A1 reveals that in most OECD countries wage rates account for the largest portion of earnings inequality, explaining 55-63% of earnings variance on average across the countries, while variation in annual working hours contributed to about 28-40% on average. We will return to this issue in Section 2.2 when interpreting the main results of the study.

Figure 1 reveals that household earnings inequality, in terms of the Gini coefficient, has increased noticeably in most OECD countries over time. There is a more consistent

Figure 1. **Evolution of equivalent household earnings inequality (Gini coefficient)**



Note: Samples are restricted to the working-age population (25-64 years) living in a household with a working-age head and positive earnings. Equivalent household earnings are calculated as the sum of earnings from all household members (including elderly and young adults if they lived in a household with a working-age head), adjusted for differences in household size with an equivalence scale (square root of household size).

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

trend among those countries which report gross rather than net earnings (Panel A). Norway and Sweden initially had low inequality levels but experienced a considerable increase over the years, while Canada, the United Kingdom and the United States started with relatively high levels of inequality which further increased by the end of the period. Trends are more diverse among countries which report net earnings only (Panel B). In Luxembourg and Poland, the Gini of household earnings rose more than 7 percentage points over the past decades, while in some countries, such as Greece, Ireland and Hungary, earnings inequality was stable or even fell. For the latter countries, it can however not be disentangled with the data at hand to which extent such a modest change (or decline) in household earnings inequality reflects the impact of labour market and demographic developments or was a combined result of changing market/family trends and tax systems.

1.1. The determinants of rising household earnings inequality

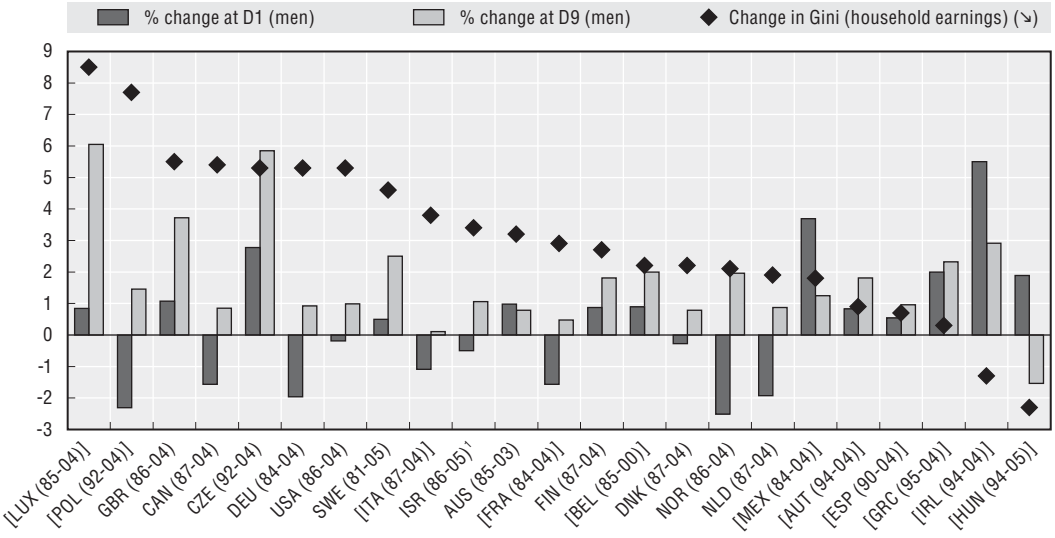
What drives changes in household earnings inequality? Previous research (for a review, see OECD, 2011, p. 199) suggests that inequality of household earnings is affected by two broad types of determinants: labour market factors and household formation factors. The former is often captured by changes in wage dispersion as well as employment rates, while the latter may be modelled by two additional influences: assortative mating, i.e. the degree to which individuals marry within their own income group; and household structure.[3] This sub-section examines changes in both labour market and demographic factors from the mid-1980s to the mid-2000s.

1.1.1. *Trends in men's earnings distribution*

Figure 2 presents the annual percentage change of real earnings[4] among men in the bottom and the top deciles. The distribution of male earnings has become more dispersed in a large majority of the countries studied. In ten countries (such as Poland, Canada and Germany), rising male earnings inequality was a result of growth in real earnings in the top decile combined with a decrease in the bottom decile (see also Table A1 in the Annex).

Figure 2. **Dispersion of men's earnings**

Annual percentage changes in men's real earnings at the bottom and top decile and percentage point changes in Gini coefficients of household earnings



Note: Earnings refer to net earnings for countries in brackets and to gross earnings for other countries. Men's earnings refer to working-age men (25-64) with positive annual earnings. Sample refers to working-age persons in households with positive earnings.

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

Changes in household earnings inequality are smaller in countries where the growth in men's earnings dispersion is less pronounced. The Gini coefficient of household earnings changed very little in Austria, Spain and Greece where the growth of earnings in the top and bottom deciles was either modest or increased at a similar rate. In Ireland, men's earnings increased at both ends of the earnings distribution, but more so in the bottom decile resulting in a drop in household earnings inequality. Such a pattern is also observed in Mexico, though this did not move in hand with decreased overall household earnings inequality, suggesting other important factors at play. In Hungary, which experienced a notable drop in household net earnings inequality, earnings inequality actually decreased among men as real earnings declined in the top decile and rose in the bottom (for interpretation of the results for Hungary, see Box 1 below).

1.1.2. *Trends in employment rates*

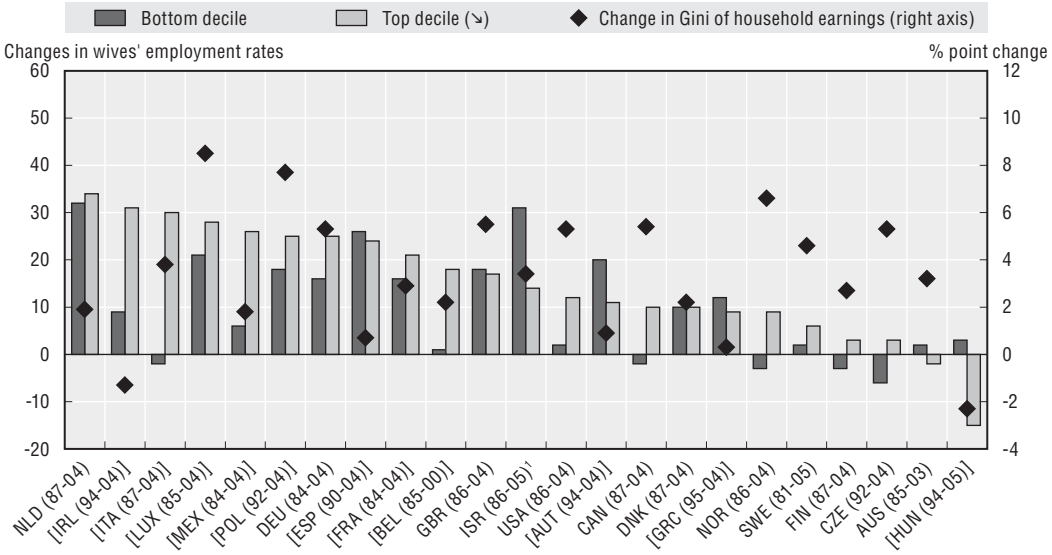
The other important trend affecting household earnings inequality was the substantial increase in female employment rates. Indeed, women's employment rates rose substantially in most OECD countries, exceeding 10 percentage points in 14 of the 23 countries under study, with the largest increases seen in Luxembourg and Spain

(Table A2).[5] Contrary to female employment, male employment rates reveal no obvious trend.

While changes in men's and women's labour market outcomes (i.e. employment rate and wages) undoubtedly reshaped the distribution of household income, their relative role in explaining household earnings inequality also depends on how the family is formed and the extent to which it has changed over time. For instance, there is increasing evidence on the relation between wives' work decisions and husbands' earnings.[6] To investigate this issue, in Figure 3 we look at changes in wives' employment rates by husbands' earnings deciles among couple households with a working husband. In most countries, employment rates increased more among wives of men in the top than in the bottom earnings decile. This was particularly the case in Italy, Mexico, Belgium, Canada and Norway. By contrast, employment rates of wives of low-wage earners increased relatively more in Austria, Hungary and Israel.

Figure 3. **Female employment rates increased the most among wives of top earners**

Wives' employment rates by husbands' earnings (top and bottom decile), couple households



Note: Sample for employment rates restricted to couple households with a working husband. Earnings refer to net earnings for countries in brackets and to gross earnings for other countries.
1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.
Source: Authors' calculations from the Luxembourg Income Study (LIS).

Figure 3, however, also suggests no apparent link between trends in wives' employment rates and husbands' earnings on the one hand, and trends in overall household earnings inequality on the other. For instance, a growing association between wives' employment rates and husband's earnings status is not only observed in countries with a noticeable increase in earnings inequality such as Norway, Canada, Italy and the United States but also in countries with less of an inequality change such as Ireland, Mexico and Belgium. This suggests, at first sight, that the observed higher growth in participation rates of wives of top-earner husbands is not a prime candidate for explaining trends in household earnings inequality.

1.1.3. Trends in assortative mating

There is also a literature that discusses the increasing resemblance of earnings or educational background between husbands and wives, the phenomenon described as "assortative mating". Past research has found that the increased marital sorting contributed a nontrivial portion to widening household income inequality (Cancian et al., 1993; Blackburn and Bloom, 1995; Cancian and Reed, 1999; Hyslop, 2001; Schwartz, 2010).[7]

A straightforward way to measure the extent of assortative mating is to look at simple linear correlation between husbands and wives earnings. This can be captured by Pearson correlation coefficients as presented in Figure A1 in the Annex. The sample covers only couple households with at least one person working. Overall, the correlation coefficients have increased notably over time in 20 out of 23 countries (except Czech Republic, Finland and Hungary), suggesting that there is a general trend toward stronger marital sorting by earnings. The increase in correlation coefficients is most pronounced in Italy, Mexico, Norway and Poland.

To examine the level and development of earnings relationships between spouses more directly, Annex Figure A2 shows working wives' real annual earnings, ranked by husbands' earnings deciles. If there is indeed a growing trend of "assortative mating" (either along educational or occupational characteristics),[8] one would see higher earnings correlations among household members which in turn would accentuate earnings inequality between households. Figure A2 indicates that the level of wives' earnings increases continuously when moving up the ladder of husbands' earnings, especially in the top three deciles. This trend is a departure from the past; in the mid-1980s, wives' annual earnings were still rather equally distributed across the husbands' earnings spectrum in many countries.

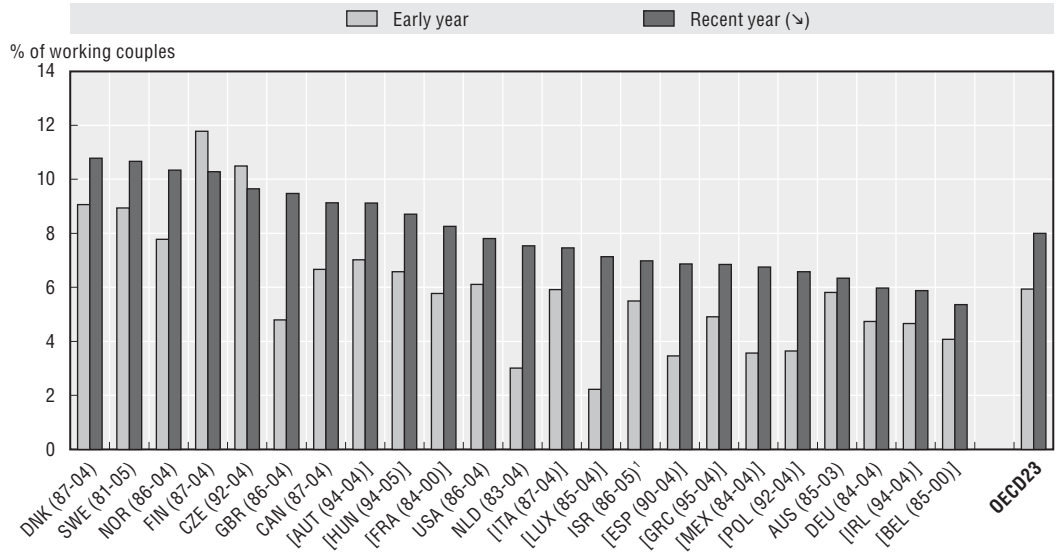
The greatest changes took place in the English-speaking countries, Luxembourg, Norway, Poland and Sweden. In the United Kingdom, for example, the earnings gap between wives of husbands in the top and the bottom decile was about GBP 3 900 in 1987, and this gap increased to GBP 10 200 in 2004 (both figures are expressed in 2005 constant values of national currency). The earnings gap almost tripled in Norway and Poland. In most countries, wives of men in the top deciles benefited most from earnings increases. Poland is a particularly striking example: working wives' earnings rose by almost two-thirds in the top decile, while there was no sizeable increase in the first five deciles.

There is, however, another group of countries which bucked the trend. In Italy and Mexico, the already existing strong correlation between men's and their wives' earnings did not increase further. In Finland, it decreased (when excluding the top decile). And in Austria and Germany, the correlation continues to be weak.

In order to build a summary measure of the degree of marital sorting that can be used for the decomposition analysis described below, we follow Fortin and Schirle (2006) to define assortative mating by the likelihood of a person in earnings decile *i* to be married to a spouse in earnings decile *j*, according to their respective earnings distribution.[9] In general, we find that assortative mating using this measure has increased in nearly all OECD countries (Figure 4). On average for the 23 OECD countries, the share of workers married to a person in the same earnings decile grew from about 6% in the mid-1980s to 8% in the mid-2000s. Luxembourg stands out with the largest increase, from 2.3% in 1985 to 7.4% in 2004. Significant increases were also recorded in the Netherlands and the United Kingdom (more than 4.5 points) and in Mexico, Poland and Spain (between 3.0 and

Figure 4. **A higher degree of assortative mating**

Per cent of workers in earnings decile i with a spouse in the same earnings decile



Note: Refers to couple households with both partners working. Earnings refer to net earnings for countries in brackets and to gross earnings for other countries.

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

3.5 points). The Czech Republic and Finland are the only two countries that experienced a drop in the degree of assortative mating over the past two decades. Levels of assortative mating in terms of earnings have been converging across countries, with the highest levels recorded in the Nordic countries.[10]

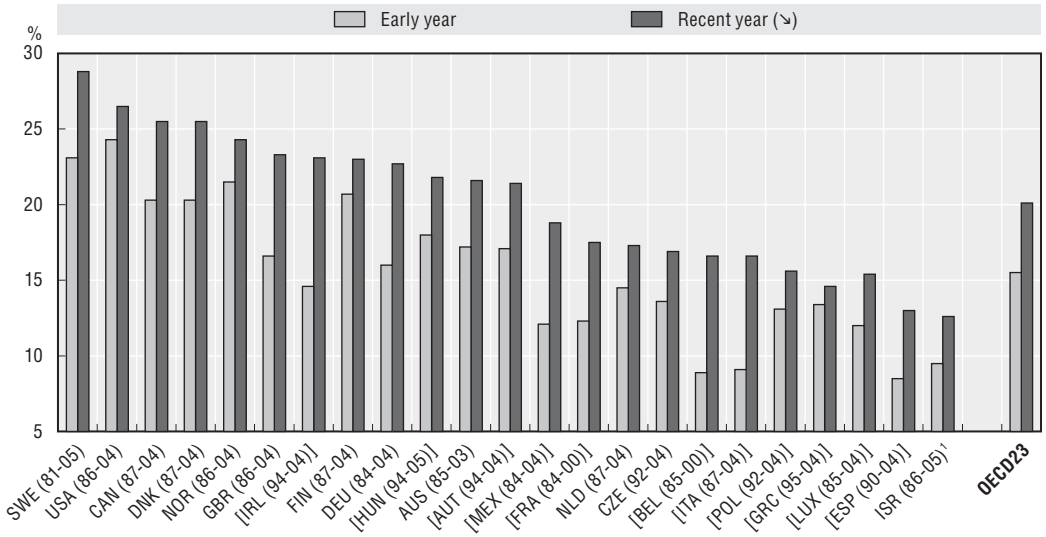
1.1.4. Trends in household composition

Another major change that has been happening at the household level and which may affect inequality is the shift towards more single-headed (i.e. single-parent, single unattached or single with unrelated persons) households.[11] Single-headed households are more common in the Nordic countries and in Canada and the United States where they make up about 25% and more of all working-age households (Figure 5). The share of this household type has increased across the board in all OECD countries under study, on average by almost 5 percentage points. By the mid-2000s, this household type accounted for more than 15% of all households in 20 out of the 23 countries under study. Some single-headed households are more likely to have low earnings (single parents) while others may more often be found among high earners (prime-age singles). An increase in the share of single-headed households therefore could contribute to widening the household earnings dispersion.

2. Explaining changes in household earnings inequality

This section discusses the empirical approach used to quantify the distributional impact of the aforementioned factors. Specifically, the analysis below decomposes the overall change in household earnings inequality among working-age households and assesses the relative impacts of changes in *i*) earnings dispersion among male workers;[12]

Figure 5. **The share of single-headed households has increased in all OECD countries**



Note: Single-headed households refer to single parents with children under 18, singles and singles with unrelated adults. Sample refers to all working-age households (head aged 25-64 years old). Earnings refer to net earnings for countries in brackets and to gross earnings for other countries.
1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.
Source: Authors' calculations from the Luxembourg Income Study (LIS).

ii) male employment share; *iii)* female employment share; *iv)* the degree of assortative mating;[13] and *v)* household structure.[14]

The decomposition method is based on Daly and Valletta (2006) and DiNardo et al. (1996). The point of departure of the method is to develop a counterfactual earnings distribution keeping driving factors, other than family formation, constant (i.e. changes in employment and earnings). This hypothetical earnings distribution allows us to derive the level of inequality that would have prevailed at the end of the period had the general labour market conditions (in terms of men's earnings and labour supply of males and females) remained unchanged. The difference between this counterfactual earnings distribution and actual earnings inequality then represents a starting point for understanding the role of family formation. The impacts of other factors are then obtained based on the "conditional re-weighting procedure". This technique has been used in recent studies (e.g. Chen and Corak, 2008; Daly and Valletta, 2006; Chiquiar and Hanson, 2005) and is similar in spirit to the Oaxaca-Blinder decomposition (Oaxaca, 1973).[15]

2.1. An illustration of results: Canada

We use Canadian samples (1987 and 2004) to illustrate the conditional reweighting and decomposition procedure introduced above and described in Annex Table A3. Panels (A)-(E) in Figure 6 below display the density of equivalent household earnings for these two years in *primary-order decomposition sequences*. Each panel adjusts an additional modelled factor to its 1987 levels, and the impact of a given factor can then be assessed by comparing the differences between the counterfactual distribution with the actual and prior distribution.

The solid line and the dashed line in Panel (A) represent the original density of equivalent earnings for the years 1987 and 2004, respectively. They show that the distribution of household earnings across working-age households became more unequal

in Canada over the years, as density moved from the middle to both tails. The increase in household earnings inequality in Canada is also documented by the summary indicators presented in Tables 1 and A1. The dotted line in Panel (A) delineates the counterfactual density for 2004 with adjustment of men's earnings to 1987 levels. The differences between the dashed and dotted lines therefore reveal the effect of the changing dispersion of men's earnings alone. It shows that the distribution of equivalent household earnings would have been clearly less dispersed if the structure of men's earnings were held constant at its 1987 levels: the adjusted distribution moved density from both tails to the middle.

Panel (B) further adjusts for changes in the employment rate of men. The effect of changing the conditional distribution of men's labour supply appears to have had almost no impact. If any, it reduced the density in the lower half and marginally increased the mass in the upper middle of the distribution, suggesting a very limited contribution to the overall increase in household earnings inequality.

Panel (C) displays the sequential effect of the increase in women's employment rates. The entire adjusted density shifts uniformly toward the left with a relatively greater mass of density in the lower tail. This suggests that the increase in women's employment rates had an appreciable equalising effect as it reduced the density in the left and moved it to the middle and the upper part of the distribution. Panel (C) also suggests that the change in this factor likely contributed a notable gain in median income between these periods.

The effect of the growing tendency to assortative mating is shown in Panel (D). The impact is more visible in the lower tail of the distribution as the adjusted distribution shifted density from the lower tail to the middle. Inequality would have been somewhat lower in the absence of trends to increased assortative mating. This indicates a disequalising effect of this factor though the effect appears lower than for the factor of men's earnings dispersion [in Panel (A) above].

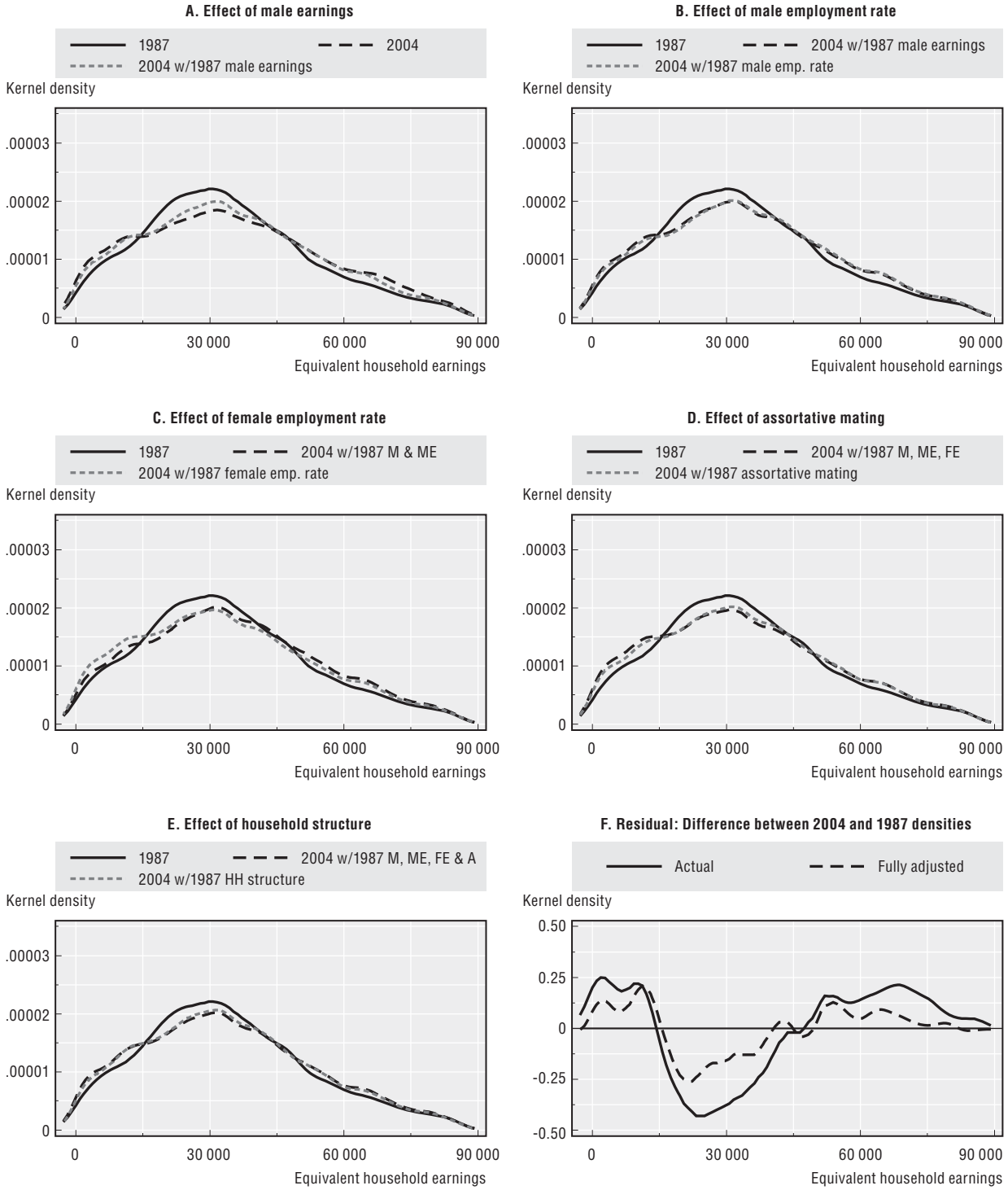
Panel (E) brings in the effect of the changing household structure. This seems to have had a fairly moderate but disequalising impact. The adjusted distribution appears to be less dispersed with a slightly reduced density mass in both tails and a corresponding increase of the mass in the middle of the distribution.

Finally, the residual effect is illustrated in Panel (F), which displays the difference between the adjusted distribution (accounting for the five aforementioned factors) and the original 1987 distribution (i.e. the dashed line). If our controlled factors fully accounted for the observed change in the distribution of equivalent household earnings, we would have obtained a flat line instead. The difference between the dashed line and the flat line therefore represents the residuals. Compared with the difference between the actual distributions for 1987 and 2004 (solid line), Panel (F) shows that accounting for the five factors explains a substantial share (about 50%) in the changing distribution of household earnings between the two periods: the sizable mass that is presented at the bottom, middle and the upper portions of the 2004 distribution has greatly reduced.

2.2. Results of the decomposition analysis

The quantitative assessment of the contribution of each explanatory factor to the changes in the distribution of household earnings is shown in Table 1, in Panel A for countries reporting gross earnings and in Panel B for countries reporting only net earnings. It is important to interpret results for the two samples of countries separately because first-order effects of changes to the tax system impact on changes in the distribution of *net*

Figure 6. **Density of equivalent household earnings, Canada (1987-2004), primary-order decomposition**



Note: Samples are restricted to working-age households with positive household earnings. *M* refers to male earnings, *ML* and *FL* are male and female employment rates, respectively, *A* refers to assortative mating and *S* to household structure. Earnings are expressed in 2005 national currency.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

earnings but have not been modelled in the decomposition – these effects will thus appear in the unobserved residuals, to the contrary of the first country panel.

The first two rows of Table 1 display the actual levels of inequality for the earlier and the most recent year, for two alternative summary inequality indicators, the Gini coefficient and the D9/D1 ratio. The third row shows the changes in the measures between the two years. Decomposition results are presented in the following rows:[16] these numbers

Table 1. **Factors influencing changes in household earnings inequality**

Panel A. Countries reporting gross earnings

	Australia		Canada		Czech Republic		Denmark		Finland		Germany	
	(1985, 2003)		(1987, 2004)		(1992, 2004)		(1987, 2004)		(1987, 2004)		(1984, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.318	4.757	0.335	5.965	0.283	3.669	0.280	4.236	0.295	4.355	0.300	3.725
Recent year	0.350	5.815	0.389	8.222	0.336	4.954	0.302	4.837	0.322	5.380	0.353	6.687
Change	0.032	1.059	0.054	2.257	0.053	1.285	0.022	0.601	0.027	1.025	0.053	2.962
Contribution to change in inequality												
Primary-order decomposition												
1. Men's earnings dispersion	0.017	0.288	0.022	1.188	0.005	0.009	0.009	0.311	0.009	0.221	0.023	1.589
	(.547)	(.272)	(.413)	(.526)	(.092)	(.007)	(.424)	(.517)	(.325)	(.216)	(.429)	(.537)
2. Male employment	0.012	0.544	0.004	0.138	0.002	0.155	0.002	0.080	0.007	0.358	0.017	0.815
	(.367)	(.514)	(.069)	(.061)	(.034)	(.120)	(.101)	(.133)	(.262)	(.350)	(.327)	(.275)
3. Female employment	-0.004	-0.241	-0.013	-0.806	0.007	0.239	-0.008	-0.337	0.002	0.077	-0.017	-0.622
	-(.111)	-(.228)	-(.249)	-(.357)	(.136)	(.186)	-(.378)	-(.561)	(.055)	(.075)	-(.318)	-(.210)
4. Assortative mating	0.006	0.387	0.008	0.595	-0.004	-0.116	0.007	0.372	-0.004	-0.200	0.008	0.352
	(.196)	(.365)	(.152)	(.264)	-(.065)	-(.091)	(.304)	(.619)	-(.155)	-(.195)	(.147)	(.119)
5. Household structure	0.006	0.300	0.004	0.393	0.003	0.139	0.007	0.248	0.004	0.249	0.009	0.398
	(.190)	(.283)	(.080)	(.174)	(.052)	(.108)	(.336)	(.413)	(.151)	(.243)	(.166)	(.134)
6. Residual	-0.006	-0.219	0.029	0.749	0.040	0.861	0.005	-0.073	0.010	0.319	0.013	0.429
	-(.190)	-(.207)	(.535)	(.332)	(.751)	(.670)	(.212)	-(.121)	(.362)	(.312)	(.248)	(.145)
	Israel[1]		Netherlands		Norway		Sweden		United Kingdom		United States	
	(1986, 2005)		(1987, 2004)		(1986, 2004)		(1981, 2005)		(1986, 2004)		(1986, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.398	7.324	0.305	3.668	0.256	3.242	0.285	4.017	0.329	5.259	0.367	7.229
Recent year	0.432	9.130	0.326	4.831	0.322	5.630	0.331	5.609	0.384	7.255	0.420	7.988
Change	0.034	1.806	0.021	1.163	0.065	2.388	0.047	1.593	0.055	1.996	0.054	0.759
Contribution to change in inequality												
Primary-order decomposition												
1. Men's earnings dispersion	0.016	0.732	0.024	0.835	0.033	1.429	0.020	0.690	0.026	1.173	0.025	0.625
	(.456)	(.408)	(1.121)	(.718)	(.509)	(.598)	(.437)	(.433)	(.474)	(.587)	(.459)	(.823)
2. Male employment	0.014	0.475	-0.002	-0.057	0.009	0.348	0.009	0.412	0.002	0.132	-0.005	-0.298
	(.424)	(.264)	-(.107)	-(.049)	(.136)	(.146)	(.188)	(.259)	(.027)	(.066)	-(.097)	-(.393)
3. Female employment	-0.020	-0.178	-0.040	-0.440	-0.001	-0.020	-0.004	-0.140	-0.020	-0.722	-0.004	-0.170
	-(.597)	-(.099)	-(1.847)	-(.378)	-(.020)	-(.008)	-(.081)	-(.088)	-(.365)	-(.362)	-(.069)	-(.224)
4. Assortative mating	0.006	0.139	0.005	0.226	0.006	0.160	0.006	0.256	0.006	0.322	0.005	0.229
	(.165)	(.077)	(.237)	(.195)	(.087)	(.067)	(.135)	(.161)	(.116)	(.161)	(.097)	(.302)
5. Household structure	0.001	0.168	0.002	0.227	0.003	0.111	0.006	0.194	0.001	0.292	0.002	0.188
	(.032)	(.094)	(.102)	(.196)	(.044)	(.046)	(.118)	(.122)	(.018)	(.146)	(.030)	(.248)
6. Residual	0.018	0.460	0.032	0.371	0.016	0.361	0.010	0.181	0.040	0.800	0.031	0.185
	(.521)	(.256)	(1.493)	(.319)	(.243)	(.151)	(.203)	(.114)	(.730)	(.401)	(.580)	(.244)

Note: Numbers in parentheses show the share of the explained change in total change.

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

Table 1. **Factors influencing changes in household earnings inequality** *(cont.)*

Panel B. Countries reporting net earnings

	Austria		Belgium		France		Greece		Hungary		Ireland	
	(1994, 2004)		(1985, 2000)		(1984, 2000)		(1995, 2004)		(1994, 2005)		(1994, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.325	4.522	0.256	3.233	0.329	4.639	0.343	5.039	0.410	7.714	0.374	7.653
Recent year	0.334	4.597	0.278	3.439	0.358	5.878	0.346	5.000	0.387	7.427	0.361	6.632
Change	0.009	0.075	0.022	0.206	0.029	1.239	0.003	-0.039	-0.023	-0.287	-0.013	-1.021
Contribution to change in inequality												
Primary-order decomposition												
1. Men's earnings dispersion	0.012	0.091	0.020	0.042	0.024	0.746	-0.001	-0.357	-0.017	-2.410	-0.005	-0.409
	(1.341)	(1.213)	(1.000)	(.205)	(.819)	(.602)	-(.152)	(8.925)	(.738)	(8.225)	(.370)	(.400)
2. Male employment	-0.007	-0.346	0.010	0.260	0.000	0.050	0.002	0.080	-0.021	-2.210	0.002	0.258
	-(.791)	-(4.613)	(.488)	(1.268)	-(.003)	(.040)	(.636)	-(2.000)	(.933)	(7.543)	-(.148)	-(.253)
3. Female employment	-0.011	-0.227	-0.026	-0.491	-0.023	-0.420	-0.017	-0.578	0.003	0.242	-0.015	-0.620
	-(1.209)	-(3.027)	-(1.271)	-(2.395)	-(.799)	-(.339)	-(5.061)	(14.450)	-(.147)	-(.826)	(1.119)	(.607)
4. Assortative mating	0.002	-0.101	0.008	0.166	0.003	-0.036	0.011	0.249	-0.003	-0.520	0.003	0.029
	(.231)	-(1.347)	(.409)	(.810)	(.101)	-(.029)	(3.364)	-(6.225)	(.129)	(1.775)	-(.193)	-(.028)
5. Household structure	0.000	0.059	-0.005	-0.097	0.002	0.115	0.000	0.093	0.004	1.255	0.001	0.245
	-(.011)	(.787)	-(.246)	-(.473)	(.063)	(.092)	(.121)	-(2.325)	-(.156)	-(4.283)	-(.074)	-(.240)
6. Residual	0.013	0.599	0.013	0.325	0.024	0.786	0.007	0.473	0.011	3.350	0.001	-0.524
	(1.440)	(7.987)	(.621)	(1.585)	(.819)	(.634)	(2.091)	-(11.825)	-(.498)	-(11.433)	-(.074)	(.513)

	Italy		Luxembourg		Mexico		Poland		Spain	
	(1987, 2004)		(1985, 2004)		(1984, 2004)		(1992, 2004)		(1990, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.325	4.155	0.254	3.204	0.454	9.700	0.331	4.962	0.325	4.748
Recent year	0.363	5.190	0.340	5.363	0.472	10.230	0.408	7.497	0.331	5.073
Change	0.038	1.035	0.086	2.159	0.018	0.533	0.076	2.535	0.007	0.325
Contribution to change in inequality										
Primary-order decomposition										
1. Men's earnings dispersion	0.021	0.642	0.022	1.087	0.001	-1.316	0.030	1.231	0.008	0.261
	(.556)	(.620)	(.254)	(.503)	(.080)	-(2.468)	(.396)	(.486)	(1.258)	(.803)
2. Male employment	0.005	0.168	0.003	0.134	0.002	0.419	-0.006	-0.098	-0.002	-0.035
	(.136)	(.162)	(.038)	(.062)	(.109)	(.786)	-(.079)	-(.039)	-(.273)	-(.108)
3. Female employment	-0.008	-0.110	-0.020	-0.235	-0.010	-1.411	-0.003	-0.091	-0.016	-0.233
	-(.214)	-(.106)	-(.228)	-(.109)	-(.549)	-(2.646)	-(.033)	-(.036)	-(2.394)	-(.717)
4. Assortative mating	0.013	0.280	0.022	0.430	0.014	1.695	0.011	0.284	0.017	0.465
	(.350)	(.271)	(.257)	(.199)	(.783)	(3.179)	(.139)	(.112)	(2.500)	(1.431)
5. Household structure	0.000	0.066	0.003	0.086	0.006	0.090	0.006	0.367	0.002	0.090
	(.008)	(.064)	(.033)	(.040)	(.331)	(.169)	(.080)	(.145)	(.258)	(.277)
6. Residual	0.006	-0.011	0.056	0.657	0.004	1.056	0.038	0.841	-0.002	-0.223
	(.164)	-(.011)	(.646)	(.304)	(.246)	(1.980)	(.496)	(.332)	-(.348)	-(.686)

Note: Numbers in parentheses show the share of the explained change in total change.

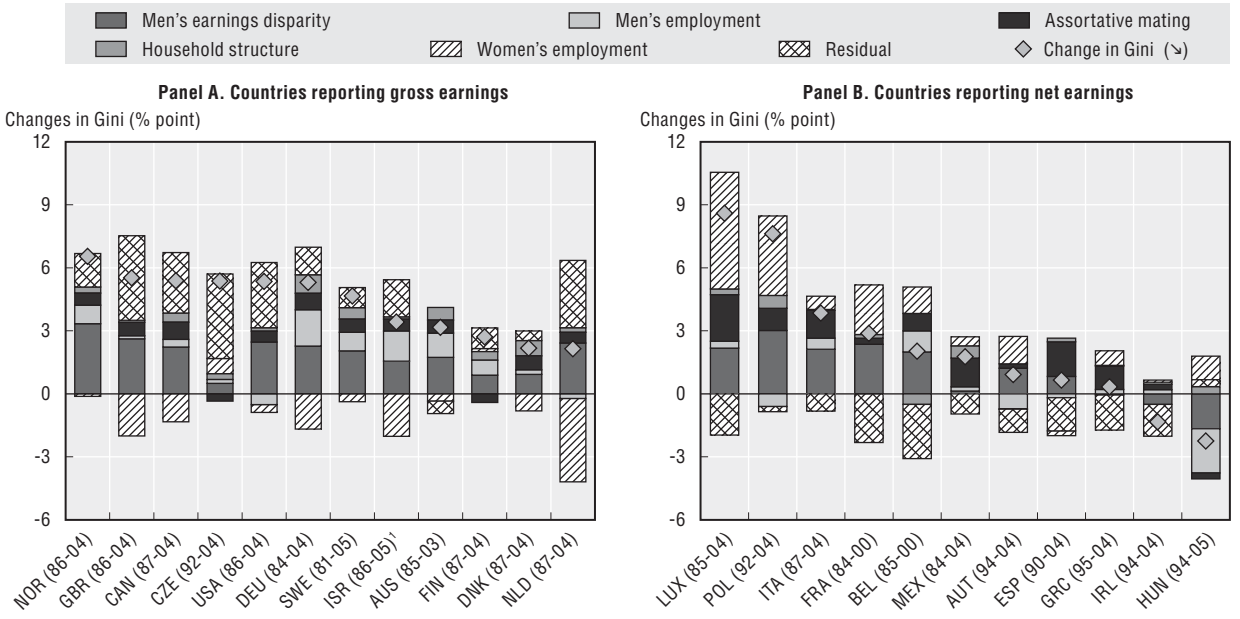
1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

show the amount of change that can be attributed to changes in the explanatory factors, and those in parentheses report each factor's contributory share to the total change in the household inequality measures. Visual presentations of these contributions to changes in the Gini coefficient are presented in Figure 7.

Among countries reporting gross earnings, four main findings emerge from the summary presentation in Panel A, Figure 7. First, the increase in men's earnings disparities

Figure 7. **Explaining changes in household earnings inequality: Contributions of labour market and demographic factors**



Note: Samples are restricted to the working-age population (25-64 years) living in a household with a working-age head. Equivalent household earnings are calculated as the sum of earnings from all household members, adjusted for differences in household size with an equivalence scale (square root of household size).

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

is the main factor driving household earnings inequality, contributing between one-third and half to the overall increase. Second, the increase in women's employment had an equalising effect in nearly all countries. Third, the effect of changing men's employment rates had little impact on the trend in household earnings inequality, with the major exception of Australia, Germany and Israel. Fourth, demographic factors (assortative mating and household structure changes), while contributing positively to increased household earnings inequality, had much more modest effects, contributing less than 20% to the overall increase. These patterns hold for all countries.

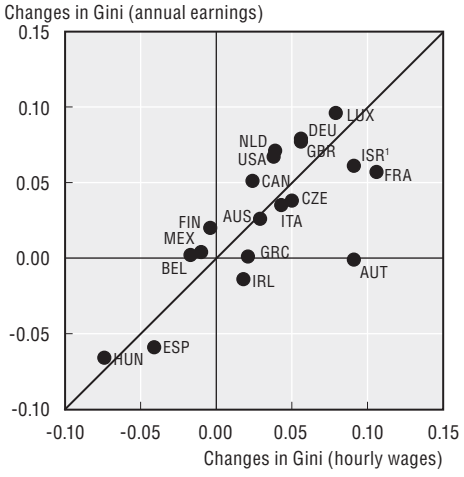
Finally, the contribution of other factors not captured here ("residuals") is higher in the Czech Republic, the Netherlands and the United Kingdom, and lower in the Nordic countries. In the United Kingdom, for instance, more than 70% of the increase in the Gini coefficient of household earnings remains unexplained. On the other hand, the decomposition analysis seems to capture most contributors to the household earnings inequality in Denmark and Norway. Overall, the decomposition results suggest a more modest contribution of demographic relative to labour-market factors and are generally in line with findings from country-specific studies in many respects.[17]

There is more diversity across the sample of countries for which only net earnings estimates are available. Overall changes in household earnings inequality ranged from an increase of over 8 percentage points (Luxembourg) to a 1 to 2 point decrease in Ireland and Hungary (for the specificity of Hungarian results, see Box 1). Demographic factors had more of an impact on trends in household earnings inequality than among the panel of countries reporting gross earnings above, in particular in Luxembourg, Mexico, Italy and Spain. Nonetheless, the increase in men's earnings disparities remains the main

contributor to household earnings inequality in five of the 11 countries. The rise in women's employment rates had a sizeable equalising effect, especially in France, Belgium and Luxembourg. The extent of unobserved factors impacting overall inequality (the residual) is higher among most of the countries in this sample, as the net earnings obviously include the effect of changes to the tax system.

However, the above analysis requires further investigation. First, rising annual earnings dispersion among men is a combined result of both widening wage dispersion and changing work patterns (e.g. more part-time or part-year employment). To gauge what underpins the increase in men's earnings dispersion, Figure 8 plots changes in inequality of hourly wages against changes in inequality of annual earnings for men for selected OECD countries. In an extreme scenario where all employees work the same number of hours per year, the extent of changes in annual earnings inequality would be determined solely by changes in hourly wage distribution, and all countries would lie along the 45° line. This is not the case in reality. In about half of the countries the increase in annual earnings inequality exceeded the increase in inequality measured by hourly wages, suggesting that changing work patterns can play a role in the upward earnings inequality trends, in particular for Canada, the United States, the Netherlands and Finland. For Luxembourg, on the contrary, the increase in annual earnings inequality is more attributable to rising dispersion in hourly wages. Changing work patterns can also lead to more equally distributed annual earnings if low-paid workers have gained more hours worked and/or high-paid workers have reduced hours. Examples of this include Austria and Greece, and to a lesser extent, France and Israel.

Figure 8. **Changes in inequality of hourly wages versus changes in inequality of annual earnings, mid-1980s to mid-2000s, male workers**



Note: Samples are restricted to male workers (25-64 years) with positive hourly wages and annual earnings during the reference year. Changes refer to a period from the mid-1990s to mid-2000s for AUT, CZE, HUN, IRL and GRC; from mid-1980s to 2000 for BEL and FRA.

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

Similarly, for the role of assortative mating the analysis above does not distinguish whether this is occurring mostly because the participation rate of women partnering high-earning men rose particularly rapidly, or because they were already working and their earnings converged to those of their partners. While the exact answer to this question is

beyond the scope of the paper, comparing Table 1 and Figure 3 may provide some hints. For instance, assortative mating contributed about more than 1 percentage point increase to Gini coefficients for both Italy and Greece over time. However, the former experienced a marked increase in the employment rates among wives of high earnings men, while the comparable figure for the latter was rather modest. This may imply that a rising employment rate among women of high-earning partners is a more important driver of assortative mating for Italy, but earnings convergence among couples already working may be a better-candidate scenario for explaining increased marital sorting for Greece. Nevertheless, future analysis is required on this issue.

Finally, the impact of changing household structures (i.e. the trend away from traditional couple households with children towards single-headed households or households with no children) on the distribution of household earnings can be obscured due to the use of equivalence scales.[18] For instance, the relative earnings position of families may improve over time without a rise in *unequivalised* earnings if there is a drop in the average number of dependent children. This would arithmetically reduce the inequality impact of changing household structures due to the rising share of smaller-sized households (e.g. single parents), as the use of equivalisation of total household earnings moved such households further up the earnings distribution.

Box 1. **The specificity of decomposition results for Hungary**

Hungary stands out among all the countries under study as it has registered a moderate decline in household earnings inequality between 1994 and 2005. The moderate fall in earnings inequality is, according to some authors, linked to a series of policy reforms in 2002/03 which raised the wages of all public sector employees (approximately 20% of the Hungarian labour force) by 50%. Telegdy (2006) documented that, prior to the change, the wages of civil servants were lagging behind the salaries earned in the private sector in all occupation groups and at every educational level. The findings above suggest that the changing structure of men's earnings alone has led to a 1.7 percentage-point decline in the Gini coefficient of households net earnings during this period, accounting thus for three-quarters of the decline.

Moreover, given the fact that the public sector often favours employees from more disadvantaged groups (such as new entrants and the elderly), the wage increase may induce higher participation among these groups, and in turn reduce earnings inequality. This is confirmed in the results above in that the increase in men's labour supply further contributed a large part to the decrease in the Gini coefficient.

Finally, despite a tendency toward assortative mating, which matches the OECD average, this factor also contributed to decreasing household net earnings inequality. On the other hand, household structure changes drove earnings inequality up, as did changes in the employment rates of women: Hungary is the only country in the sample in which the employment rate of men grew more than that of women (twice as much).

3. Robustness analysis

One potential problem of the decomposition technique applied above is that the estimated impacts of explanatory factors rely on assumptions about the particular order for the primary decomposition. For instance, the analysis considered household structure last in the decomposition as it assumes that changes in this factor do not affect labour

market choices, but that changes in labour market outcomes (e.g. women's labour force participation) do affect family formation, e.g. by delaying fertility decisions and thus influencing household structures. Similarly, the approach above places women's participation before assortative mating in the decomposition order, assuming that the change in the degree of marital sorting does not have an impact on women's participation decisions. In reality, men's and women's employment rates as well as assortative mating are interdependent.

Although the preceding "primary" order seems a reasonable way to proceed and has been applied in similar types of analyses, it may still over- or underestimate some impacts if there is joint causation in the distribution of factors under examination. For instance, increasing marital sorting might increase (or decrease) the chance of family dissolution and thus have an influence on household structures. On the other hand, it can also be argued that it is the change to the household structure that made marital sorting more feasible. To address such possibilities and the sensitivity of the results, estimates from reverse-order decomposition are presented in Table 2.[19]

The sensitivity analysis shows that results are robust. The increased dispersion of men's earnings remains the most important factor in accounting for household earnings inequality even when it is considered last in the decomposition. Its quantitative impacts are roughly the same as in the primary-order and analysis. Among countries reporting gross earnings, the contribution of men's earnings disparities to household earnings inequality is similarly between one-third and half to the overall increase. The impact of the changing household structure is somewhat larger in magnitude, at the expense of assortative mating and men's and women's employment, suggesting that these three factors are likely to be interdependent. Nevertheless, the inequality-reducing effect of rising female employment remains visible in most countries. The contribution of the residuals is also similar and particularly large in the same countries as when using the "primary" order for the decomposition (the Netherlands, the United Kingdom). In the net income countries for which inequality increased over time, men's earnings were also the major explanatory factor behind the increase, except in Mexico where changes in household structure were more important. Similar to gross income countries, the reversed order of the decomposition leads to a decline in the importance of men's earnings and assortative mating (which contributes negatively to inequality in many cases) while changes in household structure become more prominent.

4. Summary and conclusions

How did the increase in earnings inequality among individuals translate into changes in *household* earnings inequality? The latter takes into account the pooling of earnings of the different household members, changes of labour force participation of men and women, as well as changes in household structures. Overall, the analysis in this article, based on a decomposition technique, finds that labour-market-related trends explain a much larger portion of household earnings inequality development than demographic or "societal" factors.

The dispersions of male earnings have become wider in 20 of the 23 countries under study. In ten of them, this was due to an increase in real earnings in the top decile combined with a decrease in the bottom decile. Female employment rates have substantially increased since the mid-1980s, especially in countries with low starting

Table 2. **Factors influencing on changes in household earning inequality, robustness test**

Panel A. Countries reporting gross earnings

	Australia		Canada		Czech Republic		Denmark		Finland		Germany	
	(1985, 2003)		(1987, 2004)		(1992, 2004)		(1987, 2004)		(1987, 2004)		(1984, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.318	4.757	0.335	5.965	0.283	3.669	0.280	4.236	0.295	4.355	0.300	3.725
Recent year	0.350	5.815	0.389	8.222	0.336	4.954	0.302	4.837	0.322	5.380	0.353	6.687
Change	0.032	1.059	0.054	2.257	0.053	1.285	0.022	0.601	0.027	1.025	0.053	2.962
Contribution to change in inequality												
Reverse-order decomposition												
1. Household structure	0.019	0.855	0.003	0.376	-0.004	0.035	0.010	0.453	0.005	0.294	0.013	0.993
	(.600)	(.808)	(.052)	(.167)	-(.067)	(.027)	(.440)	(.754)	(.192)	(.287)	(.251)	(.335)
2. Assortative mating	0.000	0.088	0.003	0.147	0.000	-0.007	0.001	0.080	-0.001	-0.015	0.002	-0.013
	(.010)	(.083)	(.054)	(.065)	(.007)	-(.005)	(.056)	(.133)	-(.018)	-(.015)	(.028)	-(.004)
3. Female employment	-0.004	-0.085	-0.005	-0.320	0.004	0.184	-0.006	-0.197	0.002	0.089	-0.008	-0.591
	-(.117)	-(.080)	-(.098)	-(.142)	(.067)	(.143)	-(.264)	-(.328)	(.077)	(.087)	-(.157)	-(.200)
4. Male employment	0.003	0.127	0.001	0.168	0.002	0.128	0.003	0.112	0.003	0.089	0.005	0.497
	(.105)	(.120)	(.026)	(.074)	(.037)	(.100)	(.120)	(.186)	(.096)	(.087)	(.096)	(.168)
5. Men's earnings dispersion	0.019	0.292	0.023	1.138	0.011	0.085	0.009	0.226	0.008	0.250	0.028	1.646
	(.594)	(.276)	(.432)	(.504)	(.207)	(.066)	(.435)	(.376)	(.292)	(.244)	(.534)	(.556)
6. Residual	-0.006	-0.219	0.029	0.749	0.040	0.861	0.005	-0.073	0.010	0.319	0.013	0.429
	-(.190)	-(.207)	(.534)	(.332)	(.749)	(.670)	(.213)	-(.121)	(.362)	(.311)	(.247)	(.145)
	Israel[1]		Netherlands		Norway		Sweden		United Kingdom		United States	
	(1986, 2005)		(1987, 2004)		(1986, 2004)		(1981, 2005)		(1986, 2004)		(1986, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.398	7.324	0.305	3.668	0.256	3.242	0.285	4.017	0.329	5.259	0.367	7.229
Recent year	0.432	9.130	0.326	4.831	0.322	5.630	0.331	5.609	0.384	7.255	0.420	7.988
Change	0.034	1.806	0.021	1.163	0.065	2.388	0.047	1.593	0.055	1.996	0.054	0.759
Contribution to change in inequality												
Reverse-order decomposition												
1. Household structure	0.011	0.741	-0.002	0.058	0.018	1.224	0.016	1.034	0.008	0.552	-0.002	0.006
	(.331)	(.412)	-(.085)	(.051)	(.272)	(.513)	(.336)	(.649)	(.143)	(.277)	-(.036)	(.008)
2. Assortative mating	0.001	-0.009	0.001	0.004	0.002	0.196	0.001	-0.010	0.002	0.059	0.007	0.297
	(.041)	-(.005)	(.038)	(.003)	(.031)	(.082)	(.024)	-(.006)	(.031)	(.030)	(.133)	(.392)
3. Female employment	-0.017	-0.256	-0.038	-0.793	-0.005	-0.160	-0.001	-0.056	-0.017	-0.701	-0.006	-0.310
	-(.501)	-(.142)	-(1.793)	-(.688)	-(.073)	-(.067)	-(.015)	-(.035)	-(.312)	-(.351)	-(.112)	-(.409)
4. Male employment	0.005	0.301	-0.002	-0.222	0.002	0.097	0.002	0.094	0.000	-0.009	-0.002	0.009
	(.155)	(.168)	-(.113)	-(.192)	(.031)	(.041)	(.047)	(.059)	-(.002)	-(.005)	-(.041)	(.012)
5. Men's earnings dispersion	0.016	0.560	0.031	1.734	0.033	0.670	0.019	0.350	0.023	1.295	0.025	0.570
	(.455)	(.312)	(1.446)	(1.505)	(.497)	(.281)	(.405)	(.220)	(.411)	(.649)	(.475)	(.752)
6. Residual	0.018	0.460	0.032	0.371	0.016	0.361	0.010	0.181	0.040	0.800	0.031	0.185
	(.519)	(.256)	(1.507)	(.322)	(.243)	(.151)	(.203)	(.114)	(.728)	(.401)	(.581)	(.244)

levels. The increase exceeds 10 percentage points in 14 out of 23 countries, with the strongest increases in the Netherlands, Luxembourg and Spain. In most countries, the rise in employment rates among the wives of men was greater in the top than in the bottom earnings decile.

Increasingly, people are married to spouses with similar earnings levels, known as "assortative mating". This trend was observed in all countries bar the Czech Republic and Finland. On average assortative mating increased by 2 to 6 percentage points, depending on whether a stricter or broader definition is used. Further, the share of single-headed

Table 2. **Factors influencing on changes in household earning inequality, robustness test** *(cont.)*

Panel B. Countries reporting net earnings

	Austria		Belgium		France		Greece		Hungary		Ireland	
	(1994, 2004)		(1985, 2000)		(1984, 2000)		(1995, 2004)		(1994, 2005)		(1994, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.325	4.522	0.256	3.233	0.329	4.639	0.343	5.039	0.410	7.714	0.374	7.653
Recent year	0.334	4.597	0.278	3.439	0.358	5.878	0.346	5.000	0.387	7.427	0.361	6.632
Change	0.009	0.075	0.022	0.206	0.029	1.239	0.003	-0.039	-0.023	-0.287	-0.013	-1.021
Contribution to change in inequality												
Reverse-order decomposition												
1. Household structure	-0.007	-0.365	0.005	-0.020	0.004	0.197	0.010	0.360	-0.008	-2.540	-0.002	0.156
	-(.714)	-(4.867)	(.262)	-(.097)	(.135)	(.159)	(3.188)	-(9.231)	(.367)	(8.789)	(.161)	-(.153)
2. Assortative mating	-0.002	-0.131	-0.002	0.008	-0.002	-0.004	-0.006	-0.295	-0.004	-0.012	0.003	0.081
	-(.198)	-(1.747)	-(.089)	(.039)	-(.069)	-(.003)	-(1.750)	(7.564)	(.164)	(.042)	-(.182)	-(.080)
3. Female employment	-0.005	-0.044	-0.018	-0.042	-0.021	-0.448	-0.005	-0.015	0.002	-0.031	-0.008	-0.205
	-(.571)	-(.587)	-(.881)	-(.205)	-(.719)	-(.362)	-(1.438)	(.385)	-(.097)	(.107)	(.584)	(.201)
4. Male employment	-0.001	-0.043	0.005	0.049	-0.002	-0.103	0.001	-0.010	-0.009	-0.562	0.005	0.218
	-(.077)	-(.573)	(.243)	(.236)	-(.080)	-(.083)	(.219)	(.256)	(.376)	(1.945)	-(.365)	-(.214)
5. Men's earnings dispersion	0.010	0.059	0.017	-0.113	0.026	0.811	-0.004	-0.552	-0.016	-0.494	-0.012	-0.747
	(1.121)	(.787)	(.842)	-(.549)	(.913)	(.655)	-(1.375)	(14.154)	(.686)	(1.709)	(.876)	(.732)
6. Residual	0.013	0.599	0.013	0.325	0.024	0.786	0.007	0.473	0.011	3.350	0.001	-0.524
	(1.440)	(7.987)	(.624)	(1.576)	(.819)	(.634)	(2.156)	-(12.128)	-(.496)	-(11.592)	-(.073)	(.513)

	Italy		Luxembourg		Mexico		Poland		Spain	
	(1987, 2004)		(1985, 2004)		(1984, 2004)		(1992, 2004)		(1990, 2004)	
	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1	Gini	D9/D1
Early year	0.325	4.155	0.254	3.204	0.454	9.700	0.331	4.962	0.325	4.748
Recent year	0.363	5.190	0.340	5.363	0.472	10.230	0.408	7.497	0.331	5.073
Change	0.038	1.035	0.086	2.159	0.018	0.533	0.076	2.535	0.007	0.325
Contribution to change in inequality										
Reverse-order decomposition										
1. Household structure	0.011	0.524	0.014	0.446	0.013	0.607	0.005	0.357	-0.001	0.116
	(.292)	(.506)	(.165)	(.207)	(.746)	(1.142)	(.069)	(.141)	-(.091)	(.357)
2. Assortative mating	0.002	0.038	-0.011	-0.295	-0.003	-0.127	0.010	0.269	0.000	-0.046
	(.060)	(.037)	-(.129)	-(.137)	-(.164)	-(.239)	(.127)	(.106)	-(.030)	-(.142)
3. Female employment	-0.005	-0.172	-0.006	0.192	0.001	0.011	-0.007	-0.163	0.001	0.212
	-(.133)	-(.166)	-(.064)	(.089)	(.056)	(.020)	-(.090)	-(.064)	(.182)	(.652)
4. Male employment	0.003	0.178	0.000	0.000	0.000	0.248	-0.001	-0.059	0.002	0.070
	(.076)	(.172)	(.002)	(.000)	(.011)	(.466)	-(.017)	-(.023)	(.242)	(.215)
5. Men's earnings dispersion	0.021	0.478	0.033	1.159	0.002	-1.263	0.032	1.290	0.007	0.196
	(.542)	(.462)	(.380)	(.537)	(.107)	-(2.375)	(.415)	(.509)	(1.045)	(.603)
6. Residual	0.006	-0.011	0.056	0.657	0.004	1.056	0.038	0.841	-0.002	-0.223
	(.164)	-(.011)	(.646)	(.304)	(.243)	(1.986)	(.495)	(.332)	-(.348)	-(.686)

Note: Numbers in parentheses show the share of the explained change in total change.

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

(i.e. single-parent, single unattached, or single with unrelated) households has grown in all OECD countries under study by an average of 5 percentage points.

Three main findings emerge from the decomposition analysis for the group of countries reporting gross earnings. First, the increase in men's gross earnings disparities is the main factor driving household gross earnings inequality, contributing between one-third and half to the overall increase. Second, the increase in women's employment had an

equalising effect in all countries in that it contributed negatively to overall household gross earnings inequality. Third, demographic factors (i.e. assortative mating and household structure changes), while contributing positively to increasing gross household earnings inequality, had much more modest effects.

There is more diversity across the sample of countries for which only net earnings estimates are available. The demographic factors had somewhat more of an impact on trends in household net earnings inequality. Nonetheless, the increase in men's net earnings disparities remains the main contributor in six of the ten countries. The extent of unobserved factors impacting overall inequality is higher among most of the countries in this group, as it includes the effect of changes to the tax system.

Robustness analyses suggest that the estimated effects of the three labour market factors on changes in household earnings inequality display very similar patterns regardless of which decomposition order is used. The contributions of changing household formation practices, however, are somehow more sensitive to the order of decomposition, with a larger estimated inequality-enhancing impact of changing household structures when it is considered first in the decomposition.

Notes

1. The definition of household refers to all members living in the same dwelling unit regardless of whether or not they are related to each other by blood or marriage. Young adults (16-24) as well as older workers (65+) were excluded in this study to avoid the difficulty of disentangling the effects of labour supply (as thus earnings) due to schooling as well as retirement behaviour.
2. To measure the individual's economic well-being derived from household earnings, the total household earnings are standardised through an equivalence scale in order to adjust for differences in household composition. Following OECD convention, the equivalence scale is defined as the square root of household size (see *www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf*). Total household earnings include earnings from all household members. That means earnings of elderly (65+) and young adults (16-24) who lived in a working-age head household are counted in total household earnings "attributed" to each household member even though the elderly and younger individuals are not included in the sample.
3. There are other factors that are not considered in the analysis below because of lack of data which may affect trends in the distribution of household earned income. One example is changes in the composition of the workforce driven by international migration. Empirical studies on the impact of migration on wage disparities remain largely inconclusive (see, for example, Borjas et al., 1997; Card, 2005).
4. These include full-time and part-time earnings, as well as income from self-employment.
5. Note that the employment rates here refer to the proportion of workers in the working-age population. Workers are defined as persons who receive positive annual earnings regardless of the hours and weeks worked. This is different from the common LFS definition that defines employment as working at least one hour during a brief period (either one week or one day).
6. Juhn and Murphy (1997), for instance, find that the increase in female labour supply over time (either in terms of participation or hours worked) has been strongly non-uniform among all married women in the United States, with wives of high-paid husbands experiencing more pronounced increases in labour market activities than wives of low-paid husbands. Morissette and Hou (2008) also report similar findings for Canada. Esping-Andersen (2009) observes, for five OECD countries, that women's employment participation increased to a much larger extent at the top end of the income distribution, contributing to increased household income inequality.
7. Nevertheless, Cancian et al. (1993) and Cancian and Reed (1998) suggest that wives' earnings equalise the distribution of family income and Harkness (2010) finds an inverse relationship between female employment and income inequality for a sample of 17 OECD countries.

8. The extent of marital sorting may well reflect a more general pattern of educational (or occupational) homogamy. Therefore, another strand of research on assortative mating uses measures of husbands' and wives' education levels (see, for instance, Worner, 2006).

9. That is, we first create decile categories for mens's and women's earnings distributions, separately, for all workers. Then we assign a husband (wife) to earnings decile *i* if his (her) annual earnings falls into decile *i* of men's (women's) earnings distribution. This can be presented by cross-tabulations (10×10) showing husbands' and wives' earnings deciles for each year, respectively. The most rudimentary measure of assortative mating therefore is simply the summation of the diagonal elements.

10. In a separate specification (not shown), we broaden our definition by defining assortative mating as the likelihood of a person in earnings decile *i* to be married to a spouse in the same *or* the adjacent earnings decile *j*, where $|j - i| <= 2$. The overall pattern as well as country rankings remains very similar. With the broader definition, between one-third and half of earners are living with spouses in the same gender-specific earnings quintile. These results are also in line with findings in other empirical literature that used educational homogamy (usually five categories) as a measure for assortative mating. See, for instance, Halpin and Chan (2002) for the United Kingdom, and Worner (2006) for Australia.

11. Karoly and Burtless (1995), Burtless (1999), and Daly and Valletta (2006), for instance, suggest that the increase in single-headed families is responsible for a sizeable proportion of the spread in overall income inequality in the United States. Peichl et al. (2010) find that the changing household composition in Germany between 1991 and 2007 was associated with increasing inequality but the effect was stronger for pre-tax household income inequality than after accounting for taxes. Focusing on family earnings in Canada, Lu et al.. (2011) show that about 20% (30%) of the growth in inequality between 1980 and 1995 (1995 and 2005) can be explained by changing family composition. By contrast, Jäntti (1996) finds that demographic shifts cannot be assigned any major role in the increase in inequality in five OECD countries (including Canada) over the 1980s.

12. Here, only male earnings dispersion is considered. Female wage dispersion is not included in the analytical framework as the evolution of women's wage distribution ties closely to rising women's labour force participation which is one of the labour-market related behavioural changes that is investigated. Because of this correlation, past empirical research similarly did not include women's wage dispersion in such decomposition analyses (e.g. Daly and Valletta, 2006).

13. The degree of assortative mating is described by the likelihood of a husband in earnings decile *i* being married to a wife in earnings decile *j*, according to their respective earnings distribution. This can be presented by a 10×10 cross-tabulation. In the counterfactual exercise, we assigned each dual-earner household into one of the 10 categories, according to their relative degree of marital sorting using information from the 10×10 cross-tabulation. That is, we assign the highest value "10" to households where husbands and wives earnings are in the same decile. Then, the value "9" is given to households with a husband in earnings decile *i* married to a wife in the immediately adjacent earnings decile *j*, where $|j - i| = 1$. Similarly, the value "8" is assigned to households whose earnings as a couple are two deciles apart, $|j - i| = 2$, and the remaining categories are defined accordingly.

14. Household structure is defined according to five mutually exclusive types: *i)* couple households with children; *ii)* couple households without children; *iii)* single-parent households; *iv)* single unattached persons; and *v)* single persons with other adults.

15. Since the decomposition analysis also investigated factors other than earnings, the estimations of the conditional reweighting functions are based on the sample of *all working-age individuals* (including non-workers). Although reweighting factors were estimated for each individual, only workers with positive earnings were used to create the counterfactual distributions of household earnings. See Annex B in Chen et al. (2013) for the detailed decomposition procedure.

16. The decomposition proceeds according to the primary order of effects as described in Table A3 in the Annex.

17. Daly and Valletta (2006), for instance, found that men's earnings contributed the largest share to the change in equivalent family *income* between 1969 and 1989 in the United States (64%), rising female labour supply had a moderate equalising effect and changing family structures had a disequalising effect. Pencavel (2006) also drew similar conclusions from US data for 1968 to 2001, with assortative mating playing a negligible role in accounting for the growth in family earnings inequality over time. For Canada, Lu et al. (2011) showed that 22% of the increase in family earnings inequality between 1980 and 2005 was explained by changing men's wage dispersion, while demographic changes played a rather moderate role. Worner (2006) found that 2-6% of the

increase in inequality of household weekly gross earnings between 1986 and 2003 in Australia can be attributed to assortative mating, a contribution increasing to 4-7% for a broader definition. By contrast, changing patterns in labour force participation explain roughly one-third of the increase in earnings inequality.

18. This article uses the so-called square-root equivalence scale to adjust household earnings and other income for the household size in order to take account of economies of scale (see *www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf*). This scale assumes that the economic needs of a household with four persons is twice as large as that of a one-person household.

19. See Chen et al. (2013, Annex, B.4) for the derivation of re-weighting functions for reverse-order decomposition.

References

Blackburn, M. and D. Bloom (1995), "Changes in the Structure of Family Income Inequality in the United States and Other Industrial Nations during the 1980s", in Polacheck, Solomon W., ed., *Research in Labour Economics,* 14, pp. 141-170.

Borjas, G.R., R. Freeman and L. Katz (1997), "How Much do Immigration and Trade affect Labour Market Outcomes?", *Brookings Paper on Economic Activity*, Vol. 1, pp. 1-67.

Burtless, Gary (1999), "Effects of Growing Wage Disparities and Changing Family Composition on the U.S. Income Distribution", *European Economic Review,* 43, pp. 853-865.

Burtless, Gary (2011), "Demographic Transformation and Economic Inequality" in *The Oxford Handbook of Economic Inequality*, Wiemer Salverda, Brian Nolan and Timothy M. Smeeding, eds., New York, Oxford University Press, pp. 435-454.

Cancian, Maria, Sheldon Danziger and Peter Gottschalk (1993), "Working Wives and the Distribution of Family Income" in *Rising Tides: Rising Inequality in America*. Sheldon Danziger and Peter Gottschalk, eds., New York, Russell Sage Foundation.

Cancian, Maria and Deborah Reed (1998), "Assessing the Effects of Wives' Earnings on Family Income Inequality", *The Review of Economics and Statistics*, 80(1), pp. 73-79.

Cancian, Maria and Deborah Reed (1999), "The Impact of Wives' Earnings on Income Inequality: Issues and Estimates", *Demography*, 36, pp. 173-184.

Card, D. (2005), "Is the New Immigration Really so Bad?", *Economic Journal*, Vol. 115, No. 507, pp. 303-3023.

Chen, Wen-Hao and M. Corak (2008), "Child Poverty and Changes in Child Poverty", *Demography* 45(3), pp. 537-553.

Chen, Wen-Hao, Michael Forster and Ana Llena-Nozal (2013), "Determinants of Households Earnings Inequality: the Role of Labour Market Trends and Changing Household Structure", *LIS Working Papers* Series.

Chiquiar, Daniel and Gordon H. Hanson (2005), "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States", *Journal of Political Economy*, 113(2), pp. 239-281.

Daly, Mary C. and Robert G. Valletta (2006), "Inequality and Poverty in the United States: The Effects of Rising Dispersion of Men's Earnings and Changing Family Behavior", *Economica*, 73(289), pp. 75-98.

DiNardo, John, Nicole M. Fortin and Thomas Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach", *Econometrica,* 64, pp. 1001-1044.

Esping-Anderson, G. (2009), *The Incomplete Revolution – Adapting Welfare States to Women's New Roles*, Polity Press.

Fortin, N. and T. Schirle (2006), "Gender Dimensions of Changes in Earnings Inequality in Canada" in *Dimensions of Inequality in Canada*, David A. Green and Jonathan R. Kesselman, eds.,Vancouver, UBC Press, pp. 307-346.

Gottschalk, P. and S. Danziger (2005), "Inequality of Wage Rates, Earnings and Family Income in the United States, 1975-2002", *The Review of Income and Wealth*, 51(2), pp. 231-254.

Halpin, B. and T. Chan (2002), "Union Dissolution in the United Kingdom", *International Journal of Sociology,* 32, pp. 76-93.

Harkness, S. (2010), "The Contribution of Women's Employment and Earnings to Household Income Inequality: A Cross-Country Analysis", *Luxembourg Income Study Working Paper* 531.

Hyslop, D. (2001), "Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings", *American Economic Review*, 91, pp. 755-777.

Jäntti, M. (1996), "Inequality in Five Countries in the 1980's: The Role of Demographic Shifts, Markets and Government Policies", *Luxembourg Income Study Working Paper* 146.

Juhn, Chinhui, and Kevin M. Murphy (1997), "Wage Inequality and Family Labor Supply", *Journal of Labor Economics,* 15, pp. 72-97.

Karoly, Lynn A. and Gary Burtless (1995), "Demographic Change, Rising Earnings Inequality, and the Distribution of Personal Well-Being, 1959-1989", *Demography,* 32(3), pp. 379-406.

Lu, Yuqian, Rene Morissette and Tammy Schirle (2011), "The Growth of Family Earnings Inequality in Canada, 1980-2005", *The Review of Income and Wealth*, 57(1), pp. 23-39.

Luxembourg Income Study (LIS) Database, *www.lisdatacenter.org* (multiple countries: microdata runs completed between 21 July 2010 and 3 August 2011), LIS, Luxembourg.

McCall, Leslie and Christine Percheski (2010), "Income Inequality: New Trends and Research Directions", in *Annual Review of Sociology*, Vol. 36, pp. 329-347.

Morissette, R. and F. Hou (2008), "Does the Labour Supply of Wives Respond to Husbands' Wages? Canadian Evidence from Micro Data and Grouped Data", *Canadian Journal of Economics*, 41(4), pp. 1185-1200.

Myles, J. (2010), "The Inequality Surge: Changes in the Family Life Course are the Main Cause" in *Inroads: A Journal of Opinion*, Quebec, January.

Oaxaca, R (1973), "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review,* 14, pp. 693-709.

OECD (2008), *Growing Unequal? Income Distribution in OECD Countries*, OECD Publishing, Paris.

OECD (2011), *Divided we Stand – Why Inequality Keeps Rising*, OECD Publishing, Paris.

Parker, S. (1995), "The Structure of the Earnings Distribution: How are Households and Individuals Related?", *Bulletin of Economic Research*, Vol. 47, Issue 2, pp. 127-142.

Peichl, A., N. Pestel and H. Schneider (2010), "Does Size Matter? The Impact of Changes in Household Structure on Income Distribution in Germany", *SOEP Papers 280*, DIW Berlin.

Pencavel, John (2006), "A Life Cycle Perspective on Changes in Earnings Inequality among Married Men and Women", *The Review of Economics and Statistics,* 88(2), pp. 232-242.

Saunders, Peter (2005), "Reviewing Recent Trends in Wage Income Inequality in Australia", in Isaac, Joe, Lansbury, Russell D. (eds), *Labour Market Deregulation – Rewriting the Rules*, Sydney, The Federation Press, pp. 68-88.

Schwartz, Christine (2010), "Earnings Inequality and the Changing Association between Spouses' Earnings" in *American Journal of Sociology*, Vol. 115, No. 5, pp. 1524-57.

Telegdy, Almos (2006), "The Effect of the Public Sector Wage Increase on the Public-Private Relative Wages" in *In Focus: Wages New Developments*, P. Galasi and G. Kezdi (eds.), Hungary.

Worner, Shane Mathew (2006), "The Effects of Assortative Mating on Income Inequality: A Decompositional Analysis", The Australian National University Centre for Economic Policy Research, *Discussion Paper* No. 538.

ANNEX A

Additional tables and figures

Table A1. **Decomposition of the variance of log annual earnings, mid-2000s**

	Var(ln_annual earnings)		Var(ln_hourly wages)		Var(ln_annual hours)		2xCov(ln_hwage, ln_ahours)	
	(1)		(2)		(3)		(4)	
Panel A. Countries reporting gross earnings								
Australia 2003[1]	0.460	(1.00)	0.210	(0.457)	0.255	(0.554)	-0.005	-(0.011)
Canada 2004	1.539	(1.00)	0.934	(0.607)	0.222	(0.144)	0.383	(0.249)
Czech Republic 2004[1]	0.416	(1.00)	0.300	(0.721)	0.055	(0.132)	0.061	(0.147)
Finland 2004	1.085	(1.00)	0.553	(0.510)	0.233	(0.215)	0.298	(0.275)
Germany 2004	1.089	(1.00)	0.441	(0.405)	0.333	(0.306)	0.315	(0.289)
Israel 2005[2]	0.769	(1.00)	0.504	(0.655)	0.198	(0.257)	0.066	(0.086)
Netherlands 2004	0.877	(1.00)	0.394	(0.449)	0.286	(0.326)	0.197	(0.225)
United Kingdom 2004[1]	0.700	(1.00)	0.347	(0.496)	0.229	(0.327)	0.123	(0.176)
United States 2004	0.972	(1.00)	0.600	(0.617)	0.218	(0.224)	0.154	(0.158)
Average	0.879		0.476	(0.546)	0.225	(0.276)	0.177	(0.177)
			Corr(AE, hw) = 0.91		Corr(AE, ah) = 0.43			
Panel B. Countries reporting net earnings								
Austria 2004	0.532	(1.00)	0.386	(0.726)	0.267	(0.502)	-0.121	-(0.227)
Belgium 2000	0.358	(1.00)	0.209	(0.584)	0.139	(0.388)	0.010	(0.028)
France 2000	0.654	(1.00)	0.273	(0.417)	0.308	(0.471)	0.073	(0.112)
Greece 2004	0.440	(1.00)	0.318	(0.723)	0.191	(0.434)	-0.069	-(0.157)
Hungary 2005	0.498	(1.00)	0.299	(0.600)	0.156	(0.313)	0.043	(0.086)
Ireland 2004	0.604	(1.00)	0.264	(0.437)	0.340	(0.563)	0.000	(0.000)
Italy 2004	0.326	(1.00)	0.238	(0.730)	0.137	(0.420)	-0.049	-(0.150)
Luxembourg 2004	0.582	(1.00)	0.330	(0.567)	0.200	(0.344)	0.052	(0.089)
Mexico 2004[2]	0.846	(1.00)	0.813	(0.961)	0.142	(0.168)	-0.108	-(0.128)
Spain 2004	0.529	(1.00)	0.280	(0.529)	0.208	(0.393)	0.041	(0.078)
Average	0.537		0.341	(0.627)	0.209	(0.400)	-0.013	-(0.027)
			Corr(AE, hw) = 0.78		Corr(AE, ah) = 0.31			

Note: Samples are restricted to all paid workers (aged 25-64) with positive wages and positive hours worked during the reference year. Following Blau and Kahn (2009), here we decompose the variance of logarithm annual earnings (AE) into three components: the variance of hourly wages (hw), the variance of annual hours (ah), and the covariance of the two components, as follows $Var(lnAE) = var\ (\ln_{hw}) + var(\ln_{ah}) + 2\text{cov}(\ln_{hw},\ln_{ah})$.

1. Hourly wage is calculated based on imputed weeks worked.
2. Hourly wage is calculated based on working 52 weeks. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: OECD Secretariat calculations from the Luxembourg Income Study (LIS).

Table A2. **Labour market and family formation factors impacting on household earnings inequality**

Panel A. Countries reporting gross earnings

	Australia		Canada		Czech Republic		Denmark		Finland		Germany	
	1985	2003	1987	2004	1992	2004	1987	2004	1987	2004	1984	2004
Gini of household equivalent earnings	0.32	0.35	0.34	0.39	0.28	0.34	0.28	0.3	0.3	0.32	0.31	0.35
1. Labour market factors												
Share of males working	0.85	0.8	0.91	0.89	0.85	0.86	0.89	0.86	0.9	0.85	0.9	0.87
Share of females working	0.58	0.64	0.71	0.78	0.72	0.7	0.79	0.82	0.86	0.84	0.55	0.75
Change in annual earnings of men	-2.5%		-4.2%		54.1%		9.1%		27.3%		9.6%	
At the bottom 10%	17.6%		-26.6%		33.3%		-4.6%		14.9%		-39.2%	
At the top 10%	14.1%		14.5%		70.2%		13.4%		30.8%		18.5%	
2. Family formation factors												
Household structure												
% couple w/ kids under 18	61.5	51.3	56.2	47.1	61.3	47.5	55.4	49	54.4	47.9	52.5	45.5
% couple w/o kids under 18	21.3	27.1	23.5	27.4	25.1	35.6	24.3	25.5	24.9	29.1	31.1	31.8
% single parent	5.5	9.2	6.8	8.4	5.6	6	6.4	8.9	5.3	6.7	4.3	6.9
% single or other household types	11.7	12.4	13.5	17.1	8	10.9	13.9	16.6	15.4	16.3	12.1	15.8
Assortative mating (all couple households)												
Wives' employment rates												
Husbands earnings in the top decile	0.66	0.64	0.70	0.80	0.81	0.84	0.81	0.91	0.89	0.92	0.47	0.72
Husbands earnings in the bottom decile	0.68	0.70	0.74	0.72	0.72	0.66	0.68	0.78	0.82	0.79	0.63	0.79
% of spouses in the same decile	15.7	17.4	10.8	13.1	16.8	18.2	12.3	15.1	14.3	14.3	10.6	14.6
% of spouses within (+/-2) deciles	40.3	45.6	40.8	47.7	51.3	49.9	45.6	52.2	52.3	49.2	36.0	39.4

	Israel[1]		Netherlands		Norway		Sweden		United Kingdom		United States	
	1986	2005	1987	2004	1986	2004	1981	2005	1986	2004	1986	2004
Gini of household equivalent earnings	0.4	0.43	0.31	0.33	0.26	0.32	0.29	0.33	0.33	0.38	0.37	0.42
1. Labour market factors												
Share of males working	0.83	0.78	0.81	0.86	0.95	0.90	0.94	0.89	0.80	0.81	0.82	0.87
Share of females working	0.48	0.62	0.38	0.70	0.82	0.86	0.86	0.85	0.50	0.69	0.68	0.74
Change in annual earnings of men	-1.9%		11.4%		20.1%		42.9%		44.1%		-1.3%	
At the bottom 10%	-9.5%		-32.7%		-45.1%		12.0%		19.4%		-3.3%	
At the top 10%	20.2%		14.8%		35.3%		60.0%		67.0%		17.8%	
2. Family formation factors												
Household structure												
% couple w/ kids under 18	79.4	69.8	60.0	53.1	57.3	51.9	55.5	47.8	58.5	45.9	53.7	50.4
% couple w/o kids under 18	11.1	17.6	25.5	29.6	21.2	23.8	21.5	23.4	25.0	30.8	22.0	23.1
% single parent	4.9	5.7	4.5	4.8	8.8	8.9	7.3	9.7	7.4	10.6	11.5	12.7
% single or other household types	4.6	6.9	10.0	12.5	12.7	15.4	15.8	19.1	9.2	12.7	12.8	13.8
Assortative mating (all couple households)												
Wives' employment rates												
Husbands earnings in the top decile	0.59	0.73	0.35	0.69	0.78	0.87	0.89	0.95	0.52	0.69	0.52	0.64
Husbands earnings in the bottom decile	0.32	0.63	0.33	0.65	0.83	0.80	0.80	0.82	0.49	0.67	0.69	0.71
% of spouses in the same decile	15.4	21.5	14.3	13.9	8.6	12.6	10.4	13.7	20.2	17.5	14.2	12.4
% of spouses within (+/-2) deciles	44.9	51.4	35.5	43.9	39.1	47.8	44.0	50.2	43.2	48.3	42.4	45.5

Table A2. **Labour market and family formation factors impacting on household earnings inequality** *(cont.)*

Panel B. Countries reporting net earnings

	Austria		Belgium		France		Greece		Hungary		Ireland	
	1994	2004	1985	2000	1984	2000	1995	2004	1994	2005	1994	2004
Gini of household equivalent earnings	0.33	0.33	0.26	0.28	0.33	0.36	0.34	0.35	0.41	0.39	0.37	0.36
1. Labour market factors												
Share of males working	0.82	0.85	0.80	0.73	0.71	0.85	0.80	0.84	0.62	0.71	0.81	0.80
Share of females working	0.56	0.67	0.41	0.59	0.48	0.70	0.38	0.51	0.54	0.58	0.47	0.62
Change in annual earnings of men		2.4%		27.4%		-2.0%		23.4%		2.0%		34.2%
At the bottom 10%		8.3%		13.4%		-25.0%		18.0%		20.8%		86.6%
At the top 10%		18.1%		30.0%		7.6%		20.9%		-16.9%		29.1%
2. Family formation factors												
Household structure												
% couple w/ kids under 18	53.9	48.8	52.2	55.9	59.0	56.5	57.0	51.7	55.9	46.9	70.1	55.2
% couple w/o kids under 18	29.0	29.9	38.9	27.5	28.7	26.0	29.6	33.7	26.1	31.3	15.4	51.7
% single parent	5.5	5.6	2.7	6.3	4.1	6.7	3.0	2.7	6.8	6.0	5.4	11.3
% single or other household types	11.6	15.8	6.2	10.3	8.2	10.8	10.4	11.9	11.2	15.8	9.2	11.8
Assortative mating (all couple households)												
Wives' employment rates												
Husbands earnings in the top decile	0.58	0.69	0.34	0.52	0.44	0.65	0.37	0.46	0.69	0.54	0.30	0.61
Husbands earnings in the bottom decile	0.52	0.72	0.53	0.54	0.49	0.65	0.34	0.46	0.56	0.59	0.43	0.52
% of spouses in the same decile	17.8	18.1	17.8	22.2	23.8	16.2	16.0	17.5	25.3	25.7	18.1	17.6
% of spouses within (+/-2) deciles	45.6	46.0	40.0	49.1	50.6	50.4	43.6	49.0	55.3	55.0	43.9	44.2

	Italy		Luxembourg		Mexico		Poland		Spain	
	1987	2004	1985	2004	1984	2004	1992	2004	1990	2004
Gini of household equivalent earnings	0.33	0.36	0.25	0.34	0.45	0.47	0.33	0.41	0.33	0.33
1. Labour market factors										
Share of males working	0.83	0.82	0.84	0.87	0.93	0.92	0.67	0.78	0.91	0.86
Share of females working	0.38	0.54	0.31	0.62	0.31	0.48	0.48	0.60	0.31	0.59
Change in annual earnings of men		0.7%		72.3%		9.1%		0.6%		8.5%
At the bottom 10%		-18.5%		16.1%		73.9%		-27.7%		7.6%
At the top 10%		1.8%		115.0%		24.9%		17.5%		13.4%
2. Family formation factors										
Household structure										
% couple w/ kids under 18	60.3	48.3	60.1	54.2	83.8	70.6	69.9	60.7	68.5	50.2
% couple w/o kids under 18	30.6	35.0	27.9	30.5	4.1	10.6	17.0	23.6	23.0	36.8
% single parent	2.4	3.1	4.1	4.0	9.8	14.0	5.2	6.7	3.3	3.3
% single or other household types	6.7	13.5	7.9	11.4	2.3	4.8	7.9	9.0	5.2	9.7
Assortative mating (all couple households)										
Wives' employment rates										
Husbands earnings in the top decile	0.33	0.63	0.25	0.53	0.20	0.46	0.33	0.54	0.32	0.56
Husbands earnings in the bottom decile	0.43	0.41	0.50	0.71	0.30	0.36	0.33	0.49	0.34	0.60
% of spouses in the same decile	17.0	20.6	12.5	14.5	9.7	11.1	13.8	17.8	15.8	17.2
% of spouses within (+/-2) deciles	45.2	53.8	33.8	44.6	35.1	40.7	40.1	47.3	38.0	46.6

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

Table A3. **Weights used in the density decomposition**

A. Primary-order decomposition

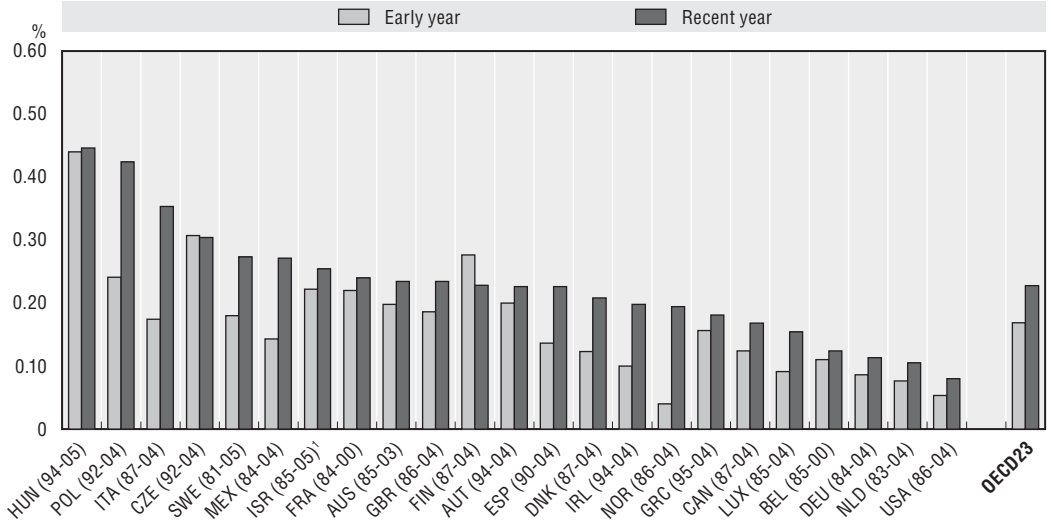
Order	Counterfactual distribution of equivalent household earnings in 2005 held constant factors to 1985 levels in the following order	Equivalent household earnings measure	Weight
	1. Original 2005 distribution	Y_{05}	W_{05}
M	2. (1) with 1985 men's earnings	Y_{85}^{m}	W_{05}
L^M	3. (2) with 1985 male emp. rate	Y_{85}^{m}	$W_{05} \cdot \lambda_{L^M\|L^F,A,S}$
L^F	4. (2) with 1985 female emp. rate	Y_{85}^{m}	$W_{05} \cdot \lambda_{L^M\|L^F,A,S} \cdot \lambda_{L^F\|A,S}$
A	5. (3) with 1985 assortative mating	Y_{85}^{m}	$W_{05} \cdot \lambda_{L^M\|L^F,A,S} \cdot \lambda_{L^F\|A,S} \cdot \lambda_{A\|S}$
S	6. (4) with 1985 household structure	Y_{85}^{m}	$W_{05} \cdot \lambda_{L^M\|L^F,A,S} \cdot \lambda_{L^F\|A,S} \cdot \lambda_{A\|S} \cdot \lambda_{S}$

B. Reverse-order decomposition

	1. Original 2005 distribution	Y_{05}	W_{05}
S	2. (1) with 1985 household structure	Y_{05}	$W_{05} \cdot \lambda_{S\|A,L^F,L^M}$
A	3. (2) with 1985 assortative mating	Y_{05}	$W_{05} \cdot \lambda_{S\|A,L^F,L^M} \cdot \lambda_{A\|L^F,L^M}$
L^F	4. (3) with 1985 female emp. rate	Y_{05}	$W_{05} \cdot \lambda_{S\|A,L^F,L^M} \cdot \lambda_{A\|L^F,L^M} \cdot \lambda_{L^F\|L^M}$
L^M	5. (3) with 1985 male emp. rate	Y_{05}	$W_{05} \cdot \lambda_{S\|A,L^F,L^M} \cdot \lambda_{A\|L^F,L^M} \cdot \lambda_{L^F\|L^M} \cdot \lambda_{L^M}$
M	6. (4) with 1985 men's earnings	Y_{85}^{m}	$W_{05} \cdot \lambda_{S\|A,L^F,L^M} \cdot \lambda_{A\|L^F,L^M} \cdot \lambda_{L^F\|L^M} \cdot \lambda_{L^M}$

Note: Y_{05} refers to equivalent household earnings in 2005, W_{05} is the original survey sampling weights, and λs are estimated conditioning weightings.

Figure A1. **Correlation coefficients between husband and wife's earnings, couple households with at least one person working**



Note: Figures refer to couple working-age households (head aged 25-64 years old) with at least one partner working.
1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.
Source: Authors' calculations from the Luxembourg Income Study (LIS).

Figure A2. **Working wives' annual earnings by husband's earnings decile, couple households, mid-1980s and mid-2000s**

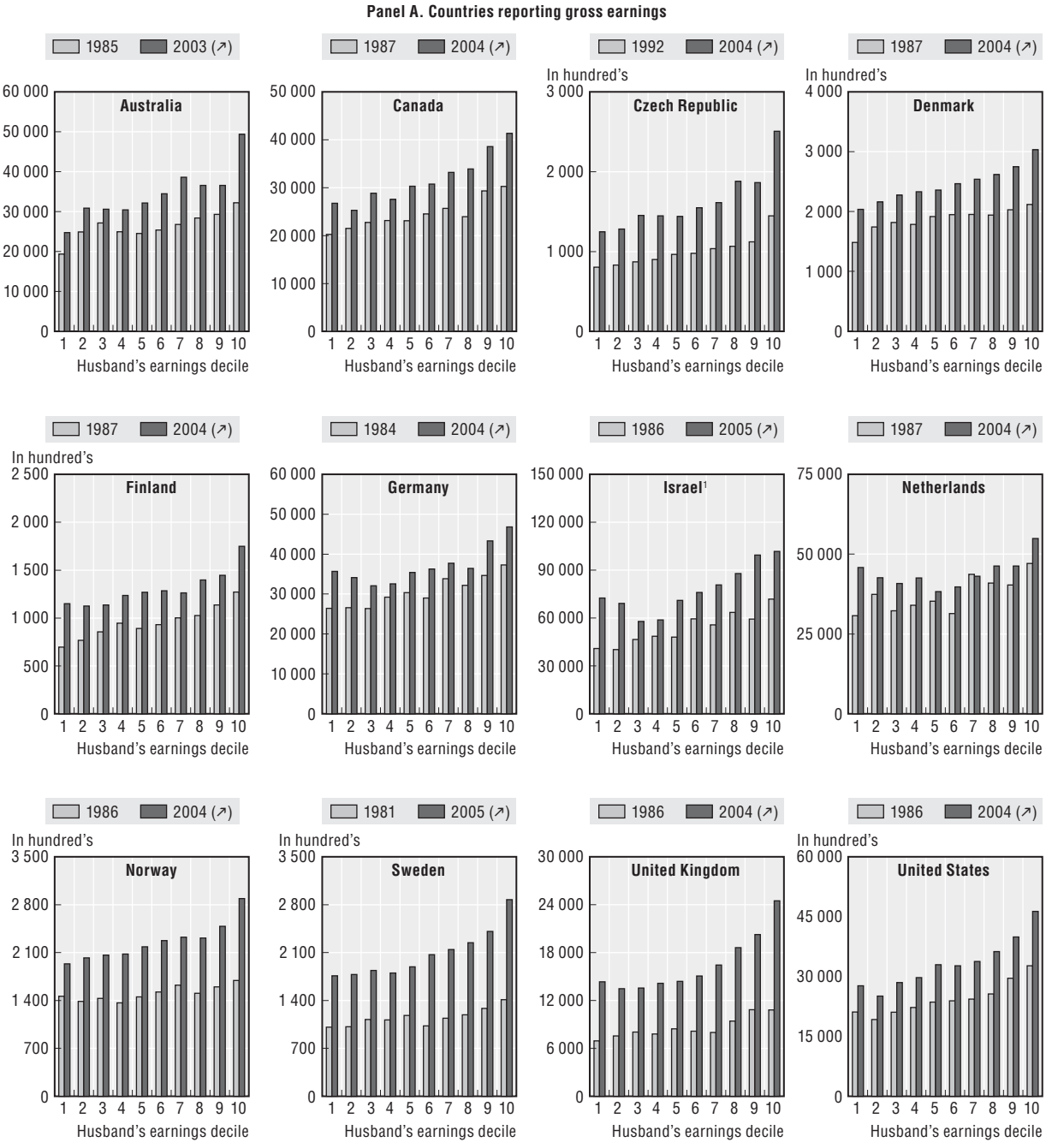
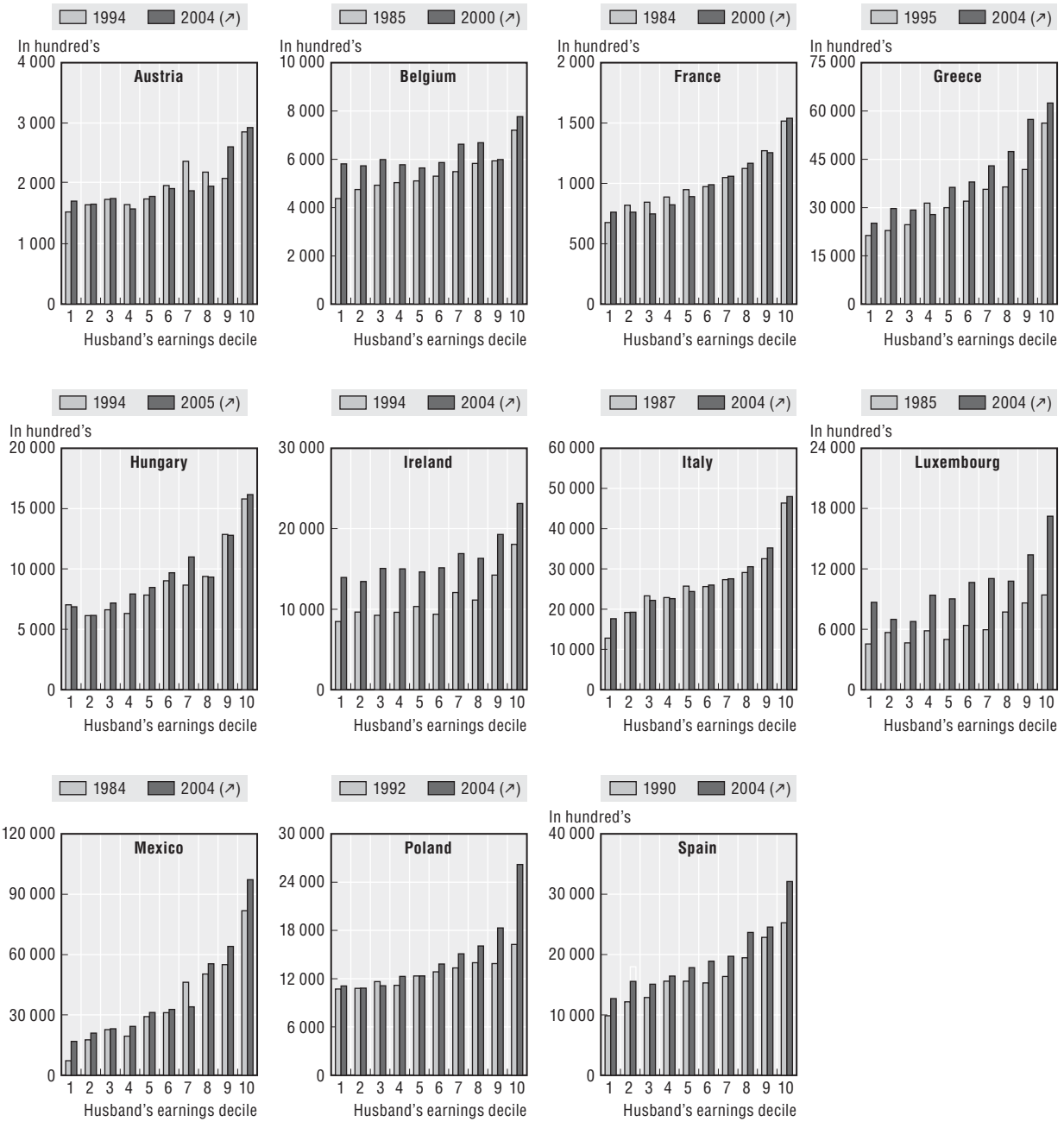


Figure A2. **Working wives' annual earnings by husband's earnings decile, couple households, mid-1980s and mid-2000s** *(cont.)*

Panel B. Countries reporting net earnings



Note: Figures refer to couple working-age households (head aged 25-64 years old) with both partners working. Amounts are in national currencies (constant values of 2005).

1. Information on data for Israel: *http://dx.doi.org/10.1787/888932315602*.

Source: Authors' calculations from the Luxembourg Income Study (LIS).

OECD Journal: Economic Studies
Volume 2013
© OECD 2014

Towards global carbon pricing: Direct and indirect linking of carbon markets

by
Rob B. Dellink, Stéphanie Jamet, Jean Chateau *and* Romain Duval*

Emissions trading systems (ETS) can play a major role in a cost-effective climate policy framework. Both direct linking of ETSs and indirect linking through a common crediting mechanism can reduce costs of action. We use a global recursive-dynamic computable general equilibrium model to assess the effects of direct and indirect linking of ETS systems across world regions. Linking of domestic Annex I ETSs leads to moderate aggregate cost savings, as differences in domestic permit prices are limited. Countries benefit directly from linking by either buying permits and avoiding investing in high-cost mitigation options, or by exploiting relatively cheap mitigation options and selling permits at a higher price. Although the economy of the main permit sellers, such as Russia, is negatively affected by the real exchange rate appreciation that is induced by the large export of permits, on balance they also still benefit from linking. The cost-saving potential for developed countries of well-functioning crediting mechanisms appears to be very large. Even limited use of credits would nearly halve mitigation costs; cost savings would be largest for carbon-intensive economies. However, one open issue is whether these gains can be fully reaped in reality, given that direct linking and the use of crediting mechanisms both raise complex system design and implementation issues. The analysis in this paper shows, however, that the potential gains to be reaped are so large, that substantial efforts in this domain are warranted.

JEL classification: H23, O41, Q54

Keywords: Climate mitigation policy, emissions trading systems, general equilibrium models, linking carbon markets

* Contact: Rob Dellink (rob.dellink@oecd.org), OECD Environment Directorate. The authors would like to thank Jean-Marc Burniaux, Jan-Corfee-Morlot, Helen Mountford, Cuauhtemoc Rebolledo-Gomez, Sébastien Jean and Nick Johnstone for their input, suggestions and comments. The views expressed in this paper are those of the authors and do not necessarily represent the views of the OECD or of its member countries.

Note by Turkey: The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue".

Note by all the European Union member states of the OECD and the European Union: The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

Putting a price on greenhouse gas (GHG) emissions through price mechanisms such as carbon taxes and emissions trading systems (ETSs) can go a long way towards building up a cost-effective climate policy framework. Indeed these price instruments are intrinsically cost effective as they equalise marginal abatement costs across individual emitters while giving them continuing incentives to search for cheaper abatement options through both existing and new technologies.

While experience with GHG emissions taxes remains limited (e.g. Sweden), the international community has made some progress towards implementing ETSs. Several domestic/regional GHG emission trading or cap-and-trade schemes are already in place (including the EU ETS in the EU, RGGI in the United States, the Norwegian ETS and emission trading in Japan) or are emerging (see OECD, 2012 for a full overview). These vary significantly in terms of their target, size, and other design features. At present there are virtually no direct links between these systems, other than the link between the EU and Norwegian ETSs. However, in 2012, the European Commission and Australia agreed to start linking the newly-introduced Australian carbon price and the EU ETS from 2015 and to achieve full linking by 2018. As more ETSs are expected to emerge in the future, direct linking is likely to gain prominence.

Linking can also occur "indirectly" when multiple ETSs allow part of the emission reductions to be achieved in countries outside the ETS, for example through a common offset or crediting mechanism. The existence of "flexibility mechanisms" (such as the Clean Development Mechanism (CDM) of the Kyoto Protocol) enable emission-reduction commitments under the ETSs to be met by undertaking emission reductions – often referred to as offsets – in other geographical areas. When different ETSs accept credits from a common crediting mechanism, this induces some convergence in permit prices and thereby some indirect linking between the ETSs.

Compared with a fragmented approach, under which a number of regions would meet their emission reduction objectives in isolation, direct or indirect linking of ETSs can reduce mitigation costs by fostering partial or even full (in the case of direct linking) convergence in carbon prices, and thus in marginal abatement costs, across the different ETSs (Jaffe and Stavins, 2007, 2008). Linking can also reduce carbon leakage, which arises when emission reductions in one set of countries are partly offset by increases in the rest of the world. This is straightforward for indirect linking, which lowers carbon prices in all regions covered by ETSs. Direct linking may also reduce carbon leakage by alleviating the very high carbon prices and the associated leakage that could otherwise prevail in certain regions. Other significant, but difficult to quantify, gains from (direct) linking arise from enhanced liquidity of permit markets.

The economic aspects of linking ETSs have received a fair amount of attention in the literature already. Tuerk et al. (2009) provide a good non-technical introduction to the main issues involved. A more fundamental analysis of different architectures of linking is given in e.g. Flachsland et al. (2009) and Jaffe and Stavins (2007, 2008). Grull and Taschini (2012)

analyse the conditions under which linking will lead to full convergence in carbon prices. Anger (2008) provides a numerical assessment of linking the EU ETS to other ETS systems and investigates the role of the CDM, but uses a fairly limited two-sector partial equilibrium model. Mehling et al. (2011) assess whether the main features of the US regional ETSs would be an obstacle to linking with the EU ETS.

This paper aims at numerically assessing the economic effects of direct and indirect linking of ETSs in the Annex I regions.[1] To this end, we use the OECD global recursive-dynamic computable general equilibrium model ENV-Linkages. ENV-Linkages provides numerical projections for GHG emissions, economic activity and economic growth in a multi-sector, multi-gas setting for the major world regions. The global general-equilibrium approach ensures that all major feedback mechanisms that link the economies are taken into account.

Our results first confirm the general rule that the greater the difference in carbon prices across countries prior to linking, the larger the cost savings from linking. Countries with higher pre-linking carbon prices gain from abating less and buying cheaper permits. Countries with lower pre-linking prices benefit from abating more and selling permits, although their economy may be negatively affected by the real exchange-rate appreciation triggered by the large permit exports (the Dutch disease effect). Under an illustrative 20% emission cut (relative to 1990 levels) in each Annex I region by 2020, if domestic Annex I ETSs were linked, permit buyers would include Canada, Australia-New Zealand and, to a lesser extent, the European Union and Japan. Russia would be the main seller.

Secondly, the cost-saving potential for developed countries of well-functioning crediting mechanisms appears to be very large, reflecting the vast low-cost abatement potential in a number of emerging and developing countries. Even limited use of CERs amounting to 20% of Annex I emission reductions would already almost halve mitigation costs in these countries, and raising this cap on offset credit use would bring further cost reductions. Cost savings would be largest for the more carbon-intensive Annex I economies, such as Australia, Canada and Russia. China has the potential to be by far the largest seller, and the United States the largest buyer in the offset credit market, each of them accounting for about half of transactions by 2020 under the illustrative mitigation scenario mentioned above.

Thirdly, the mitigation cost savings from direct linking of ETSs in Annex I countries are much smaller than those from indirect linking through a well-functioning crediting mechanism. This finding essentially reflects the much greater heterogeneity in marginal abatement costs between Annex I and non-Annex I countries than among Annex I countries themselves, as well as the much greater low-cost abatement potential available in non-Annex I countries. In fact, under the illustrative scenario considered in this paper, allowing Annex I regions to meet up to 50% of their domestic commitments through the use of offsets would trigger major carbon price convergence, thereby exhausting the gains from linking and making direct linking of little additional value.

These gains from linking are unlikely to be fully reaped in reality. Both direct linking and the use of crediting mechanisms require careful set-up of the regulatory issues. In particular, setting up a well-functioning sectoral or other large-scale crediting mechanism may prove a major challenge. The analysis in this paper shows, however, that the potential gains to be reaped are so large that substantial efforts in this domain are warranted.

The set-up of this paper is as follows: Section 1 briefly outlines the methodology used, with a focus on the specification of the ENV-Linkages model. Section 2 presents the results

of the numerical simulations and identifies the main outcomes of the analysis. Some sensitivity analysis is carried out in Section 3. Section 4 discusses how direct and indirect linking could be established in practice, while Section 5 concludes.

1. Methodology

1.1. Short description of the ENV-Linkages model

The OECD ENV-Linkages model is a recursive-dynamic neo-classical Computable General Equilibrium (CGE) model. It is a global economic model to study the economic impacts of environmental policies. The ENV-Linkages model is the successor to the OECD GREEN model for environmental studies (Burniaux et al., 1992). In the version of the model used here (ENV-Linkages Version 2.1, as documented in OECD, 2009), the model represents the world economy in 12 countries/regions, each with 25 economic sectors (see the Annex A), including five different technologies to produce electricity. A fuller model description is given in OECD (2009); here we just present the main features of the model.

All production in ENV-Linkages is assumed to operate under cost minimisation with an assumption of perfect markets and constant return-to-scale technology. The production technology is specified as nested CES production functions in a branching hierarchy (cf. Annex A). This structure is replicated for each output, while the parameterisation of the CES functions may differ across sectors. Adjustments are made for specific sectors, including the agricultural sectors and non-fossil fuel based electricity technologies by specifying a factor endowment specific to the sector that represents natural resource use. The nesting of the production function for the agricultural sectors is further re-arranged to reflect substitution between intensification (e.g. more fertiliser use) and extensification (more land use) of activities.

The model adopts a putty/semi-putty technology specification, where substitution possibilities among factors are assumed to be higher with new vintage capital than with old vintage capital. This implies relatively slow adjustment of quantities to price changes. Capital accumulation is modelled as in the traditional Solow/Swan neo-classical growth model.

The energy bundle is of particular interest for analysis of climate change issues. Energy is a composite of fossil fuels and electricity. In turn, fossil fuel is a composite of coal and a bundle of the "other fossil fuels". At the lowest nest, the composite "other fossil fuels" commodity consists of crude oil, refined oil products and natural gas. The value of the substitution elasticities are chosen as to imply a higher degree of substitution among the other fuels than with electricity and coal.

The structure of electricity production assumes that a representative electricity producer maximises its profit by using the five available technologies to generate electricity using a CES specification with a large value of the elasticity of substitution. The production of the non-fossil electricity technologies (net of GHG and expressed in Terawatt per hour) has a structure similar to the other sectors, except for a Leontief-structure top nesting combining a sector-specific natural resource, on the one hand, and all other inputs, on the other. This specification acts as a capacity constraint on the supply of these electricity technologies as the supply of the sector-specific natural resource is limited.

Household consumption demand is the result of static maximisation behaviour which is formally implemented as an "Extended Linear Expenditure System" (Lluch, 1973; Howe, 1975). A representative consumer in each region – who takes prices as given – optimally allocates disposable income among the full set of consumption commodities and savings. Saving is

considered as a standard good and therefore does not rely on a forward-looking behaviour by the consumer. The government in each region collects various kinds of taxes in order to finance government expenditures. Assuming fixed public savings (or deficits), the government budget is balanced through the adjustment of the income tax on consumer income.

International trade is based on a set of regional bilateral flows. The model adopts the Armington specification, assuming that domestic and imported products are not perfectly substitutable. Moreover, total imports are also imperfectly substitutable between regions of origin. Allocation of trade between partners then responds to relative prices at the equilibrium.

Market goods equilibria imply that, on the one side, the total production of any good or service is equal to the demand addressed to domestic producers plus exports; and, on the other side, the total demand is allocated between the demands (both final and intermediary) addressed to domestic producers and the import demand.

ENV-Linkages is fully homogeneous in prices and only relative prices matter. All prices are expressed relative to the *numéraire* of the price system that is arbitrarily chosen as the index of OECD manufacturing export prices. Each region runs a current-account balance, which is fixed in terms of the numéraire. One important implication from this assumption in the context of this paper is that real-exchange rates immediately adjust to restore current-account balance when countries start exporting/importing emission permits.

Emissions of CO_2 from combustion of fossil fuels are linked in ENV-Linkages directly to the related energy use by the different sectors. Process-related emissions of CO_2 are linked to the output of the emitting sector. Similarly, emissions of other greenhouse gases (N_2O, CH_4, PFCs, HFCs and SF_6) are linked to economic activity in the various sectors. In this way, emissions increase when the related activities expand. Over time, there is, however, some relative decoupling of emissions from the underlying economic activity through autonomous technical progress, implying that emissions grow less rapidly than economic activity.

Emissions can be abated through three channels: *i)* reductions in emission intensity of economic activity; *ii)* changes in structure of the associated sectors away from the "dirty" input to cleaner inputs; and *iii)* changes in economic structure away from relatively emission-intensive sectors to cleaner sectors. The *first* channel, which is not available for emissions from combustion of fossil fuels, entails end-of-pipe measures that reduce emissions per unit of the relevant input. The *second* channel includes, for instance, substitution from fossil fuels to renewable in electricity production, or investing in more energy-efficient machinery (which is represented through higher capital inputs but lower energy inputs in production). An example of the *third* channel is a substitution from consumption of energy-intensive industrial goods to services. In the model, the choice between these three channels is endogenous and driven by the price on emissions.

1.2. Calibration and baseline projection

The process of calibration of the ENV-Linkages model is broken down into three stages. *First*, a number of parameters are calibrated, given some elasticity values, on base-year (2001) values of variables. This process is referred to as the static calibration. Each of the 12 regions is underpinned by an economic input-output table for 2001, based on the *GTAP 6.2 Database* (Dimaranan, 2006). *Second*, the 2001 database is updated to 2005 by simulating the model dynamically to match historical trends over the period 2001-05; thus all variables are expressed in 2005 real USD. *Third*, the business-as-usual (BAU) baseline projection is obtained by defining a set of exogenous socio-economic drivers (demographic

trends, labour productivity, future trends in energy prices and energy efficiency gains) and running the model dynamically again over the period 2005-2050. The main elements in designing the BAU projection are described in detail in OECD (2009) and Duval and de la Maisonneuve (2010).

The BAU baseline scenario assumes that there are no new climate change policies implemented, and projects future emissions on the basis of assumptions on the long-term evolution of output growth, relative prices of fossil fuels and potential gains in energy efficiency. It thus provides a benchmark against which policy scenarios aimed at achieving emission cuts can be assessed.

The BAU projection is based on the so-called conditional convergence assumption that total factor productivity in developing countries converges towards those in developed countries over the coming decades (for details, see Duval and de la Maisonneuve, 2010). Average annual world GDP growth (in constant purchasing power parity – PPP – in 2005 USD) is assumed to be around 3.5% between 2006 and 2050. This is slightly lower than the 2000-06 average. Overall, average world GDP per capita in constant PPP USD is expected to rise more than three times between 2006 and 2050.

World emissions of the greenhouse gases carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), hydro fluorocarbons (HFCs), per fluorocarbons (PFCs) and sulphur hexafluoride (SF_6) have roughly doubled since the early 1970s, and are expected to double again, reaching about 72 gigatons CO_2 equivalent (Gt CO_2eq) in 2050, excluding emissions from land-use, land-use change and forestry. Critical drivers of projected emissions for the period until 2050 other than GDP growth include assumptions about future fossil fuel prices and energy efficiency gains in line with IEA projections (IEA, 2008). Finally, the BAU projection assumes that the EU Emissions Trading Scheme (EU ETS) will be sustained in the future, with a gradual convergence in the carbon price to USD 25 per tonne of CO_2 and a stabilisation at this level (in real terms) beyond 2012. Marginal abatement costs for emissions from fossil-fuel combustion are driven by the detailed representation of energy use in the model, as emission factors per unit of energy use are fixed (but varying across energy carriers). For the other emission sources, marginal abatement cost curves are based on information from US EPA and approximated through a substitution between output and emissions, following Hyman et al. (2002).

2. Simulation results

In order to illustrate the effects of linking different regional ETSs, the ENV-Linkages model is used here to run an illustrative mitigation scenario (A1 act). In the case of this scenario without linking, each Annex I region is assumed to use an ETS to cut its GHG emissions unilaterally below 1990 levels by 20% by 2020 and by 50% by 2050. We assume that permits are fully auctioned to all sectors in all countries covered by the ETS. On its own, this commitment would be insufficient to achieve ambitious climate objectives. World emissions would still rise by about 20% and 50% by 2020 and 2050 respectively, vs. about 85% by 2050 in a baseline scenario with no further mitigation policy action (BAU). It would, therefore, need to be fairly rapidly tightened and/or supplemented with further action, including in non-Annex I countries if a stringent long-term stabilisation target is to be met. This illustrative A1 act scenario does not directly reflect the emission reductions as pledged by Annex I countries in the Copenhagen Accord (cf. Dellink et al., 2011) but is more stylised, in order to shed maximum light on a number of lessons about the economic impacts of linking.

2.1. Direct linking of carbon markets in the Annex I regions

Direct linking occurs if the tradable permit system's authority allows regulated entities to use emission allowances from another ETS to meet their domestic compliance obligations.[2] Direct linking can be "two way" if each system recognises the others' allowances, or "one way" if one system recognises the other system's allowances but the other does not reciprocate. Linking ETSs directly tends to lower the overall cost of meeting their joint targets by allowing higher cost emission reductions in one ETS to be replaced by lower cost emission reductions in the other. Once ETSs are linked, this cost effectiveness is achieved regardless of the magnitude of the initial emission-reduction commitment across countries or regions as trading of emission allowances ensures that marginal abatement costs of individual emitters are equalised. The initial allocation of allowances does have a distributional effect as it determines the direction of the trade flows. The potential gains from linking are greater the larger the initial difference in carbon prices – and thereby in the marginal costs of reducing emissions – across individual ETSs.

Another advantage of linking is its ability to reduce competitiveness concerns in regions with higher pre-linking carbon prices. By allowing carbon prices to converge across linked schemes the distortion between domestic and foreign prices would be reduced and thus the loss in domestic output would be limited. Indeed full convergence in prices will be achieved provided the recognition of allowances is mutual and there are no limits on trading.[3]

These two key elements of linking carbon markets, i.e. improving cost-effectiveness and reducing competitiveness concerns, will be quantitatively assessed for the stylised scenario simulations in the next two sub-sections. Linking can also deliver a number of other benefits that go beyond the scope of the gains that can be illustrated with a CGE model like ENV-Linkages, such as strengthening the carbon markets through larger trading volumes. While the importance of these cannot be quantified, the other main benefits, and some reasons for concern, will be discussed in the third sub-section below.

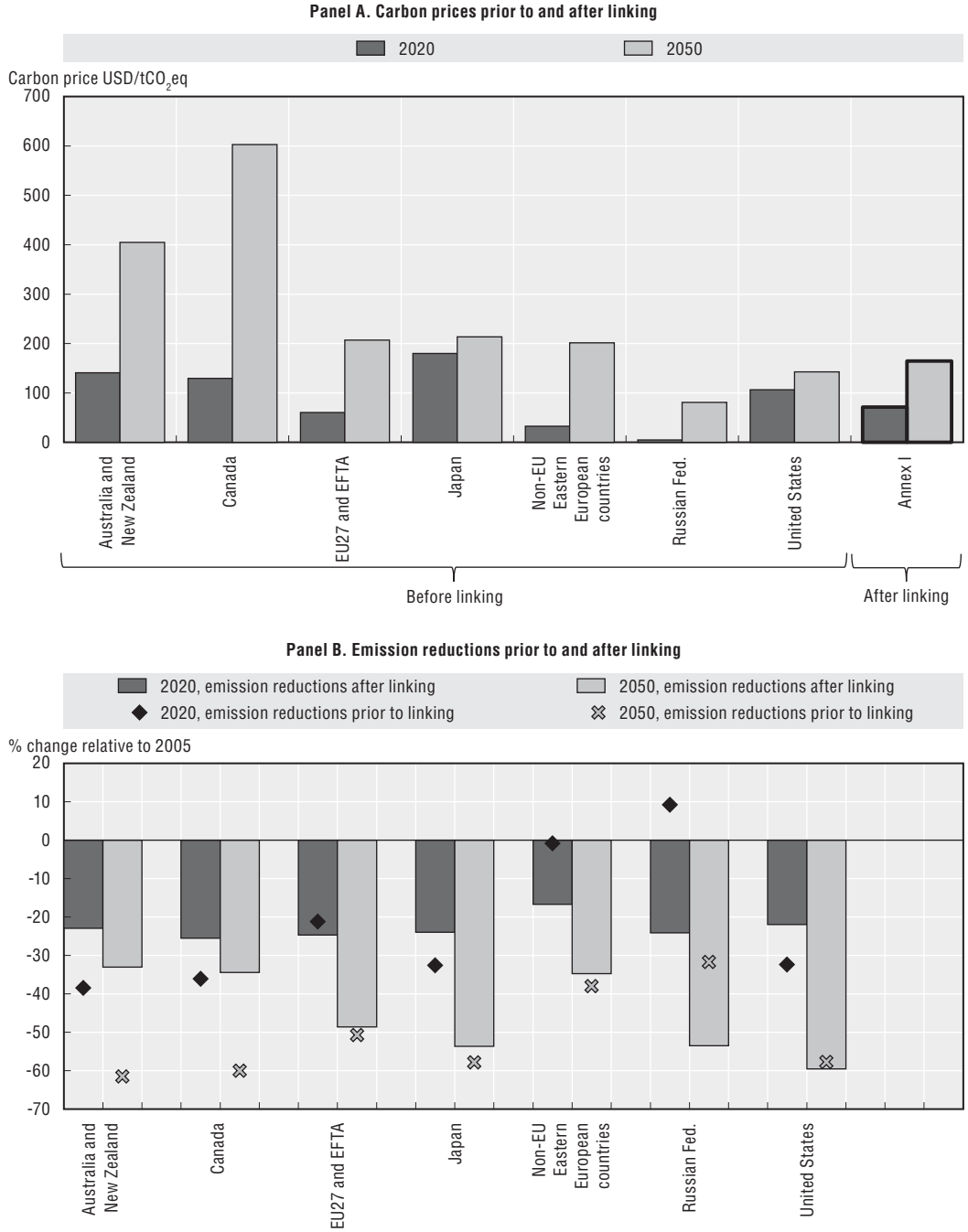
2.1.1. Improving cost-effectiveness

Simulations of the A1 act scenario under alternative assumptions about trading possibilities ("linked" or "not linked") among Annex I countries show that the reduction in mitigation costs is fairly limited in our illustrative scenario. This is mainly because carbon price differences prior to linking are estimated to be relatively small across the larger Annex I economies that account for the bulk of Annex I GDP (Figure 1, Panel A). Linking substitutes additional emission reductions in regions that had lower marginal abatement costs before linking (especially Russia) (Figure 1, Panel B) for emission increases in the others (Figure 1, Panel B).

In the unlinked case, meeting their domestic caps alone is found to cost Annex I regions about 1.5% and 2.8% of their income equivalent variation on average by 2020 and 2050, respectively (Figure 2). The associated reduction in overall mitigation costs from linking their carbon markets for Annex I countries is just under 10%, or about 0.2 percentage points of income by 2050 (Figure 2, Panel B). Larger gains from linking could be found under more heterogeneous emission-reduction commitments by Annex I countries than considered here.[4]

In addition to improving the overall cost effectiveness of the linked ETS system, linking is expected to benefit each participating region (e.g. Jaffe and Stavins, 2007). The larger the change in the carbon price after linking, the larger the income gain, other things being equal. In turn, the carbon price level prior to linking depends on the size of the country's

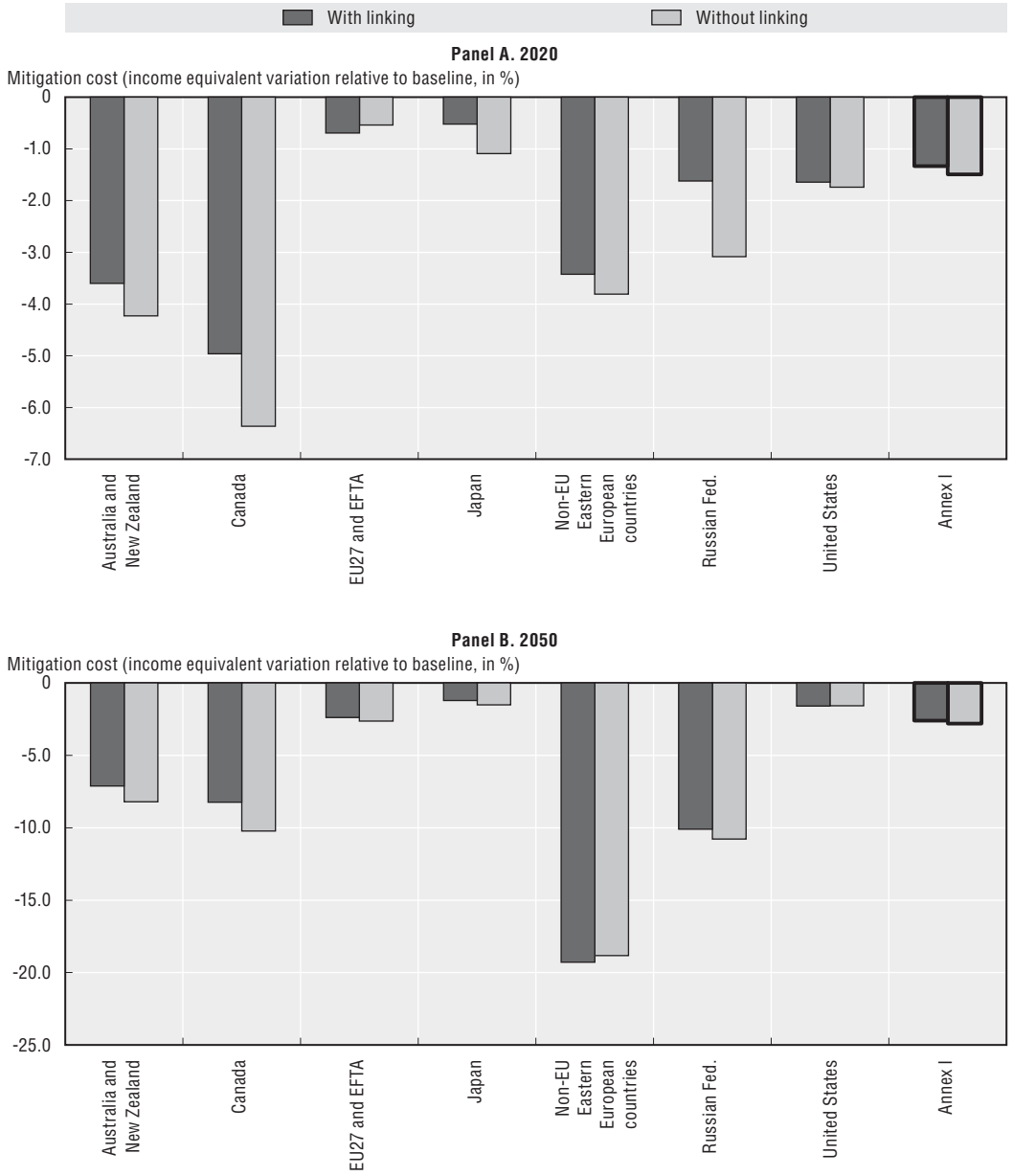
Figure 1. **Impact of linking regional Annex I ETSs on carbon prices and emission reductions in the A1 act scenario**



Note: In the A1 act scenario, Annex I regions reduce their GHG emissions below 1990 levels by 20% in 2020 and 50% in 2050.

commitment, as well as on the availability of cheap abatement opportunities. In the A1 act scenario considered here, countries with lower pre-linking carbon prices (mainly Russia) gain because the equilibrium price of the linked system exceeds their pre-existing marginal abatement costs, enabling them to abate more and sell the saved permits with a surplus while, conversely, regions with higher pre-linking carbon prices (Australia-New Zealand, Canada, Japan) benefit from the lower carbon price.

Figure 2. **Impact of linking regional Annex I ETSs on mitigation policy costs under the A1 act scenario**



While such basic reasoning suggests that permit trading among Annex I regions could benefit all participants, the multi-region case is more complicated than evaluating two-region linking. As countries cannot prevent others from linking, and some regions will always benefit from linking their carbon markets, they effectively are faced with a choice of joining a group of countries that have already linked or remaining outside. *Ceteris paribus*, the regions that are linking strengthen their competitive position *vis-à-vis* outsiders, and it may then become rational to join this group, even if the outcome is worse than in the starting situation where none of the systems are linked. These countries cannot easily avoid this welfare reduction, as the alternative, where the other Annex I countries link their markets but they remain outside this joint scheme is also not beneficial for them.[5]

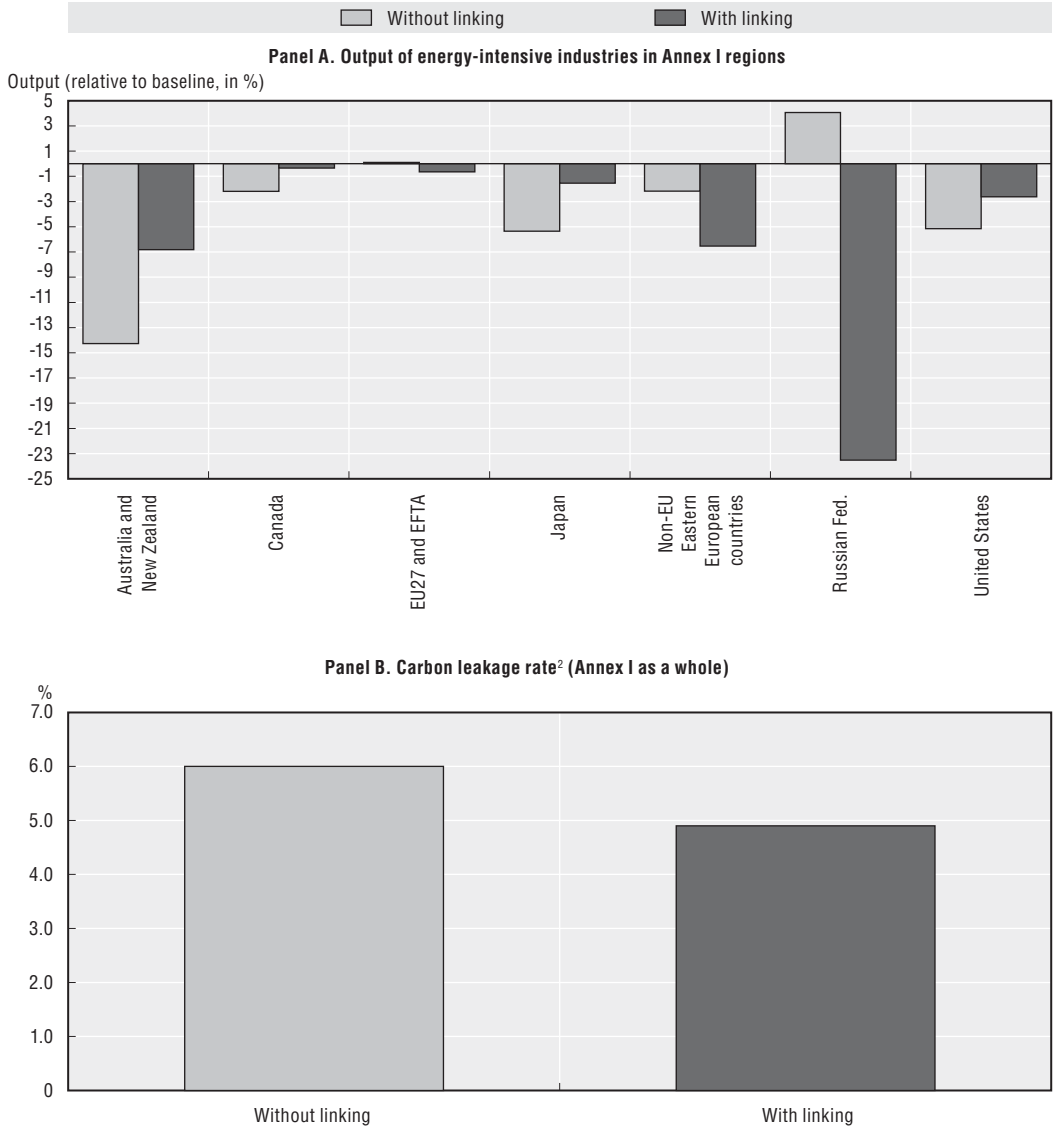
Various market imperfections complicate the picture. In our model simulation, non-EU Eastern European countries – which together form the "Rest of Annex I" region of the ENV-Linkages model – are actually found to lose on balance from linking by 2050 (Figure 2). This is partially because their large permit exports lead to a real exchange-rate appreciation, which in turn results in a fall in the exports and output of their manufacturing sector, where scrapping capital entails costs. Nevertheless, these "Dutch disease" effects are particularly large in the ENV-Linkages model, partly due to the assumption of fixed net international capital flows. The OECD ENV-Linkages model incorporates many market imperfections and distortions and, therefore, the impact of permit trading on each participating region has to be interpreted in a second-best context. One important consequence of this is that capital allocation in the BAU scenario is not optimal. As countries sell permits abroad, imports must rise and/or other exports must decline in order to satisfy the exogenous balance-of-payments constraint. Restoring the external balance requires an appreciation of the real exchange rate, which triggers costly reallocation of capital across sectors, reduces aggregate output and, in some cases, investments. This latter effect reduces future consumption. In the specific case of the non-EU Eastern European countries, their emission profile is such that in later decades they switch from being permit seller to being permit buyer. But the dynamic effects of worsened terms of trade and reduced capital accumulation then imply a negative impact on income and welfare. As ENV-Linkages is a recursive-dynamic model, there is no intertemporal evaluation of the net present value of such variations in consumption.

2.1.2. Mitigating competitiveness concerns

Linking "levels the playing field" by avoiding the most expensive emission reductions in favour of buying permits, and exploiting cheap reduction options where available. This benefits especially those sectors that are emission-intensive and trade-exposed. In practice, however, the (real gross) output losses of energy-intensive industries[6] would still be unevenly distributed across countries after linking. Those countries with the lowest pre-linking marginal abatement costs (Russia and non EU Eastern European countries, Figure 1, Panel A) face the largest losses (Figure 3, Panel A). In the unlinked case, the low carbon price in Russia implies a relative improvement in the international competitive position of the energy-intensive industries compared with its Annex I competitors. After linking, the higher carbon price in Russia affects the output of the energy-intensive industry negatively, and moreover it is profitable to reduce activity in these sectors in order to sell more emission permits on the international market.

Carbon leakage can arise: *i*) when emission-intensive, trade-exposed industries in countries with stringent mitigation policies lose market share in favour of industries located in countries that are faced with less stringent or no mitigation policies; or *ii*) when agents (consumers and producers) substitute away from emission-intensive goods in acting countries and reduce the world demand for fossil fuels. This, in turn, reduces world market prices of fossil fuels, resulting in an increased demand for fossil fuels and higher emissions in non-acting countries. The simulations find that leakage rates are relatively small, in line with the finding of Burniaux et al. (2013) that leakage rates rapidly decline when more countries are acting as competitiveness impacts become weaker. Linking across ETSs does not affect the total emissions of the linked schemes since the number of permits is simply the sum of those issued under each system, at least in the absence of strategic behaviour. However, linking can still affect carbon leakage towards uncapped countries. If linking lowers

Figure 3. **Impact of linking Annex I regional ETSs on carbon leakage and the output losses of energy-intensive industries[1] across linked regions under the A1 act scenario in 2020**



1. Energy-intensive industries include chemicals, metallurgic, other metal, iron and steel industry, paper and mineral products.
2. The carbon leakage rate is calculated as: [1 – (world emission reduction in $GtCO_2eq$)/(Annex I emission reduction objective in $GtCO_2eq$)]. It is expressed in per cent. When the emission reduction achieved at the world level (in $GtCO_2eq$) is equal to the emission reduction objective set by Annex I (in $GTCO_2eq$), there is no leakage overall, and the leakage rate is 0.

the carbon price in regions that experience higher leakage before linking, then leakage towards uncapped countries is reduced. Conversely, if linking raises the carbon price in regions that initially face low leakage rates, then leakage towards uncapped countries is increased. In the A1 act scenario, model simulations point to a small overall reduction in leakage from linking among Annex I regions (Figure 3, Panel B).

2.1.3. *Other potential benefits from direct linking*

In particular, linking schemes can improve cost effectiveness by increasing the size and liquidity of carbon markets. In the A1 act scenario presented above where Annex I ETSs are linked, the size of the carbon market is projected to reach 2.5% of Annex I GDP in 2020. A larger market size tends to dampen the impact of unanticipated shocks, thereby lowering overall carbon price volatility and enhancing incentives for firms to make emission reduction investments.[7] Furthermore, transaction costs are expected to be smaller in a larger, more liquid market, especially if some regional schemes are too small to foster the development of institutions for reducing such costs. Larger market size also reduces problems that may arise if some sellers or buyers have market power (Hahn, 1984). Finally, market liquidity can lower the cost of insuring against carbon price uncertainty by fostering the development of derivative markets.

Compared with a global emissions trading system, it has also been argued that a linked system of regional ETSs may be an easier way to reflect the principle of "common but differentiated responsibilities and respective capabilities" across regions, and thereby to extend participation to developing countries (Jaffe and Stavins, 2007). Permit allocation rules make it possible to differentiate across regional commitments and costs under a top-down approach. However, such differentiation can also be achieved through regions' own assessment of their responsibilities and reflecting their specific national circumstances – as revealed *de facto* by their target choice – under a bottom-up approach. The gains from linking Annex I ETSs to potential non-Annex I country ETSs would be larger than those achieved through linking within Annex I only, if the heterogeneity in pre-linking carbon prices (and hence in commitments or actions) between Annex I and non-Annex I is higher than within Annex I.

2.2. *Indirect linking of carbon markets through a common crediting mechanism*

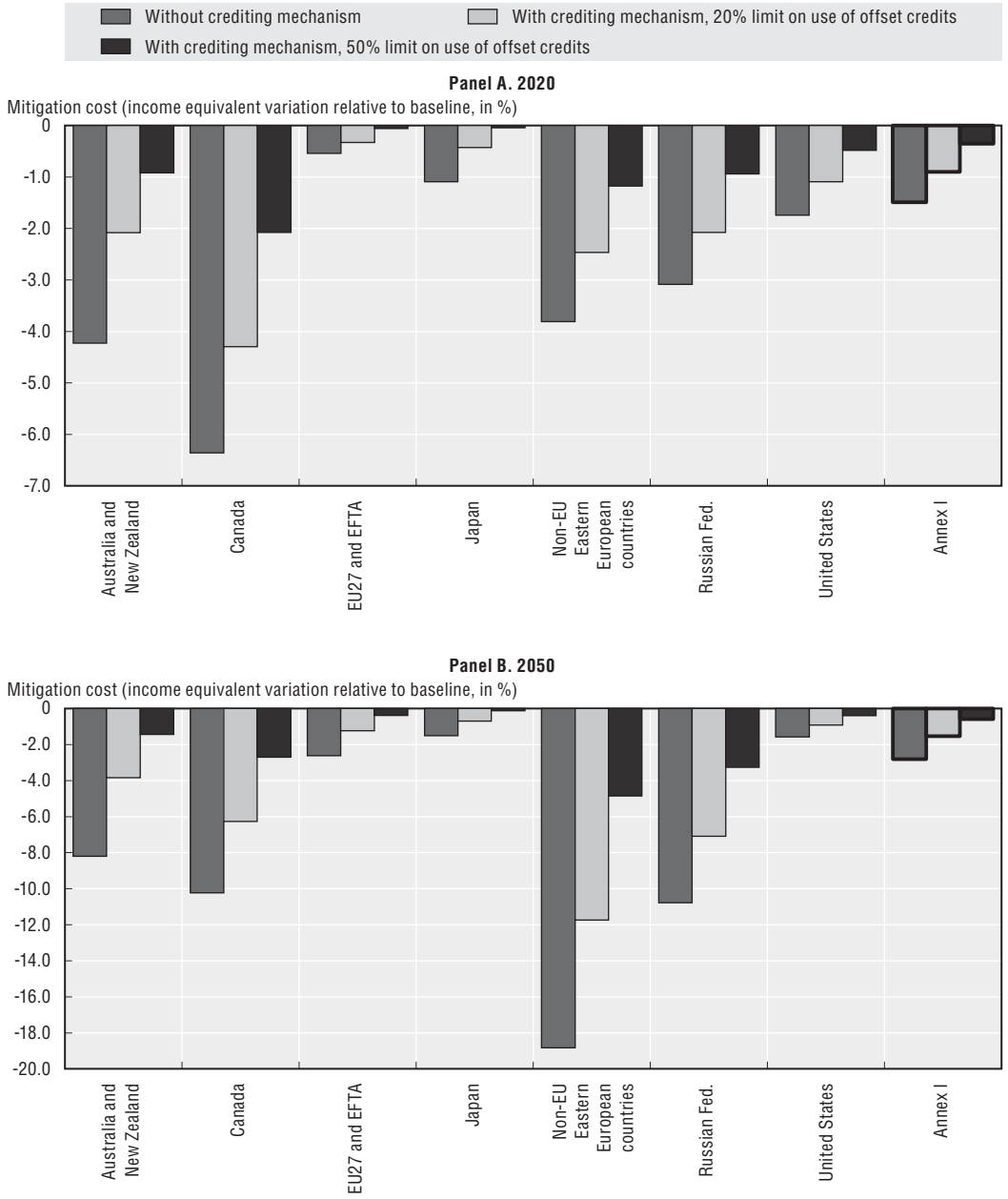
2.2.1. *Improving cost-effectiveness*

Linking can also occur "indirectly" when at least two different ETSs allow part of their emission reductions to be achieved in other countries, as can happen for example through a common crediting mechanism such asthe Clean Development Mechanism (CDM), which is one of the flexibility mechanisms of the Kyoto Protocol.

The CDM allows emission-reduction projects in non-Annex I countries – i.e. developing countries, which have no GHG emission constraints – to earn certified emission reduction (CER) credits, each equivalent to one tonne of CO_2eq. These CERs can be purchased and used by Annex I countries to meet part of their emission-reduction commitments. In principle, assuming that developing country emitters do not take on binding emission commitments in the near future, well-functioning crediting mechanisms could improve the cost-effectiveness of GHG mitigation policies in developed countries, both directly and indirectly through partial linking of their ETS as well as reduce carbon leakage and competitiveness concerns by lowering the carbon price in developed countries.[8]

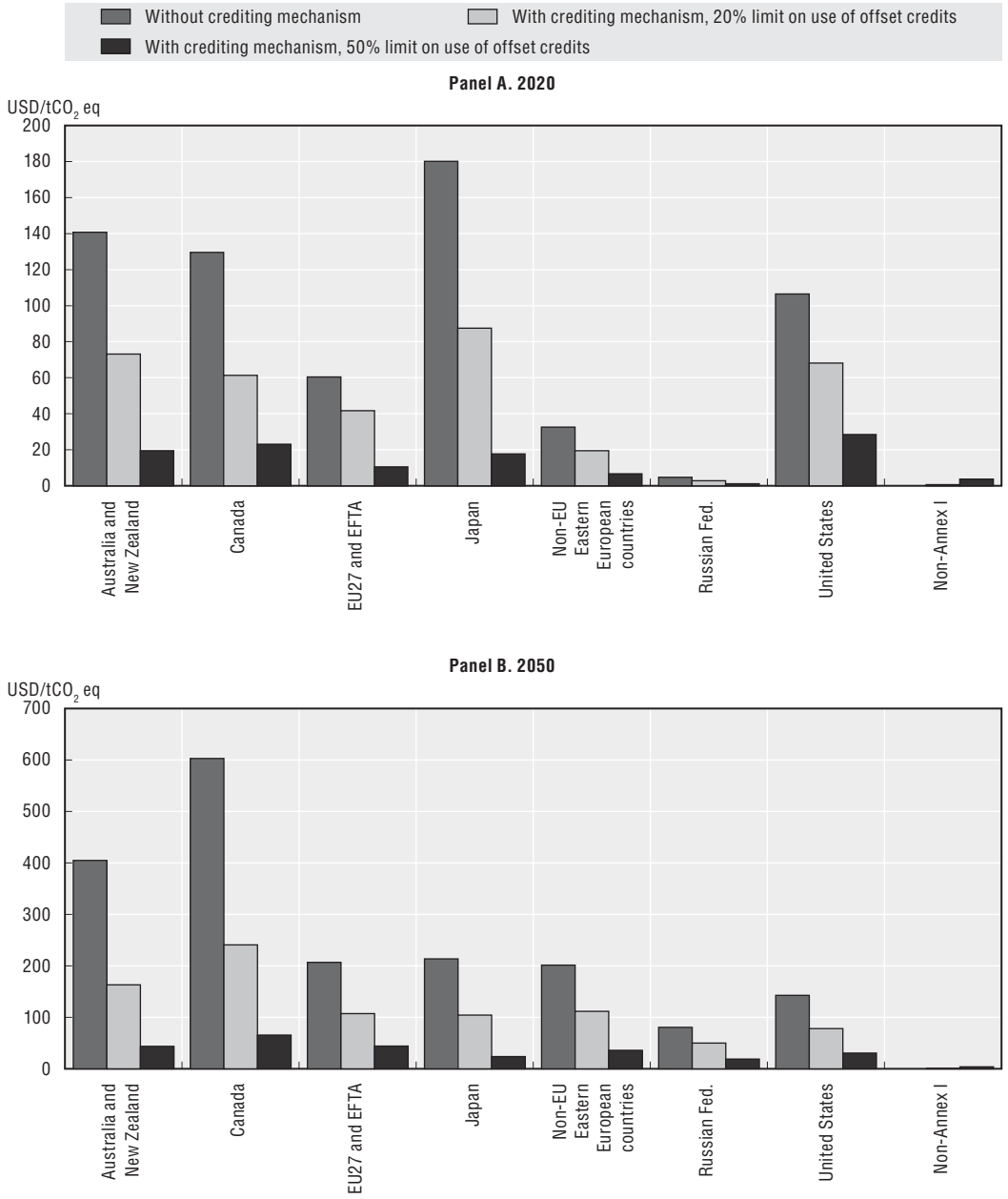
Well-functioning crediting mechanisms appear to have very large potential for saving costs, reflecting the vast low-cost abatement potential existing in a number of emerging and developing countries, particularly China. Compared with the scenario A1 act without direct linking between Annex I countries, allowing Annex I countries to meet 20% of their unilateral commitments through reductions in non-Annex I countries is estimated to reduce their mitigation costs by roughly 40% (Figure 4).

Figure 4. **Impact of allowing access to a well-functioning crediting mechanism on mitigation policy costs in each Annex I region under the A1 act scenario**



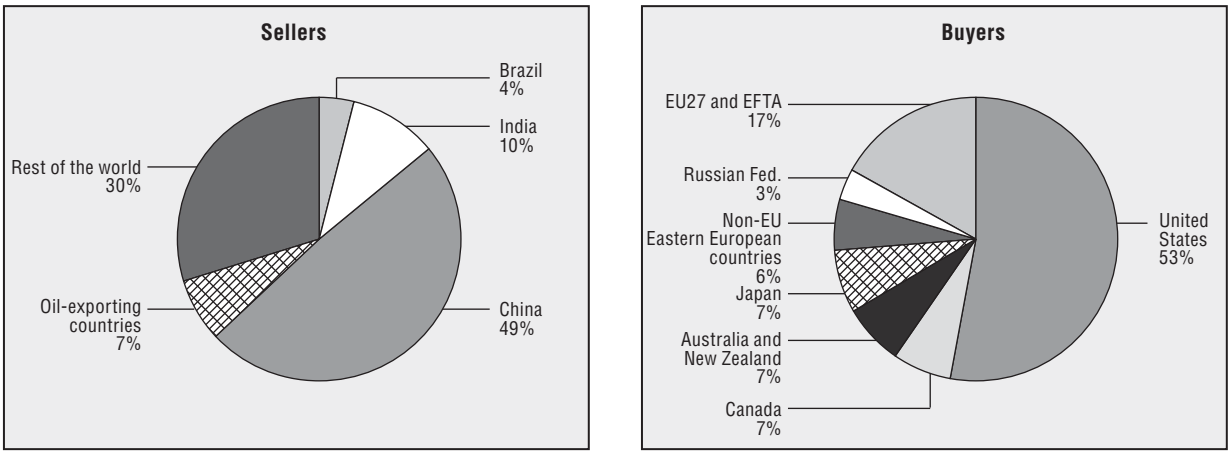
Raising the cap on offsets allowed from 20% to 50% would bring substantial further benefits. Cost savings are found to be largest for those Annex I regions that would otherwise face the highest marginal abatement costs – and, therefore, the highest carbon price levels (Figure 5) – and/or are most carbon intensive. Australia-New Zealand and Canada fall into both categories, while Russia falls into the latter. Non-Annex I regions would enjoy a slight income gain from exploiting cheap abatement opportunities and selling them profitably in the form of offset credits. In this illustrative scenario, China would be by far the largest seller and the United States the largest buyer in the offset credit market, accounting for about half of worldwide sales and purchases by 2020, respectively (Figure 6).

Figure 5. **Impact of a well-functioning crediting mechanism on carbon prices under the A1 act scenario**



The offset schemes also imply a substantial shift in the size of the different carbon markets. On the domestic markets of the Annex I countries, the revenues from carbon pricing are lower, as the increased amount of domestic emissions is more than outweighed by the reduction in the permit price, especially when the amount of offsets is limited to 20% (cf. Figure 5). Furthermore, there is less scope for trading among the Annex I countries (see below). Finally, sizable international transfers are involved in the purchase of the CER credits. As the ENV-Linkages model projects very low CER prices (due to the assumption of a well-functioning offset market), these flows do not represent a major percentage of GDP. Dellink et al. (2011) investigate the size of the carbon markets in more detail, using the same model

Figure 6. **Geographical distribution of offset credit buyers and sellers by 2020 under the A1 act scenario**



Note: Percentages apply to a 20% limit on the use of offsets, but shares are rather robust for this limit.

but focusing on the specific policies associated with the pledges made in the annexes of the Copenhagen Accord.

One caveat to these cost-saving and trade-flow estimates is that they assume a crediting mechanism with no transaction costs and no uncertainty on delivery, as is apparent from the very low projected offset prices in these simulations (Figure 5). In practice, there are numerous market imperfections and policy distortions which may prevent some of the non-Annex I abatement potential from being fully reaped. These include transaction costs and bottlenecks, information barriers, credit-market constraints, and institutional and regulatory barriers to investment in host countries.[9] The well-functioning crediting mechanism that is modelled here is largely equivalent to an international (asymmetric) ETS covering all non-Annex I countries, in which each of them is assigned a target equal to their baseline emissions. Furthermore, it is sometimes argued that purchase of CER credits may not lead to a fully proportional emission reduction in the host country, for instance due to additionality concerns. The simulations assume full integrity of the credits (cf. endnote 2); see Section 4 for a more detailed discussion.

2.2.2. Ensuring convergence in carbon prices

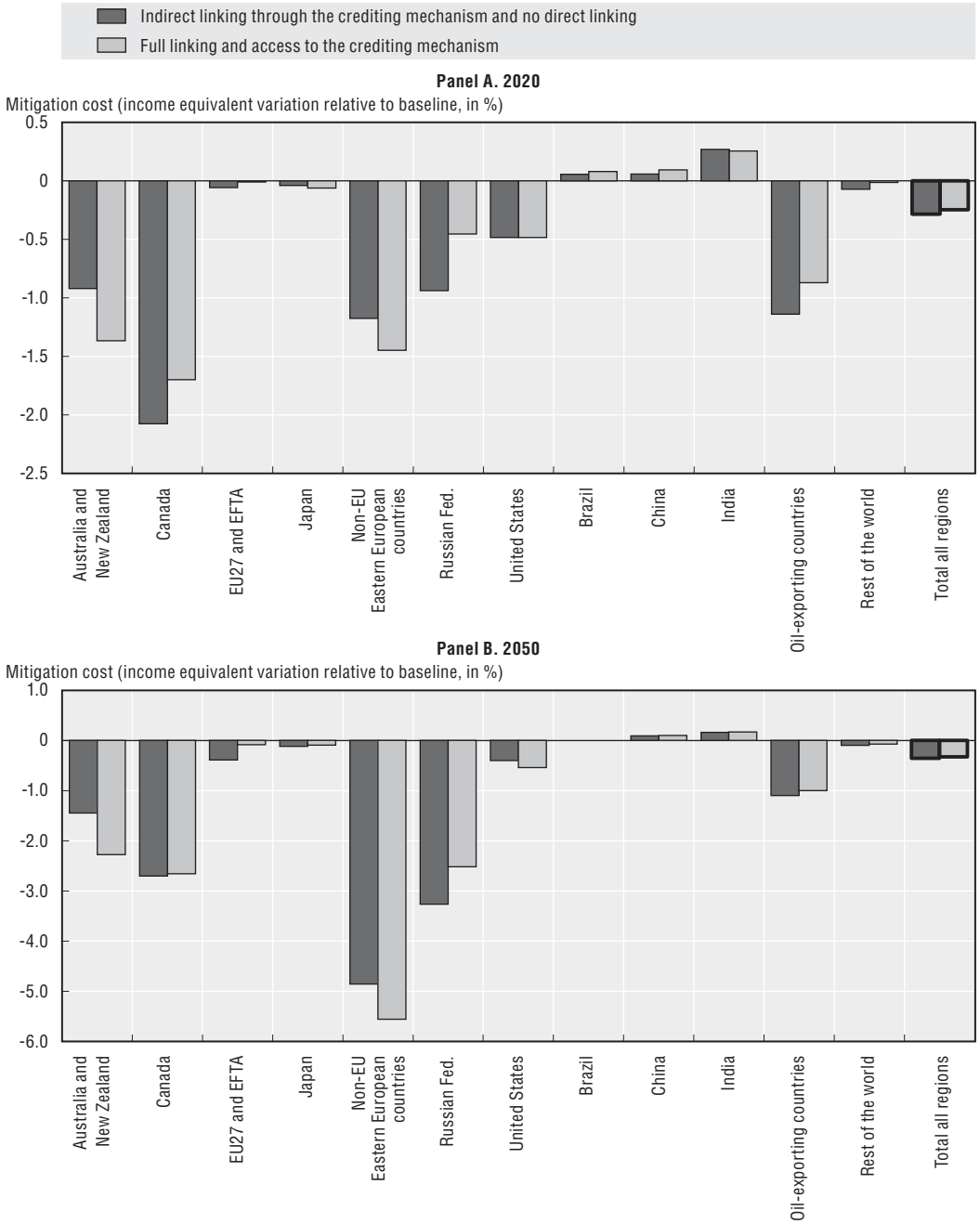
Crediting mechanisms indirectly link the ETSs of countries covered by binding emission caps if credits from a single mechanism (e.g. the CDM) are accepted in several different ETSs. Indeed, they result in partial convergence of carbon prices and marginal abatement costs across the different ETSs, which improve their cost effectiveness as a whole.

As Figure 5 clearly shows, the variance in carbon prices across Annex I regions is found to decline dramatically as the cap on the use of offsets is relaxed, becoming fairly small for instance under a 50% cap. As a result, once schemes are indirectly linked through crediting mechanisms, the additional gains from direct linking are much smaller than discussed in Section 2.1.[10]

As a matter of fact, the more lax are constraints on the use of credits, the stronger the indirect linkage between systems and the smaller the additional gains from explicit linking will be. For instance, ENV-Linkages simulations suggest that if Annex I regions are allowed to meet up to 50% of their domestic commitments through the use of offsets, the overall

additional gain from direct linking would be close to zero, although some countries would still benefit significantly (Figure 7). Full linking between ETSs implies that the 50% cap on offsets applies to the Annex I region as a whole rather than to each country individually, and is allocated across countries in a cost-effective manner. Under this set up, the region Australia-New Zealand is estimated to lose because it benefits from a smaller amount of offsets.

Figure 7. **Estimated gains from direct linking across Annex I ETSs that are already linked through a crediting mechanism**



Note: Full linking implies linking between all Annex I ETSs plus indirect linking through the crediting mechanism.

3. Alternative set ups for linking ETSs

3.1. Alternative baselines for CDM credits

The choice of a project baseline against which certified emission rights (CERs) are granted does not only have an impact on the volume of credits generated, but also matters for carbon leakage. This is because an emissions baseline established *before* the project is implemented can vary, depending on the assumptions made about policies and projects in other sectors and regions, and their effect on output and emissions within the project boundary. Three approaches can be identified in setting a baseline:

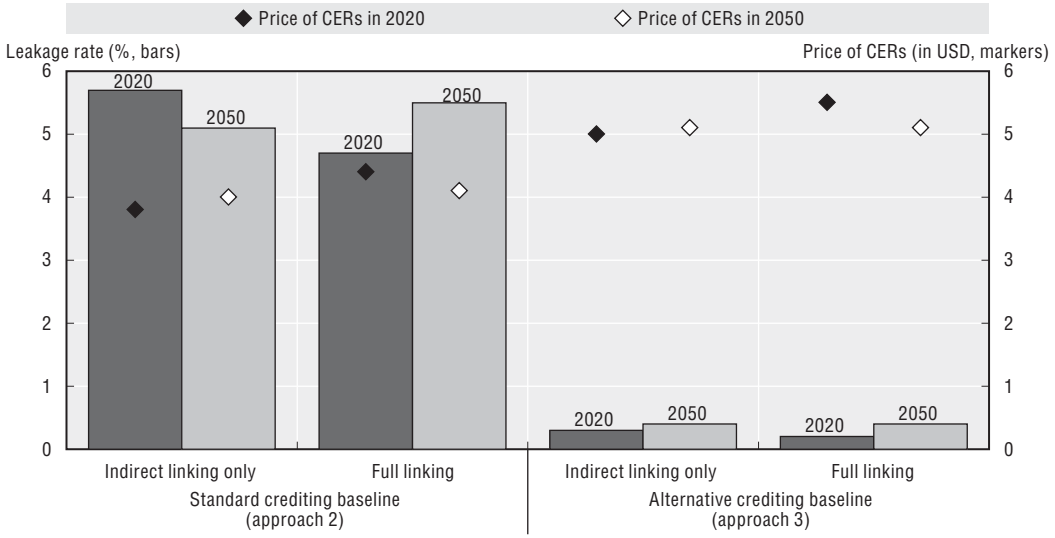
1. *Accounting* for the impact of all other CDM projects on the projects expected emissions. If these other projects lower the international carbon price and thus reduce leakage from countries covered by binding emission caps within the project boundary, they should lower the project's emission baseline. This would be the "theoretically correct" baseline under current UNFCCC guidelines. However, implementing this approach is complex and costly, and would likely remain so even under a scaled-up CDM.
2. *Excluding* the impact of all other projects on the project's expected emissions. Because of the complexity and cost of the first approach, the CDM Executive Board currently takes as the baseline the project's emission level under a scenario where some countries – currently most of the Annex I countries – have emission commitments while the rest of the world does not. This therefore does not account for the effect of other CDM projects on the output from, and therefore credit generated by, any particular project. Implicitly, this assumes that all individual CDM projects have a marginal effect on the world economy. This is the approach that is adopted in the base simulations in Section 2.
3. *Setting* the baseline as the BAU emission level in a hypothetical "*no world action*" scenario where no country has binding emission commitments. In this case, CDM projects would receive fewer credits than under Approach 1, as they would be required to more than offset any leakage within the project boundary resulting from binding emission caps in other sectors and regions. However, this approach would *ceteris paribus* imply either a lower credit volume than Approach 2, or more domestic action in host countries. This alternative addresses some of the concerns with current practice whereby "market leakage" is not taken into account in CDM baselines (cf. Vöhringer et al., 2006), and this section analyses to what extent it will affect the model results.

Under all three approaches, Kallbeken (2007) finds that the CDM lowers the carbon-price differential between countries that face binding emissions caps and other countries, and thereby reduces leakage (all other factors being equal). However, these leakage reductions are typically smaller under Approach 2 as used in Section 2. This is because Approach 2 does not account for the fact that implementing all *other* CDM projects together reduces international carbon prices, leakage and thereby the projected emissions of any other project considered.

Figure 8 shows how the alternative definition of the baseline for assigning credits using Approach 3 could virtually eliminate the leakage from the Annex I mitigation actions. The figure shows both the cases of indirect linking only (i.e. there is no direct linking of ETSs in Annex I countries) and full linking (i.e. both indirect linking through the crediting mechanism and direct linking of Annex I ETSs).

In our set-up of the model scenarios, the market for CERs is entirely demand driven: as CERs are cheaper than domestic reductions, but there is a binding limit (in this case 50% of Annex I emission reductions) on their use, the size of the CDM market is not affected by

Figure 8. **Impact of an alternative baseline for the crediting mechanism on leakage and the price of CERs**



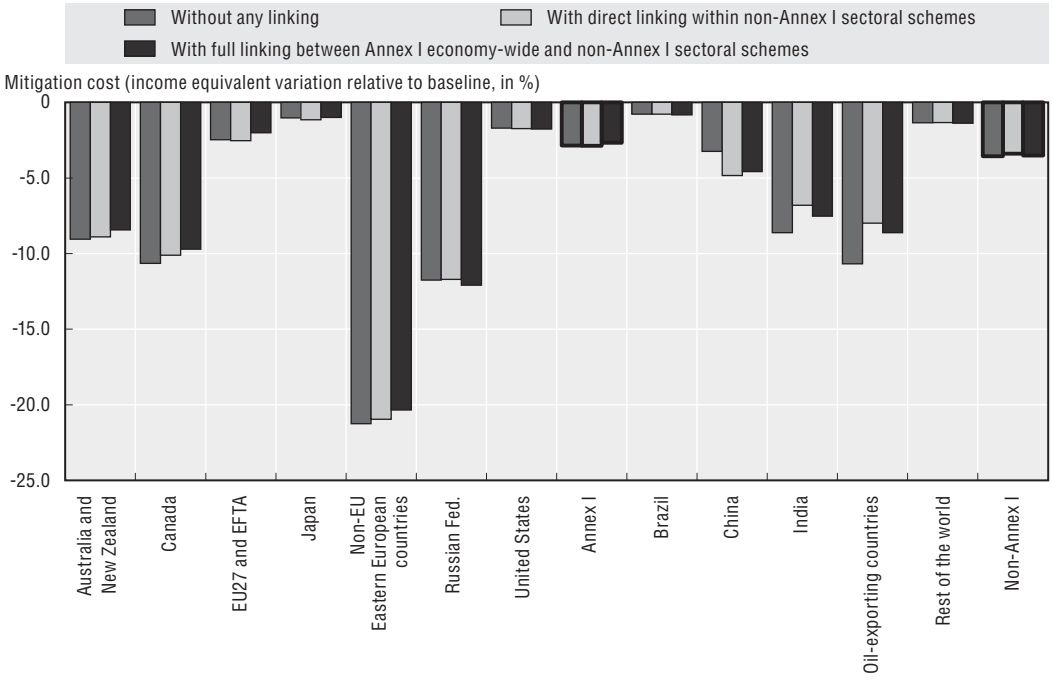
the alternative baseline assumption. Rather, the non-Annex I regions will undertake the additional domestic action to compensate for the leakage that may occur, as required to earn the maximum volume of credits. Obviously, this comes at a cost for the non-Annex I regions, and the equilibrium price of CERs increases by some 25% (Figure 8, right axis).

3.2. Alternative policies: Binding sectoral caps in non-Annex I countries

In the analysis above, mitigation actions outside the Annex I countries are based on a crediting mechanism, as no caps are placed on emissions in developing countries. In the policy arena, sectoral approaches have also been put forward as a prominent means to address competitiveness issues. To investigate how sectoral caps and crediting mechanisms compare, we investigate a binding sectoral cap on both energy intensive industries (EIIs) and power sectors in non-Annex I countries, equal to a 20% reduction from 2005 levels by 2050, and consider different scenarios regarding the type of linking between regional ETSs (no linking, linking between the sectoral schemes covering the EII industries in non-Annex I countries, and full linking to the Annex I ETS). In all cases, Annex I countries implement together the A1 act scenario through a single ETS covering their joint emissions, as discussed in Section 2.

A binding sectoral cap covering EIIs and the power sector in non-Annex I countries could substantially reduce emissions worldwide. Owing to the fast emission growth expected in non-Annex I countries, a 20% emission cut in these sectors in non-Annex I countries (from 2005 levels) would achieve a larger absolute reduction in world emissions than a 50% cut in Annex I countries by 2050. The costs associated with binding sectoral caps would vary across non-Annex I countries, as shown in Figure 9. They would depend on how stringent the target is relative to BAU, the availability of cheap abatement options (the shape of the marginal abatement cost curve), the carbon intensity of output, and whether international permit trading between EIIs in non-Annex I countries is allowed. In the illustrative scenario considered here, India is found to incur larger mitigation costs than China, mainly due to its faster projected BAU emission growth, which in turn reflects in part faster population growth. However, that gap would be reduced substantially if

Figure 9. **Mitigation costs under an international ETS in Annex I and binding sectoral caps in non-Annex I regions**



Note: All scenarios combine a 50% emission cut in Annex I (relative to 1990 levels) and a 20% cut in EIIs and the power sector in non-Annex I (relative to 2005 levels) by 2050.

international permit trading (internal linking) was allowed across non-Annex I regions. While non-Annex I countries face a smaller emission reduction relative to BAU than Annex I countries (–25% vs. –30% by 2020 and –40% vs. –60% by 2050 for total emissions) and benefit from their larger potential to reduce emissions more cheaply, non-Annex I countries would incur larger costs (more than 3% of their joint income in 2020, compared with less than 1.5% for Annex I countries), reflecting their higher carbon intensity, particularly by 2020, and the concentration of mitigation efforts in EIIs only.

Linking sectoral ETSs in non-Annex I countries to economy-wide ETSs in Annex I countries could also generate aggregate economic gains by exploiting the heterogeneity of (marginal) abatement costs between the two areas. As Figure 9 shows, these additional gains are limited, and mostly concentrated in the Annex I countries. Note also that such linking could have significant redistributive effects across countries. Therefore, allocation rules may need to be adjusted upon linking to ensure that the gains from linking are shared widely across participating countries.

4. Practical issues raised by linking

The benefits of linking ETSs are of course conditional on the success of establishing a well-functioning domestic ETS in the participating countries. ETSs are by no means infallible. Experiences with e.g. the EU ETS have shown that it may be difficult to set the permit volumes at the level that generates a desired market outcome, and that the predictability and stability of the schemes may suffer in case of large volatility, e.g. caused by economic shocks. Nonetheless, the large potential benefits of well-functioning ETSs are clearly illustrated in this paper. Furthermore, while linking ETSs either directly or indirectly

can be an important step towards the emergence of a single international carbon price, linking ETSs that are already domestically established raises a number of difficulties and risks in practice that will need to be addressed. These concerns cannot in general be analysed within the context of the model.

One of the major risks associated with a linked system of several independent and heterogeneous ETSs and crediting mechanisms is that its overall environmental performance could be weak. This is partly because the region with the lower carbon price *ex ante* has an incentive to relax its cap in order to generate additional revenue from exporting allowances – and a larger gain from linking more broadly once systems are linked (Helm, 2003; Rehdanz and Tol, 2005). In order to alleviate this, the region with the higher carbon price may also relax its target, thereby triggering a "race to the bottom". Another source of environmental concern is that linking would automatically lead to the spreading across regions of provisions to contain the cost of mitigation (cost-containment measures), such as carbon price caps ("safety valves") and the use of crediting mechanisms (see OECD, 2009; Ellis and Tirpak, 2006; Jaffe and Stavins, 2007; and Flachsland et al. 2009). For example, the use of the safety valve implies that the overall target is relaxed once the price cap is reached, while in the absence of linking the scheme without any safety valve would retain control over its own target.[11] Likewise, linking to an offset credit system whose environmental integrity is weaker than that of an ETS could raise environmental concerns in countries that have more restrictive policies regarding the use and quality of offsets. Furthermore, linking ETSs also implies that volatility on a carbon market in one particular country or region will spread automatically to the other ETSs. To some extent, this can mitigate the negative consequences of the unanticipated shock that caused the volatility in the region of origin, but it also makes the other regions more vulnerable to such a shock.

Linking would automatically lead to the spreading of a number of other design features specific to one particular scheme, such as provisions for credits to be banked or borrowed from future commitment periods. As a result, governments in the linked regions would lose control over several features of their existing ETS. The impact of linking an ETS with intensity targets to an ETS with absolute targets depends on the permit allocation rules. If the cap on emissions in the system with the intensity target is set *ex ante* on the basis of projected GDP growth, then that scheme is *de facto* equivalent to an ETS with an absolute cap, and linking does not affect overall emissions. By contrast, if the permit authority of the intensity target scheme regularly adjusts the supply of permits in order to meet its intensity target, overall emissions will fluctuate. Overall environmental performance does not have to be undermined if emissions merely fluctuate around the level that would prevail under absolute caps, but it could be affected if GDP growth is higher than anticipated, or if the intensity target system creates an incentive to increase production and emissions in order to obtain additional credits, as could be the case under firm or sector-level (as opposed to economy-wide) intensity targets.

Against this background, in order to facilitate future linking, participating governments should seek to agree on their targets and the ETS design features to be harmonised prior to linking, including cost containment measures, decisions to link to another system, and how to co-ordinate monitoring, reporting and verification efforts (OECD, 2009; Haites and Wang, 2006). However, this has not happened so far in practice. Centralised institutions that support implementation of the UNFCCC, the Kyoto Protocol, and any future protocol could help by providing an international framework to discuss issues of linking national and regional ETSs. The recent decision to link the Australian ETS to the EU ETS is an important step in this process and can provide valuable information on the opportunities and pitfalls involved.

Although linking ETSs tends in general to lower the mitigation cost of each of the participating regions, it also affects the distribution of costs within schemes. Within regions where linking leads to a carbon price increase, permit sellers gain while permit buyers lose, (and *vice versa* within regions where linking results in a carbon price decline). However, these distributional effects are similar in nature to those of international trade, and as such they do not provide a compelling argument for limiting linking. Some of the associated political-economy problems can be reduced through permit allocation rules if necessary, for instance through allowing transitory grandfathering[12] upon linking in regions with lower pre-linking carbon price.

In its current form, the CDM raises a number of concerns that are comparatively greater than those arising from direct linking and which, if not addressed, will undermine its ability to deliver the expected benefits. The so-called additionality criterion, under which only emission reductions that can be attributed to the mitigation project give rise to carbon credits (technically called Certified Emission Reductions, or CERs), is key to ensuring the environmental integrity of the CDM. Otherwise, CERs would amount to a mere income transfer to recipient countries without reducing net global GHG emissions more than would have happened without the mechanism. In practice, it has been argued that a large share of CDM projects do not bring about actual reductions in emissions (ECCP, 2007; Schneider, 2007; Wara and Victor, 2008). One way for different systems to protect their integrity *vis-à-vis* imported credits could be to use discount factors when allowing permits from linked schemes or CER credits into their system (Kollmuss et al., 2010; Schneider, 2009). The CDM can also create perverse incentives to raise initial investment and output in carbon-intensive equipment, so as to get emission credits for reducing emissions later, depending on expectations about how future baselines will be set. Another incentive problem is that the large financial inflows from which developing countries may benefit under a future CDM could undermine their willingness to take on binding emission commitments at a later stage (OECD, 2009).

A number of other risks stem from the development of increasingly large carbon markets as more countries undertake mitigation actions. For instance, the size of carbon markets is estimated to reach 2.3% of GDP in Annex I countries by 2050 under the linked A1 act scenario,[13] and 5% of world GDP under a global ETS that stabilises overall GHG concentration below 550 ppm CO_2eq.[14] Three major risks can be identified that would have to be addressed:

- *Lack of market liquidity:* Liquid primary markets foster the emergence of derivative instruments that would lower the cost for firms to insure against future carbon price uncertainty. Liquid markets would also reduce the opportunities for market manipulation. Market liquidity could be enhanced through the regular spot sales of short-term permits, allowing banking and ensuring credible commitments on future mitigation policies.
- *Risk associated with the development of derivative markets:* This risk will be partly addressed by identifying and certifying financial-market authorities responsible for carbon markets.
- *A counterparty risk* that could lead to market dysfunction as a large share of current trading is conducted through bilateral over-the-counter negotiations between participants. This risk can be mitigated by extending the access to clearinghouses and/or introducing penalties for performance failure in contracts.

5. Concluding remarks

Both direct and indirect linking of emissions trading schemes (ETSs) can help to reduce the cost of international climate mitigation action. In the long run, it is essential to achieve ambitious global emission reductions at low cost, and this paper has provided evidence that linked ETSs can play a pivotal role in this regard. Especially the opportunity to tap into the least-cost mitigation measures around the globe can reduce the cost of climate action. However, various design issues will have to be addressed for direct and indirect linking to deliver their full benefits. If these can be overcome, implementing ETSs and allowing international trading of the associated permits, i.e. linking domestic carbon markets, can achieve the dual goal of increasing the environmental ambition and the cost-effectiveness of international mitigation action. This is essential for a successful international climate policy framework in the coming years.

Notes

1. Annex I regions are those countries that have agreed to reduce their greenhouse gas emissions under the Kyoto Protocol. Although the United States never ratified the Kyoto Protocol and Canada recently withdrew, we include them in our analysis as Annex I countries.
2. Note that in the simulations, differences in integrity of the permits from different trading schemes are ignored: permits are assumed to be fully substitutable once the schemes are linked. This assumption is discussed in Section 4.
3. One-way linking (when system A recognises system B's allowances but the latter does not) ensures that the price in system A never exceeds the price in system B, and hence, would only limit competitiveness concerns for firms belonging to system A. However, under one-way credits to firms in system linking, firms in system A would be penalised by not being allowed to sell to B.
4. Sensitivity analysis confirms these small gains from direct linking, at least for relatively realistic distributions of commitments among Annex I countries.
5. Note that the analysis excludes the possibility for countries to renegotiate the emission-reduction targets when linking. Adding this option would complicate the analysis and cloud the insights on the benefits of linking existing systems.
6. Energy-intensive industries in this study include ferrous metal, chemicals, mineral products, pulp and paper and non-ferrous metals.
7. Carbon price volatility may still increase in one of the two schemes if the other is subject to larger and/or more frequent shocks, and is large enough to have significant influence on the overall carbon price after linking.
8. Additionally, indirect linking can boost clean technology transfers to developing countries and facilitate the implementation of explicit carbon-pricing policies in developing countries at a later stage by putting an opportunity cost on their GHG emissions. These elements are, however, not captured in our model simulations.
9. For instance, under a USD 20 carbon price in Annex I countries, Bakker et al. (2007) tentatively estimate that the amount of emissions abated through crediting projects in non-Annex I countries might be reduced by a factor of up to two if these barriers were taken into account.
10. Equivalently, having already linked carbon markets across the Annex I region will lower the cost reductions from indirect linking.
11. Partly for these reasons, the EU directive on linkage currently forbids linking the EU ETS to a scheme featuring a safety valve.
12. Grandfathering consists of allocating permits for free on the basis of historical emissions.
13. In the case with direct linking, this is slightly less (2.1% of Annex I GDP), as carbon prices are lower.
14. By comparison, for instance, in 2007 the US sub-prime mortgage market (total outstanding amount of sub-prime loans) amounted to about 9.5% of US GDP, or about 3% of world GDP at current exchange rates (OECD, 2007).

References

Anger, N. (2008), "Emissions Trading Beyond Europe: Linking Schemes in a Post-Kyoto World", *Energy Economics,* 30, pp. 2028-2049.

Bakker, A. et al. (2007), "Carbon Credit Supply Potential Beyond 2012: A Bottom-Up Assessment of Mitigation Options", *Energy Research Centre of the Netherlands Report*, ECN-E–07-090.

Burniaux, J.M., G. Nicoletti and J. Oliveira Martins (1992), "GREEN: A Global Model for Quantifying the Costs of Policies to Curb CO_2 Emissions", *OECD Economic Studies*, No. 19 (Winter).

Burniaux, J.M., J. Chateau and R. Duval (2013), "Is there a Case for Carbon-Based Border Tax Adjustment? An Applied General Equilibrium Analysis", *Applied Economics*, Vol. 45, No. 16.

Dellink, R.B., G. Briner and C. Clapp (2011), "The Copenhagen Accord/Cancun Agreements Emissions Pledges for 2020: Exploring Economic and Environmental Impacts", *Climate Change Economics,* Vol. 2, No. 1.

Dimaranan, B.V. (ed.) (2006), *Global Trade, Assistance, and Production: The GTAP 6 Database*, Center for Global Trade Analysis, Purdue University.

Duval, R. and C. de la Maisonneuve (2010), "Long-Run Growth Scenarios for the World Economy", *Journal of Policy Modeling,* Vol. 32, No. 1.

ECCP (2007), *European Climate Change Platform Meeting Report*, October.

Ellis, J. and D. Tirpak (2006), "Linking GHG Emission Trading Systems and Markets," OECD/IEA, *www.oecd.org/dataoecd/45/35/37672298.pdf*.

Flachsland, C., R. Marschinski and O. Edenhofer (2009), "Global Trading versus Linking: Architectures for International Emissions Trading", *Energy Policy,* 37, pp. 1637-1647.

Grull, G. and L. Taschini (2012), "Linking Emission Trading Schemes: A Short Note", *Economics of Energy and Environmental Policy*, Vol. 1, No 3.

Haites, E. and X. Wang (2006), "Ensuring the Environmental Effectiveness of Linked Emissions Trading Schemes", *Margaree Consultants Paper*, 2006-5.

Helm, C. (2003), "International Emissions Trading with Endogenous Allowance Choices", *Journal of Public Economics,* Vol. 87.

Howe, H. (1975), "Development of the Extended Linear Expenditure System from Simple Saving Assumptions", *European Economic Review,* 6, pp. 305-310.

Lluch, C. (1973), "The Extended Linear Expenditure System", *European Economic Review,* Vol. 4, pp. 21-32.

IEA (2008), *World Energy Outlook 2008*, Paris.

Jaffe, J. and R. Stavins (2007), "Linking Tradable Permit Systems for Greenhouse Gas Emissions: Opportunities, Implications, and Challenges", *IETA Report.*

Jaffe, J. and R. Stavins (2008), *Linkage of Tradable Permit Systems in International Climate*, FEEM Nota di Lavoro, 90.2008, Milan.

Kallbekken, S. (2007), "Why the CDM will Reduce Carbon Leakage", *Climate Policy,* Vol. 7, No. 3.

Kollmuss A., M. Lazarus and G. Smith (2010), "Discounting Offsets: Issues and Options", *Stockholm Environment Institute Working Paper,* WP-US-1005.

Mehling, M., A. Tuerk and W. Sterk (2011), "Prospects for a Transatlantic Carbon Market: What Next after the US Midterm Elections?", *Climate Strategies*, *www.climatestrategies.org*.

OECD (2004), *Sources of Economic Growth*, OECD Publishing, Paris.

OECD (2007), *Financial Market Trends*, No. 94, OECD Publishing, Paris.

OECD (2009), *The Economics of Climate Change Mitigation: Policies and Options for Global Action beyond 2012*, OECD Publishing, Paris.

OECD (2012), *Environmental Outlook to 2050*, OECD Publishing, Paris.

Rehdanz, K. and R.S.J. Tol (2005), "Unilateral Regulation of Bilateral Trade in Greenhouse Gas Emission Permits", *Ecological Economics*, Vol. 54, pp. 397-416.

Schneider, L. (2007), "Is the CDM Fulfilling its Environmental and Sustainable Development Objectives? An Evaluation of the CDM and Options for Improvement", *Report* prepared for WWF, Öko-Institut, Berlin.

Schneider, L. (2009), "A Clean Development Mechanism with Atmospheric Benefits for a Post-2012 Climate Regime", *International Environmental Agreements*, Vol. 9.

Tuerk, A., M. Mehling, C. Flachsland and W. Sterk (2009), "Linking Carbon Markets: Concepts, Case Studies and Pathways", *Climate Policy,* 9, pp. 341-357.

Vöhringer, F., T. Kuosmanen and R. Dellink (2006), "How to Attribute Market Leakage to CDM Projects", *Climate Policy,* Vol. 5, No. 5.

Wara, W. and D. Victor (2008), "A Realistic Policy on International Carbon Offsets", *Working Paper*, No.°74, Stanford University.

ANNEX A

Key characteristics of the ENV-Linkages model

Table A1. **ENV-Linkages model sectors**

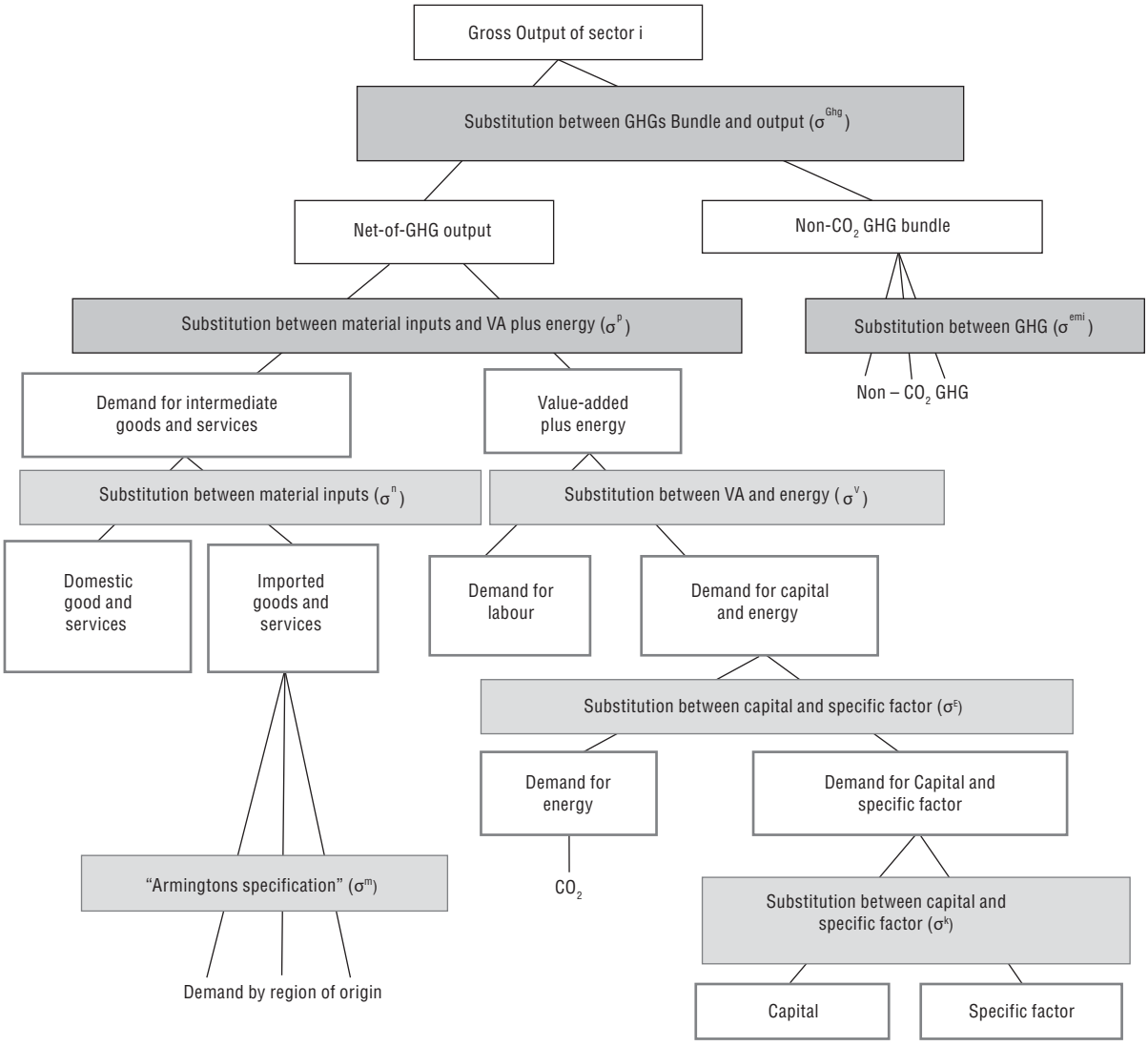
1	Rice	14	Food products
2	Other crops	15	Other mining
3	Livestock	16	Non-ferrous metals
4	Forestry	17	Iron and steel
5	Fisheries	18	Chemicals
6	Crude oil	19	Fabricated metal products
7	Gas extraction and distribution	20	Paper and paper products
8	Fossil fuel based electricity	21	Non-metallic minerals
9	Hydro and geothermal electricity	22	Other manufacturing
10	Nuclear power	23	Transport services
11	Solar and wind electricity	24	Services
12	Renewable combustibles and waste electricity	25	Construction and dwellings
13	Petroleum and coal products	26	Coal

Table A2. **ENV-Linkages model regions**

	ENV-Linkages regions	GTAP countries/regions
1	Australia-New Zealand	Australia, New Zealand
2	Japan	Japan
3	Canada	Canada
4	United States	United States
5	European Union 27 and EFTA	Austria, Belgium, Denmark, Finland, Greece, Ireland, Luxembourg, Netherlands, Portugal, Sweden, France, Germany, United Kingdom, Italy, Spain, Switzerland, Rest of EFTA, Czech Republic, Slovakia, Hungary, Poland, Romania, Bulgaria, Cyprus,[1, 2] Malta, Slovenia, Estonia, Latvia, Lithuania
6	Brazil	Brazil
7	China	China, Hong Kong
8	India	India
9	Russia	Russian Federation
10	Oil-exporting countries	Indonesia, Venezuela, Rest of Middle East, Islamic Republic of Iran, Rest of North Africa, Nigeria
11	Non-EU Eastern European countries	Croatia, Rest of Former Soviet Union
12	Rest of the world	Korea, Taiwan, Malaysia, Philippines, Singapore, Thailand, Viet Nam, Rest of East Asia, Rest of Southeast Asia, Cambodia, Rest of Oceania, Bangladesh, Sri Lanka, Rest of South Asia, Pakistan, Mexico, Rest of North America, Central America, Rest of Free Trade Area of Americas, Rest of the Caribbean, Colombia, Peru, Bolivia, Ecuador, Argentina, Chile, Uruguay, Rest of South America, Paraguay, Turkey, Rest of Europe, Albania, Morocco, Tunisia, Egypt, Botswana, Rest of South African Customs Union, Malawi, Mozambique, Tanzania, Zambia, Zimbabwe, Rest of Southern African Development Community, Mauritius, Madagascar, Uganda, Rest of Sub-Saharan Africa, Senegal, South Africa

1. *Note by Turkey:* The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue".
2. *Note by all the European Union member states of the OECD and the European Union:* The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

Figure A1. **Structure of production in ENV-Linkages**



OECD Journal: Economic Studies
Volume 2013
© OECD 2014

Using a quasi-natural experiment to identify the effects of birth-related leave policies on subjective well-being in Europe

by
Anna Cristina d'Addio, Simon Chapple, Andreas Hoherz *and* Bert Van Landeghem*

The purpose of this paper is to examine the welfare effects of birth-related leave (BRL) in terms of life satisfaction. To do so, we exploit variations in BRL policies to assess their impact on life satisfaction. The paper adds to the existing literature in various ways. First, it uses new data collected by Baldi et al. (2011) and Baldi and Chapple (2010) to describe how life satisfaction moves around the date of the reforms over time and in a number of EU countries covered in the Eurobarometer surveys. Second, the paper analyses the relation between life satisfaction and BRL in Germany and the United Kingdom with long individual panel data collected with the GSOEP and the BHPS survey. The potential endogeneity bias of the treatment effect is addressed by building a quasi-natural experiment using policy changes as the assignment rule. The results from a variety of different methods suggest that BRL polices generally have a significant positive effect on life satisfaction. Women on BRL have higher life satisfaction, controlling for observable and unobservable personal characteristics. This result is robust to alternative specifications.

JEL classification: H53, I16, J38

Keywords: Welfare, subjective well-being, difference-in-difference, birth-related leaves

* Contact author: anna.daddio@oecd.org; OECD Social Policy Division, Directorate for Employment, Labour and Social Affairs. The authors wish to acknowledge comments and feedback from various colleagues. In particular from Andrew Clark and a number of OECD colleagues including Marco Mira d'Ercole and Conal Smith; from the participants at the 31st meeting of the OECD's Working Party on Social Policy held in December 2011; from the participants at the 2011 Joint workshop FORS on Redistribution and well-being held at the University of Lausanne and IDSC of IZA. The authors are also grateful for the advice and support of Bartek Lessaer, Ralf Jacob and Radek Maly, DG Employment, European Commission. Financial support from the European Commission is gratefully acknowledged. Views expressed represent exclusively the opinions of the authors and do not necessarily correspond to the views of the OECD or of the governments of its member states. The usual disclaimers apply.

1. Overview

An important development in social policy making in many OECD countries has been the introduction and expansion of paid maternity and parental leave schemes, especially since the 1970s.

Birth-related leave (BRL) schemes have a number of aims, including improving well-being of infants and supporting child development, enhancing well-being of caregivers, keeping parents (especially women) connected to the world of paid work, and helping families with issues of work-life balance. Such schemes account for a significant proportion of family spending by governments in many OECD countries today.

The changes of BRL schemes have encompassed increases in both the duration and types of birth-related leave available to women and men and modifications to payment systems as well (Baldi et al., 2011).

While the number of good quality evaluations of these policy changes, especially in terms of their influence on short- and long-term child well-being outcomes (OECD, 2011a) is increasing, very little research has explicitly addressed the question of whether this kind of social policy, can directly influence subjective well-being (see OECD, 2010).

The present study moves into this sparsely populated territory by seeking to assess the effect of BRL policy changes on women's well-being.[1] Only Pezzini (2005) has investigated the "treatment" effect of maternity leave on life satisfaction during women's fertile age.[2] To do so, Pezzini exploited the variations in leave duration across countries and time using *Eurobarometer* data (starting in the early 1970s) and a variety of control groups (men, post-menopausal women). The results of these studies do not show any significant effect of leave expansion on life satisfaction of women during their fertile years.

Similarly to Pezzini (2005), the present paper exploits variations in leave policies to examine whether changes in those policies have a welfare effect – measured in terms of life satisfaction.

Our paper complements and adds to Pezzini (2005) in many ways. *First*, Pezzini uses the dataset compiled by Ruhm (2000) which, though very valuable, omits a number of policy changes and contains some errors about their timing. These problems are rectified in the dataset used in the present paper as shown in Baldi et al. (2010) and Baldi et al. (2011). *Second*, to investigate how life satisfaction varies around the dates of relevant BRL reforms in various European countries a descriptive approach relying also on Eurobarometer data is presented. However, the data used in our paper cover a much longer period compared with Pezzini (2005), i.e. from 1973 to 2008 rather than from 1975 to 1998. *Third*, to assess whether BRL policy changes have had a welfare effect in Germany and the United Kingdom and to control for potential endogeneity bias, our paper uses those policy changes to build a quasi-natural experiment: women on BRL before the changes are considered as "untreated" while those on BRL after the policy changes are considered as "treated". The empirical analysis rests on individual panel data extracted from the German

Socio Economic Panel (SOEP) and the British Household Panel Survey (BHPS). These two surveys have much better data, both in terms of sample size and questionnaire stability, which matters for the measurement of subjective well-being.

The structure of the paper follows. Section 2 shortly reviews the literature, while Section 3 illustrates some trends in BRL policies in OECD countries. Section 4 presents a descriptive analysis of the relation between life satisfaction and BRL policy changes in some European countries covered in the Eurobarometer surveys. Section 5 describes the empirical strategy adopted in the paper and presents the econometric specifications used. Section 6 discusses the data and variables used for the analysis. Section 7 discusses the main results, while Section 8 draws some conclusions.

2. Review of the literature on birth-related leave and well-being

As many empirical and theoretical studies highlight, there is no universal definition of subjective well-being. On the one hand, psychologists tend to identify subjective well-being with "quality of life" comprising a "combination of feeling good and functioning effectively". On the other hand, economists and sociologists tend to use a narrower definition excluding some dimensions (such as physical and mental functioning) which are in contrast considered in the "Quality of Life" approach (see OECD, 2010; 2011a, b, c).

Two distinct components enter the definition of subjective well-being (Diener, 1984). They are: *i*) affective well-being (AWB); and *ii*) cognitive well-being (CWB). The first component (i.e. AWB) refers to emotional states and feelings, while the second may be seen as retrospective evaluation either of the overall quality of life (i.e. global life satisfaction) or of specific life domains (e.g. job satisfaction or marital satisfaction).

While acknowledging the ongoing debate on whether one component is better suited than the other for policy evaluation, this paper uses (overall) life satisfaction to measure subjective well-being because of its availability in both the SOEP and BHPS surveys and in Eurobarometer as well (see OECD, 2010; and OECD, 2011c).[3]

The paper tries to investigate the "welfare effects" – measured in terms of subjective well-being – of a particular type of social policy, i.e. birth-related leave (BRL). As discussed in OECD (2010, 2011b, 2011c), the study of policy measures in terms of individuals' satisfaction is a very modern development in economics. The research in this field is rapidly expanding. The reasons for this growing interest are many. Among those the acknowledgement of the need to go beyond GDP and the development of datasets for the analysis are particularly important (see OECD, 2010; OECD, 2011c).

For the analysis of the effect of BRL on life satisfaction, only Pezzini (2005) may be cited. A number of studies have instead looked at the relation between the length of BRL and some dimensions closely related to subjective well-being – such as mental health and self-reported health outcomes.

The conclusions regarding the effect of leave on these two outcomes are somewhat mixed. For example, Chatterji and Markowitz (2005) investigate the impact of the length of maternity leave on maternal health using two measures of depression and a measure of overall health status in a sample of working mothers in the United States. They use instrumental variables models to account for the possible endogeneity of the return-to-work decision. Their findings suggest that delaying the return to work reduces depressive symptoms but not the probability of clinical depression or hospital outpatient visits after childbirth.

In a later paper, Chatterji and Markowitz (2008) consider the impact of leave on maternal depression, overall health status and substance use. They again adopt an instrumental variables approach with county-level employment conditions and state-level maternity leave policies as instruments to address reverse causality issues. Chatterji and Markowitz (2008) use data collected post Family and Medical Leave Act (in 2001) – while prior work was based on data collected in the 1980s and early 1990s – and consider several aspects of family leave, i.e. the total length of maternal leave, the total length of paid maternal leave, and the total length of the father's leave. Their results suggest that a longer maternity leave is associated with declines in depression and improvements in overall health. Additionally, if men take paternal leave maternal depression declines. The size of these effects is however small. Doubling the length of maternal leave from nine to 18 weeks reduces: 1) maternal depressive symptoms by about 5%; 2) the likelihood of mothers' experiencing severe depression by 1%; and 3) and the likelihood of mothers' reporting overall poor/fair health status by 1%.

Using two waves of the *German Socio-Economic Panel* (1992 and 1997), Frey and Stutzer (2003) find a positive (but statistically insignificant) effect of maternity leave on life satisfaction. In contrast, Liu and Skans (2009) do not find any effect of the extension of the leave duration from 12 to 15 months (which took place in 1988 in Sweden) on parental well-being – measured by divorce and mothers' mental health.[4]

Other research uses quasi-experimental methods to assess the impact of BRL on a variety of outcomes. Baker and Milligan (2005, 2007) find that doubling the length of BRL in Canada to one year resulted in a large increase in the time women take off work and in breastfeeding rates. However, when it comes to maternal health – measured alternatively by an indicator of mothers' self-reported health, an index of depression, a binary indicator of no post-partum depression, and a count of post-partum problems – the authors do not find any impact related to the extension of the duration of BRL (see also Gjerdingen and Chaloner, 1994; Hyde et al., 1995; McGovern et al., 1997).

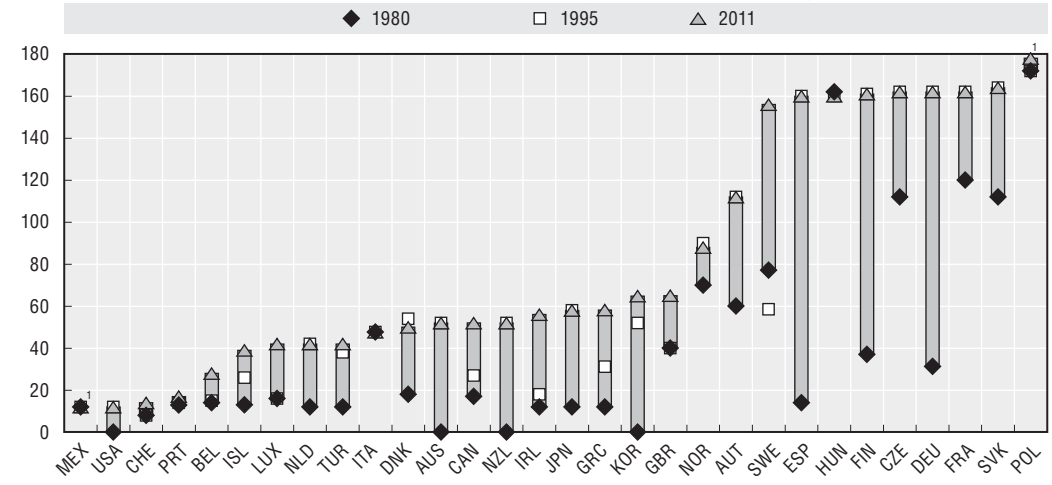
There is also a large body of literature which investigates the link between subjective well-being and other domains closely related to BRL (see Lalive and Zweimuller, 2005, for example; Staehelin et al., 2007). Some studies provide evidence of the so-called "baby honeymoon". For example, Dyrdal et al. (2010), using a very large sample of Norwegian women, report increases in life satisfaction from early pregnancy to about six months of age of the child and thereafter reductions up to the age of three of the child (see also Gjerdingen et al., 1993, 1995). Earlier research in both Britain and Germany had suggested some interesting dynamics of life satisfaction immediately before and immediately after the birth of a child (Clark et al., 2008; Clark and Oswald, 2002).

Some of the most recent literature on family policies and subjective well-being use policy changes to identify causal effects. A small but growing literature focuses on the causal effects of maternity leave and childcare policies on child well-being outcomes (Dustmann and Schönberg, 2008; Khanam et al., 2009; Tanaka, 2005; Wurtz, 2007; Zarrabi, 2009). Other studies estimate the causal impact of these policies on parental well-being. For example Brodeur and Connolly (2013) use triple-differences methods to investigate the effect of a change in childcare subsidies policy in Québec aimed at increasing subsidised childcare space. Their findings suggest that this policy led to a small decrease (1/20th of a standard deviation) in parents' self-reported well-being.

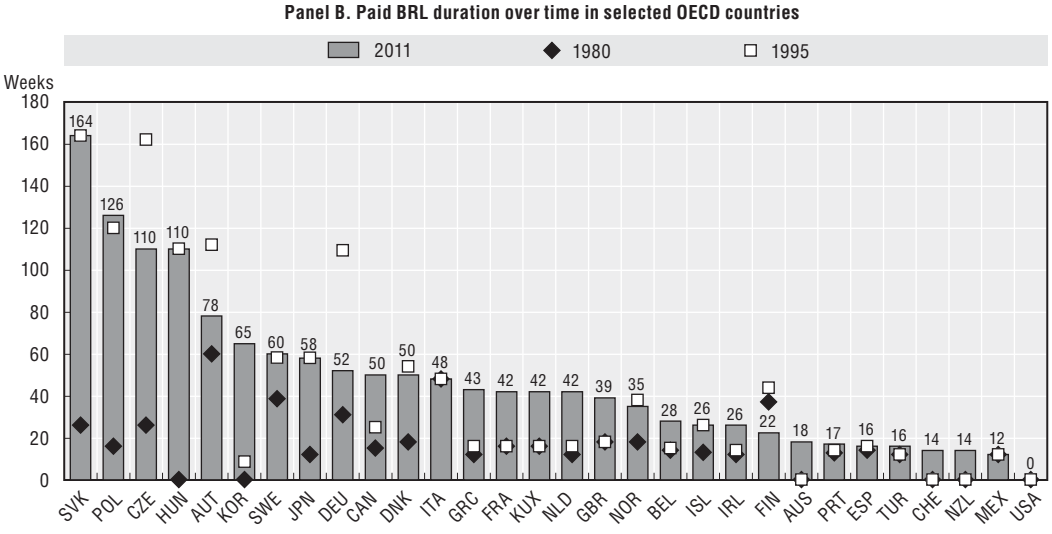
3. Trends in birth-related leave in OECD countries

Panel A of Figure 1 illustrates the trends in the duration of (paid and unpaid) BRL over time while Panel B shows the trends in the duration of paid, job-protection birth-related and post-birth leave over time.[5] The duration of leave (both paid and unpaid) has increased almost everywhere between 1980 and 2011 (top panel). The same trend is observed for the duration of paid BRL shown in the bottom panel (see Deven and Moss, 1999, 2005; Escobedo, 2008). However cross-country differences are large; for example large increases are displayed in the Slovak and the Czech Republics, Hungary, Poland and Korea. Conversely, a very small or zero-increase is observed in Italy, Portugal, Spain, Mexico and Turkey.

Figure 1. **Trends in the BRL duration (paid and unpaid) over time**



1. The bar on the chart shows the difference between the values observed in 1980 and in 2011.



Source: Based on Baldi and Chapple (2010) and on the *OECD Family Database*.

The starting level and the entitlements granted vary greatly across countries. Some countries (including Ireland, the United Kingdom and Denmark) have seen this sort of leave introduced for the first time over the time period considered. Italy, Spain and Turkey

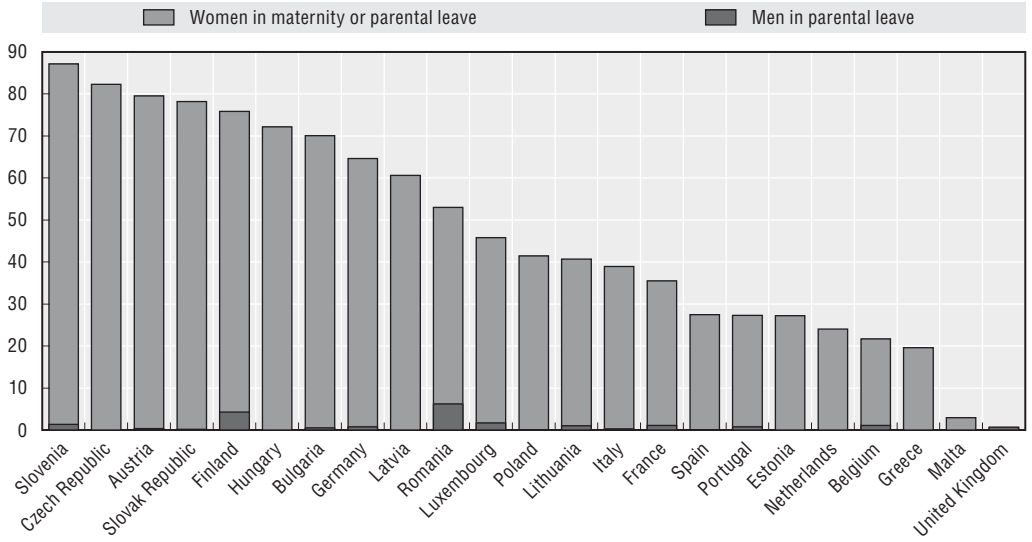
have known relatively few, and fairly small, policy changes in leave durations. In other countries, such as the Czech Republic and Germany, durations of leave rise from a few months to a time scale measured in years (see also Bird, 2001). While it is correct to state that the time trend is always upwards, in several cases there are falls in leave durations – for example in Hungary, Denmark, Finland, Poland and Sweden.

4. Is there a link between life satisfaction and birth-related leave?

To give a more international perspective to the analysis, this section analyses the effects of BRL policies on life satisfaction using Eurobarometer data for a number of EU countries experiencing major changes in those policies over the period spanning from Autumn 1973 to Spring 2008. Only women are considered in the analysis.

Figure 2 shows the parental/maternity leave take-up by employed mothers and fathers with a child below 1 year of age. It suggests that women still take the vast majority of leave on offer despite fathers' leave take-up having increased over time.[6]

Figure 2. **Leave take-up by (employed) mothers and fathers, 2006**



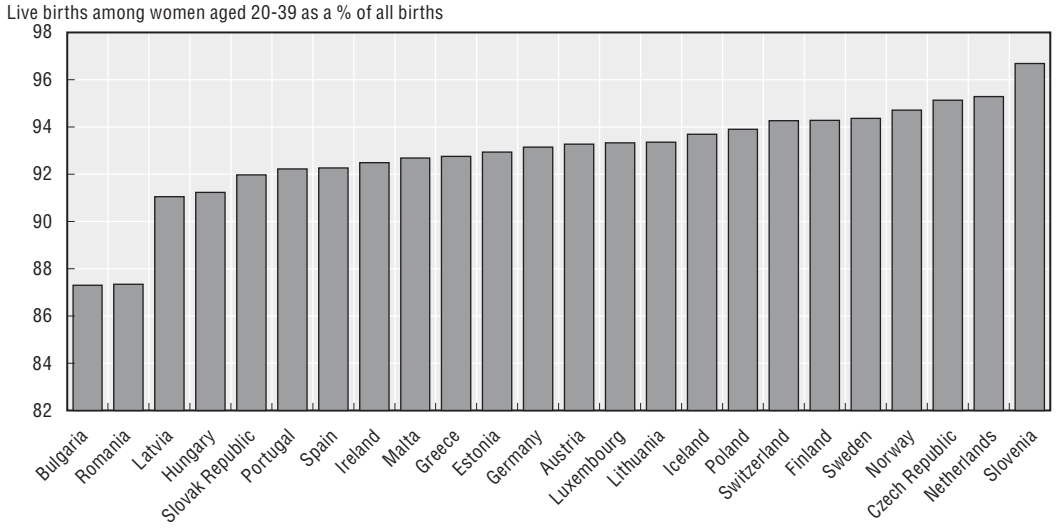
Source: OECD Family Database (Indicator PF 2.2) based on ELFS, 2006.

The analysis in this section uses data on life satisfaction reported in various Eurobarometer surveys, where life satisfaction is measured on a four-point scale – i.e. not at all satisfied, not very satisfied, fairly satisfied and very satisfied. Because these surveys do not record whether women are or are not on BRL, the assignment to the "treatment" and "control" groups, respectively, is done with respect to the age of the survey respondents.[7] This is the approach adopted for example in Pezzini (2005) who assigns women aged 15-49 to the treatment group (the "fertile" age) and those aged 50 and over to the control group.

In this paper, and differently from Pezzini, only women aged between 20 and 39 (rather than those aged 15-49) are assigned to the treatment group. Using Pezzini's definition would have resulted in the two groups overlapping over the long observation period considered. Moreover, many women included in the treatment group would have achieved their desired number of children which might have worked against finding any policy effect.

Another reason justifying the choice of the age limit adopted here is that very few women in Europe have babies after age 39: on average in the twenty four countries represented on the chart, only 7% of the total live births occur either before 20 or after 39 years of age (Figure 3). Data on the age of mothers at birth in Europe also supports the definitions of treatment and control groups used in this article.

Figure 3. **Percentage of live birth among women aged 20-39 as a % of all births, 2011, in selected EU countries**



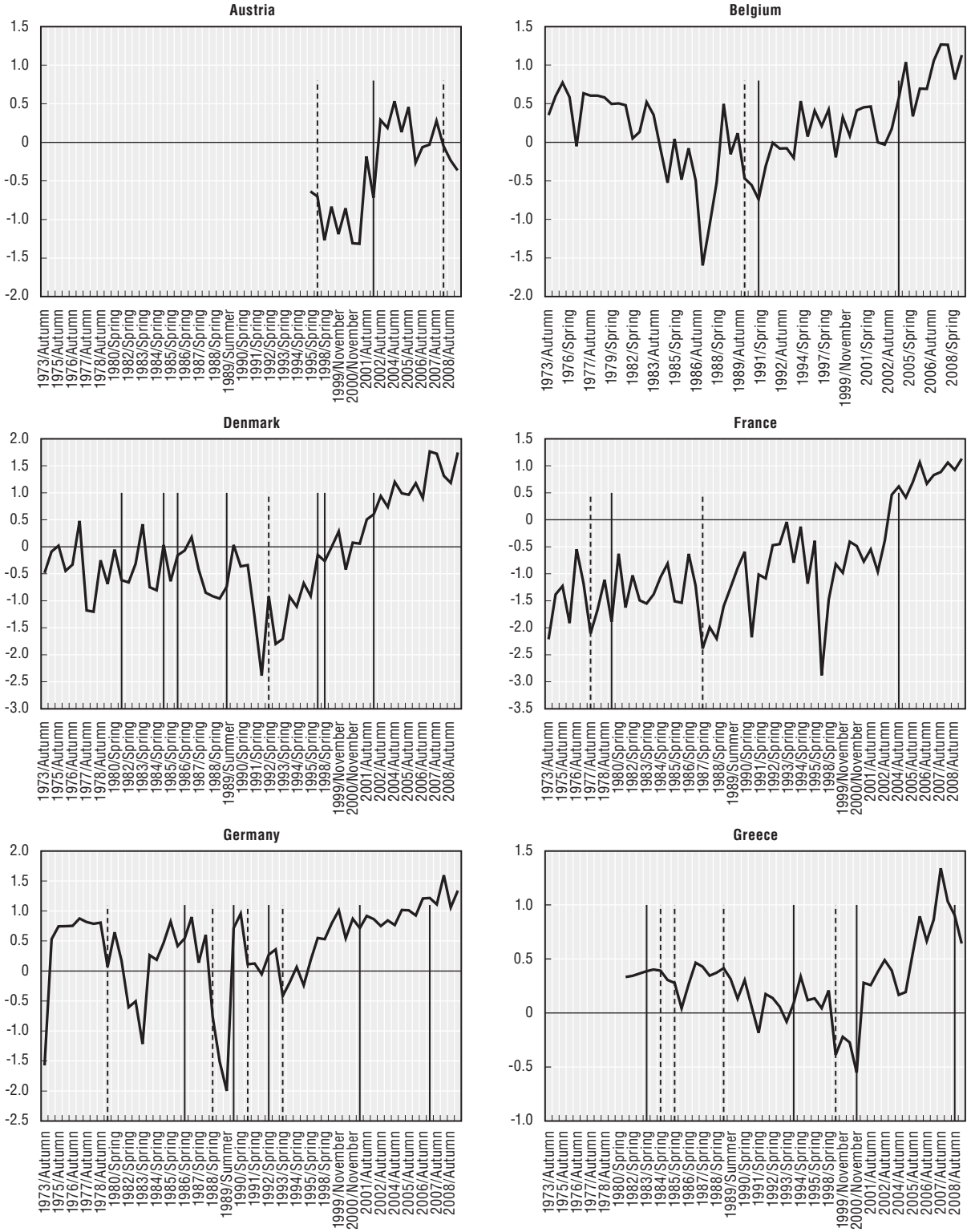
Source: Eurostat for countries with data available.

One could argue that the definition used in Pezzini (2005) is better than the one used here because in a number of countries parents are currently eligible for BRL until their child is quite old (e.g. age 8 and over in some countries), which would tend to push up the effective age of the treatment group (i.e. the women eligible for BRL). However, most parents use the vast majority of leave in the early years when their child is an infant or toddler. Data from 2006 *Labour Force Surveys* indicate, for example, that on average 46% of employed women are on maternity or parental leave when the child is aged less than one year (Figure 2). This percentage drops to 5% when the child is aged less than three.

A simple visual inspection of the difference in life satisfaction between women aged 20-39 years (i.e. the "treated") and women aged 50 years and over (i.e. the "untreated") around the policy changes does not reveal any specific trends either increasing or decreasing the relative life satisfaction of women who were potentially eligible for BRL policies. This result might be simply a statistical artefact reflecting relatively small sample sizes (a few hundred women in most cases) and high sampling errors.

To gather a better understanding of the variable of interest a mild formalisation has been adopted to smooth the data.[8] To do so, we compared the difference in life satisfaction between the two groups two periods before and two periods after the reform. The vertical lines in Figure 4 identify the date at which the reform takes place: a dotted line states that the reform did not lead to an increase in life satisfaction and a solid line indicates when it does so. The analysis suggests that life satisfaction has increased over decades in many of the countries considered.

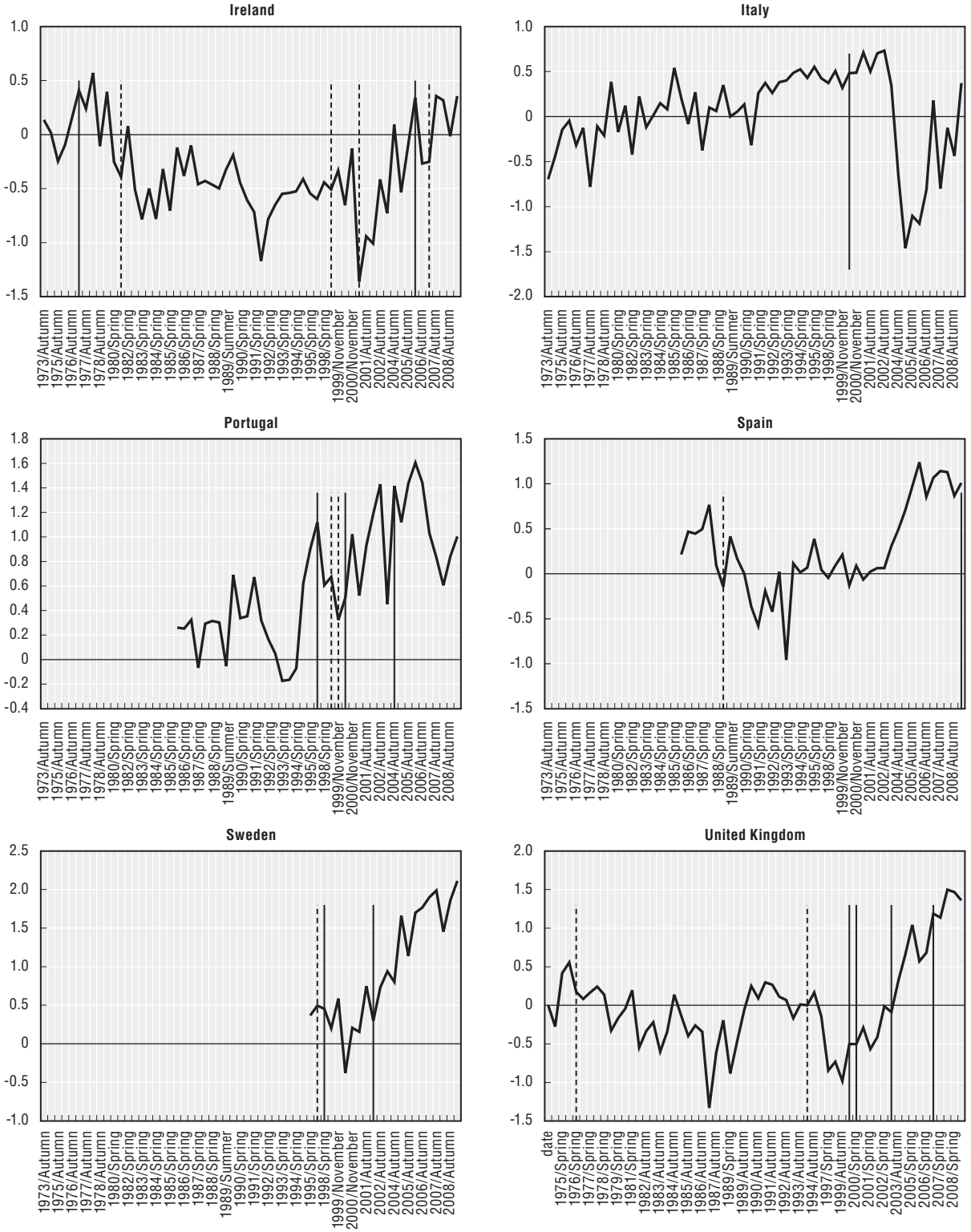
Figure 4. **Reforms of parental leave and differences of life satisfaction between treatment (women aged 20-39 years) and control groups (women aged 50> years)**



Note: The vertical lines indicate the point in time when the reform occurred. A dashed line indicates that the reform did not increase life satisfaction while a solid line does indicate an increase in life satisfaction. The analysis considers differences in life satisfaction two periods before and after the reform date.

Source: Reforms data from Baldi and Chapple (2010); data on life satisfaction from different Eurobarometer surveys (see Appendix A).

Figure 4. **Reforms of parental leave and differences of life satisfaction between treatment (women aged 20-39 years) and control groups (women aged 50> years)** *(cont.)*



Note: The vertical lines indicate the point in time when the reform occurred. A dashed line indicates that the reform did not increase life satisfaction while a solid line does indicate an increase in life satisfaction. The analysis considers differences in life satisfaction two periods before and after the reform date.
Source: Reforms data from Baldi and Chapple (2010); data on life satisfaction from different Eurobarometer surveys (see Appendix A).

However, this does not mean that trends in life satisfaction are related in a causal way to the reforms. There are good reasons to believe that most of the variations in the chart are due to other factors (such as the increased participation of women in the labour market and their different role within the household) and are unrelated to the policy effect. Only a more rigorous econometric analysis may help to shed some light on this issue.

5. Estimation strategy

The paper uses a number of econometric strategies to assess whether changes in BRL policies affect well-being. Because (relatively) long time series of individual data are necessary to adequately account for the endogeneity bias, the analysis is limited to two OECD countries, i.e. Germany and the United Kingdom.

Concerning the estimation strategy, the literature on subjective well-being generally uses ordinary least squares (OLS) for reasons of simplicity and ready interpretation (Ferrer-i-Carbonell and Frijters, 2004).

The OLS specification can be written as:

$$Y_i = \beta_0 + \beta' X_i + \alpha D_i + u_i \qquad [1]$$

where β_0 is the constant, X contains the explanatory variables, D is the treatment dummy (equal to 1 for the treated and to 0 for the untreated) and u is the error term. The choice of X (i.e. the explanatory variables) rests largely on the standard specifications used for the estimation of life satisfaction. Standard controls include employment status, number of children, nationality, years of education, age, net household income, and marital status (see OECD, 2010).

However, because of the potential endogeneity of the relation between being on BRL and life satisfaction – i.e. between D_i and Y_i in equation [2] – a model with fixed effects is also estimated. In particular, fixed personality characteristics of individuals may simultaneously give rise to both life satisfaction and decisions to be on BRL: mothers on BRL may have different time-invariant personality traits, information or preferences than mothers who are not on BRL. These traits in turn may be correlated with subjective well-being. Individual fixed effects help remove these, controlling for the unobserved factors which do not change over time (Wooldridge, 2007).

The specification estimated, which includes individual fixed effects, is in this case:

$$Y_{it} = \beta_0 + \beta' X_{it} + \alpha D_i + u_{it}, \qquad \text{with } u_{it} = \mu_i + \nu_{it} \qquad [2]$$

where μ_i are individual-specific, time-invariant effects $i = 1, \ldots, N$ denotes the individuals and $t = 1, \ldots, T$ the time periods. In equation [2], y_{it} is life satisfaction of individual i at time t, α is the parameter of interest which is associated to either $D_i = 1$ for the treated or $D_i = 0$ for the control group – i.e. the treatment is here "being on BRL"; and X_{it} are the explanatory variables.

While the use of fixed effects and multivariate controls address some of the issues associated with omitted variables and causality, it is still possible that the coefficient associated to the "treatment" will be biased upward due to high-life satisfaction mothers who are more likely to take BRL.

When assignment to the treatment or control group is random, D_i can be considered as exogenous, and the standard OLS estimator is consistent because only when $E(u_{it} \mid D_i, X_{it}) = 0$, α is the average treatment effect (ATE). This occurs in data with randomised experiments or controlled social experiments, but not in the non-random or non-

experimental data that are common in practice. When the assignment to the treatment group is non-random, the OLS estimator does not yield consistent estimates of α and β because u_{it} and D_i are correlated, which yields:

$$E(u_{it} \mid D_i, X_{it}) \neq 0 \text{ and } E(Y_{it} \mid D_i, X_{it}) \neq \beta' X_{it} + \alpha D_i \quad [3]$$

One way of addressing this issue is through instrumental variable (IV) even though good and valid instruments are difficult to find. Indeed, to be valid the instruments Z, should be: *i)* exogenous – i.e. uncorrelated with the errors such that $corr(Z'u) = 0$; *ii)* correlated with the endogenous X, i.e. $corr(Z'X) \neq 0$; and *iii)* correlated with y only through X.

$$Y_{it} = \beta_0 + \beta' X_{it} + \pi, Z_i + u_{it} \quad [4]$$

The instruments used here are the time at which policy changes occurred which are deemed to affect the probability for women to be on BRL.[9]

A fundamental problem of non-randomised policy evaluation derives from the impossibility for the observational unit of being either treated or untreated at the same time. Longitudinal data combined with (exogenous) policy changes represent a solution to this puzzle because they provide repeated measures of the same individuals before and after the policy change which may thus be seen as treated and untreated.

To assess the robustness of the estimates of the welfare effect obtained with the IV-FE estimator, a difference-in-differences estimator (DiD) is therefore used. In this model, the outcome Y_i (i.e. life satisfaction) is modelled by the following equation:

$$Y_i = \beta_0 + \alpha D_i + \gamma t_i + \delta(D_i.t_i) + u_i \quad [5]$$

where α is the treatment-group specific effect (to account for average permanent differences between treatment and control); γ is the time trend common to control and treatment groups; δ is true effect of treatment. The purpose of the programme evaluation is to find a "good" estimate of δ, given the data available.

The method considers the differences between life satisfaction before and after a treatment for participants and for non-participants. It computes thus a "double difference" (see Wooldridge, 2007). The basic idea behind the DiD estimator is that taking the difference-in-difference cancels out the difference in period effects, leaving only the impact, as desired. The assumption here is that women entering the BRL after each policy changes are "treated", those who were on BRL before the policy change are "untreated". This implies that the date of the policy change is considered as the assignment rule. Results based on this estimator are likely to differ from those illustrated in Figure 4 for a number of reasons. For example, in Figure 4 we do not control for some observed heterogeneity, the control groups are different as the data are as well.

It is important to highlight that even though the magnitudes of the estimates are not expected to be the same across the different specifications, they might happen to be so because the fixed-effect estimator on panel data may be seen as generalisation of a difference-in-difference design.

6. Data and variables

The analysis focuses on Germany and the United Kingdom because they both combine the availability of long panel data with a number of policy changes in BRL policies which can be used for the identification strategy in the quasi-experimental approach.[10]

6.1. The case of Germany

The empirical analysis for Germany uses data from the *German Socio-Economic Panel* (GSOEP) 2009 release (1984-2008). The GSOEP is a representative, longitudinal survey of more than 20 000 respondents in about 12 000 private households in Germany. The survey has been conducted every year since 1984 with the same persons and families. The sample has been amended several times to remain a long-term, longitudinal representative set of individual and household data in Germany.

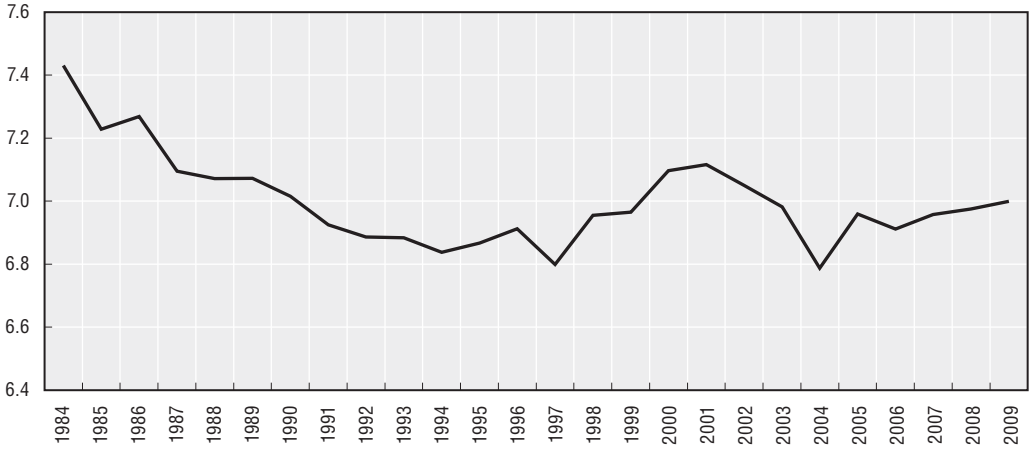
Only women are considered in the analysis because the number of men ever on BRL in the *GSOEP* is extremely low. Low male representation in part reflects a general OECD trend as illustrated in Figure 2 above.

In terms of the variable, this article is seeking to explain, in each interviewing year of the GSOEP all adult household members are asked to rank their overall life satisfaction, using an 11-point scale. The level of life satisfaction is based on responses to the question:

> *Finally, we would like to ask about your overall level of life satisfaction. Please answer again according to the following scale, "0" means completely and totally dissatisfied, "10" means completely and totally satisfied. How satisfied are you at the present time, all things considered, with your life?*

This measure is used to approximate respondents' subjective well-being over the sample period. The average female life satisfaction over the period was 6.99 with a standard deviation of 1.85. Its trend is illustrated in Figure 5.

Figure 5. **Average life satisfaction of women**

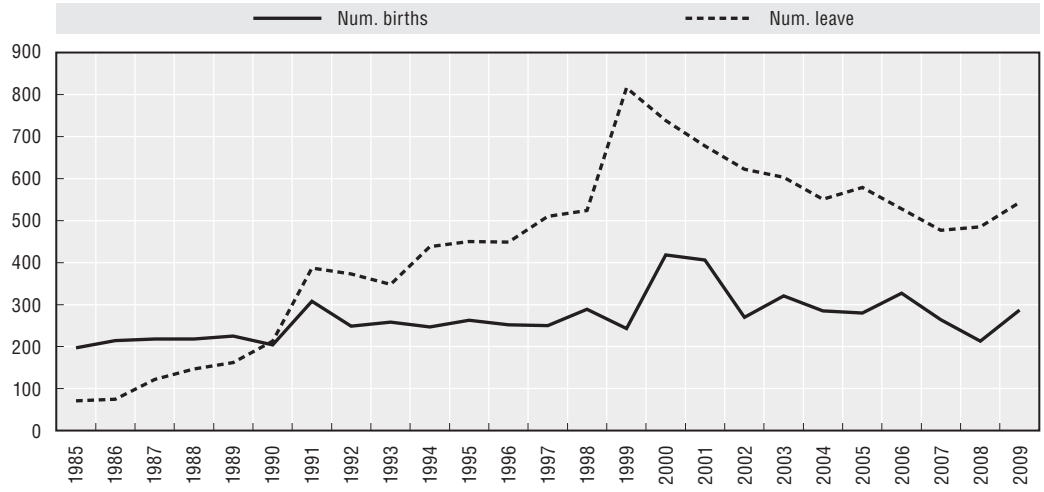


Source: German SOEP, 1984-2009.

As no single unique question for BRL is included over the entire sample period, the analysis uses a combination of questions about the particular months of the previous year in which the respondents received maternity pay (1984-90) and the particular months of the previous year in which the respondents have been on parental or maternity leave (1991-2008). The question, in both instances, relates to the January-December calendar year prior to the year the subject was interviewed, regardless of the interview month. The data are depicted in Figure 6.

Data from the *OECD Social Expenditures Database* indicate that Germany spent EUR 5 269 million on social policies in 2007; 11.8% of these expenditures were devoted to

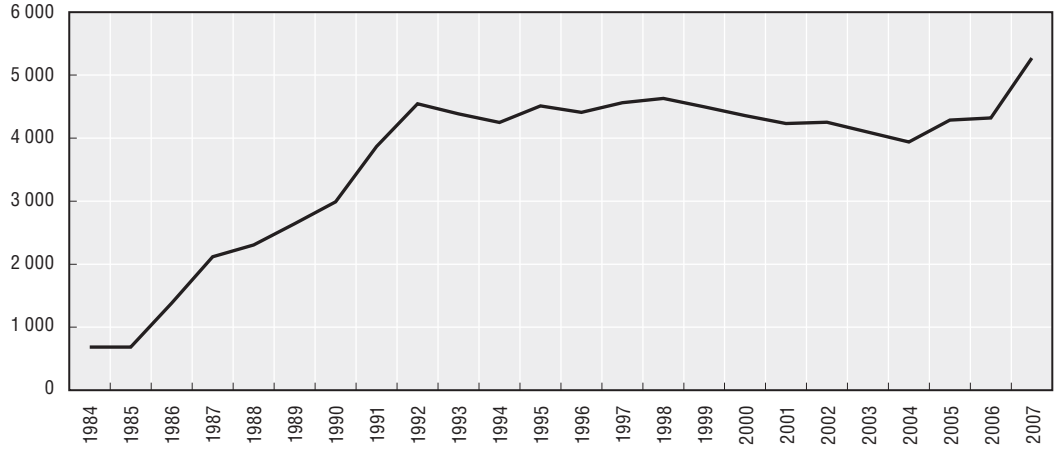
Figure 6. **Number of births to women in the German SOEP and number of women on BRL by year**



Note: Data on the number of births and leave are based on the GSOEP.
Source: German SOEP 1984-2009.

spending on birth related leave policies which represented a 7 percentage point increase from 1984. Public expenditures on BRL policies in Germany, taken from the *OECD Social Expenditure Database*, are shown in Figure 7.

Figure 7. **Public spending on BRL benefits in Germany, millions of euro**



Source: OECD SOCX Database.

For the identification strategy of the causal effect of the reforms of BRL on life satisfaction of women in the quasi-natural experiment, changes in birth-related leave policies (see Box 1) were used as an assignment rule to treatment and control groups, respectively. The treatment group includes women with children who were on BRL after the policy change, while the control group is defined by women with children who were not affected by the policy change. Eight policy changes relating to birth related leave are examined as detailed in Box 1.

Box 1. **Federal BRL policy in Germany***

(West) German maternal and parental leave policy has a long history, dating to the 1870s. In 1878, a code was introduced to forbid pregnant women working for three weeks before birth (Wikander et al., 1995). There were further amendments to the code in 1903 and 1911, which increased the leave period to six weeks and supplied women with paid time off work for two weeks before delivery (Merz, 2004). In 1924, job protection was introduced for women taking maternity leave (Jordan, 1999).

The modern expansion of BRL in (West) Germany dates to the late 1960s. From 1 January 1968, employed women expecting a child were granted fourteen weeks of mandatory maternity leave (*Mutterschutzgesetz*). Six weeks had to be taken before and eight weeks after birth. During those 14 weeks, a sickness insurance benefit was paid by the social security system at a flat rate equal to about the average salary for women workers. Employers were required to supplement this benefit to cover the woman's full salary (Merz, 2004). East Germany regulated maternal leave in 1972 introducing the "*Babyjahr*", i.e. a one-year maternity leave with a benefit equal to the level of sickness benefits. This right was granted only to mothers.

From 1 January 1979 (Merz, 2004) or 1 May (Dustmann and Schoenberg, 2008) employed women on maternity leave could opt to take an additional four months of leave immediately following maternity leave. From six weeks before to eight weeks after childbirth, mothers were paid their average income, estimated over the three months before giving birth. For the additional four months, payment was a flat rate of roughly one-third of average pre-birth earnings (Dustmann and Schoenberg, 2008). Women could not be dismissed and had the right to return to their employer, albeit not necessarily to their previous job (Merz, 2004). The prime motivation of the 1979 reform was maternal health. The later reforms would be more focused on enhancing child development (Dustmann and Schonberg, 2008).

From 1 January 1986, a new parental leave scheme was introduced. Following the 14-week maternity leave, parents were entitled to a further eight months of job protected parental leave as a shared right, with a total of post-birth job-protected leave of 10 months. The benefit paid became disconnected from an employment condition. Any new parent working less than a maximum of 15 hours per week was entitled to receive a benefit from the federal government, regardless of his or her previous labour market status. The parental benefit equaled DM 600 for the first six months of the additional eight months (about 20% of average pre-birth wages). The benefit during the seventh and eighth months was means tested, based on family income before childbirth. A married couple received the benefit as long as annual net family income was less than DM 29 400. For a single parent, this income limit was DM 23 700 per year. Each additional child increased the upper limit by DM 4 200 (Dustmann and Schonberg, 2008; Merz, 2004; Kamerman and Kahn, 1991).

There was an increase from eight months to ten months job-protected parental leave from 1 January 1988, a further rise to 13 months from 1 July 1989, and a final rise to 16 months from 1 July 1990 (Dustmann and Schonberg, 2008; Merz, 2004). Payment was extended at the same time.

The next reform was introduced from 1 January 1992. Eligible parents were entitled to take job-protected leave up to three years after the birth of their child. Payments still lasted in total for 18 months following birth, with eight weeks maternal leave and 16 months parental leave (Merz, 2004; Gauthier and Bortnick, 2001). Thereafter parents could use unpaid but job-protected parental leave up to the child's third birthday.

Box 1. **Federal BRL policy in Germany*** *(cont.)*

The 1992 reform was further extended from 1 January 1993, with the payment period for parental leave rising from 16 to 22 months, making a total of two years paid maternal and parental leave (Dustmann and Schonberg, 2008; Merz, 2004). As Dustmann and Schonberg (2008) point out, the federal government had committed to providing subsidised child care for every child over the age of three from 1996. The intent was to encourage mothers to stay at home until the child was three, and then aiding a return to work via subsidised child care thereafter.

Reforms from 1 January 2001 introduced flexibility in the parental leave system. Parents were able to choose a shorter and better paid leave (DM 900 per month for 10 months) or a longer but less well paid leave (DM 600 per month for 22 months). Benefits remained income tested. Paid leave could be used until child's second birthday. The third year of leave could be used until a child reached eight years of age (Merz, 2004).

On 1 January 2007, a new earnings-related parental leave benefit with floors and ceilings was introduced (*Elterngeld*, or "parental money"). It replaced the old, means-tested flat rate benefit (*Erziehungsgeld*). The duration of the job-protected maternal plus parental leave remained at a maximum of three years following childbirth. The parental leave payment was 67% of parent's average earnings during the year before childbirth, with a ceiling of EUR 1 800 per month and a floor of EUR 300. The new parental leave payment was for ten months, plus two extra months for the father if he used at least two months of parental leave, making 14 months of payment in total when including the eight weeks post-birth maternity leave. The maternity leave payment was included in this period, reducing the actual *Elterngeld* payment period to 12 months. It is possible to extend leave up to 24 plus four months (if each parent takes at least four months leave), with a proportionate reduction in the monthly payment rate. The actual *Elterngeld* payment period was then 28 months less the two months maternity payment, i.e. 26 months. (Moss and Korintus, 2008, p. 208). The effects of the 2007 policy change differed for rich and poor parents. Before the policy change, the poor received *Erziehungsgeld* of EUR 7 200 in total over 24 months. Post-reform they get EUR 3 600 (12 months) to EUR 4 320 (14 months). Pre-2007, richer parents also received EUR 7 200, or nothing if they were in excess of the income threshold. Post-2007, they could obtain up to EUR 25 200.

* A number of birth related leave policies are also implemented in various *Länder*. These are detailed in Appendix B. For example, while the parental leave legislation is federal, four *Länder* (Bavaria, Baden-Württemberg, Thuringia, Saxony) also pay a means-tested benefit up to the third year of parental leave.

As mentioned already, the analysis in the present paper only focuses on women. The inability to investigate the effect of BRL on fathers' life satisfaction is a major limitation of the analysis which we need to acknowledge here.

It is however difficult to examine how the take-up of leave has evolved in Germany following the implementation of these different BRL policies because of unavailability of the relevant data in many cases. However, data from the Federal Statistics Office for the year 2010 highlight that around 810 000 mothers and fathers obtained paid parental leave support. The majority of applications were filled by single parents and among these 25% were fathers.[11] But more than three-quarters of those fathers asked for parental benefit support for less than two months. In contrast, couples asked for support on average for a period of two months.

The labour market status of the mother was a decisive factor in determining who was asking for support: when the mothers were employed in the twelve months before the birth the probability that the fathers, too, take a parental leave increases (see D-Statis, 2012). 32% of the partners of women employed before the birth of the child filed an application for parental leave support in 2010 against 12% in the case of women who previously were not employed. On average the sum paid to mothers and fathers was EUR 964 – EUR 878 were paid to mothers and EUR 1 201 to fathers on average – but a third of the applicants only received the minimum amount of EUR 300. If fathers apply for paid parental leave benefits, they generally do this for the first three months of the new-born child (41%). Every fifth father opted for support after the first year of the child.

A number of controls have been used in the analysis. The descriptive statistics for the sample of women with children used in the analysis suggest that 62% of women are employed, 60% are married, 6% divorced and 1% separated, 29% are aged below 29 years, while 28% are aged between 30 and 39, 14% are aged between 40 and 44 and 27% are aged over 44 years. On average, they have been in education for 11 years; 21% live in North-Rhine Westphalia, 14% of women live in Baden-Württemberg, 3% in Berlin and 1% in Bremen. To capture "anticipation" and "actuality" effects related to childbirth, the specification includes a number of leads and lags prior to and after birth. The choice of leads and lags around birth reflects both biological factors and findings of the research on the "honeymoon period". They are introduced as controls in the regression to capture variations in subjective well-being which are not strictly related to the BRL policy.

6.2. The case of the United Kingdom

The empirical analysis for the United Kingdom uses The *British Household Panel Survey* (BHPS) – made available through the ESRC Data Archive and originally collected by the ESRC Research Centre on Micro-social Change at the University of Essex. The BHPS started in 1991 with households from 250 areas in Great Britain. In later years, samples from Wales, Scotland and Northern Ireland were added.

At the date of the analysis, the BHPS consists of 18 rounds of data, containing rich information about people's activities and socioeconomic status. Unfortunately, however, not all of these rounds contain all the information on the dependent variable – subjective well-being – that should ideally be used throughout the analysis. A general life satisfaction question is only included from the year 1996 onwards, and has been skipped in 2001, which means there are subjective well-being data for the period 1996-2000 and 2002-08. Additionally, the life satisfaction measure is not exactly comparable to the German measure. The "satisfaction with life" data for waves 6-10 and 12-18 are based on the following survey question:

> *Please tick the number which you feel best describes how dissatisfied or satisfied you are with your life overall? 1 (not satisfied at all) to 7 (completely satisfied).*

This seven point question thus differs from the 11 point life satisfaction question of the GSOEP. Across these 12 waves and all surveyed individuals, respondents give an average life satisfaction score of 5.22 on a scale of 7, with a standard deviation of 1.29. In 1.5% of the responses, an individual indicates being completely dissatisfied (a score of 1), while in 14.5% of cases respondents are completely satisfied (a score of 7).

In total, there are 2 120 individuals on BRL. Only 28 of them, however, are male. The analysis will therefore be restricted to female individuals.

The detail of BRL policy chances in the United Kingdom is presented in Box 2. In summary, there were a maximum of five policy changes to be considered during the analysis. These policy changes are seen here as exogenous, i.e. not related to people's circumstances or personality, but are likely to be correlated with the demand for parental leave. They are thus candidates for instrumental variables in regressions exploring the effect of parental leave on subjective well-being. The dates of the five policy changes are as follows: 1) 16/10/1994; 2) 15/12/1999; 3) 30/04/2000; 4) 06/04/2003; and 5) 01/04/2007.

Box 2. **BRL policy changes in the United Kingdom**

In 1973, the United Kingdom had a paid but not job-protected maternity allowance of 18 weeks (Sargeant and Lewis, 2008). From 1 June 1976, the 1975 Employment Protection Act introduced the right to return to work up to 29 weeks following birth. Maximum leave before birth was 11 weeks (thus 40 weeks job-protected leave in total). Payment was provided for 18 weeks, consisting of the first six weeks paid at 90% of earnings and the remaining 12 weeks with a flat rate. Eligibility criteria include at least two years full-time work with the same employer for 11 weeks before childbirth, or five years part-time (Zabel, 2009). Although the 1975 Employment Protection Act received parliamentary assent on 12 November 1975, it was implemented until 1 June 1976 for a right to return to work on 6 April 1977 (Fonda, 1980).

From 16 October 1994, all employed pregnant women, regardless of hours or length of service, were entitled to 14 weeks of job protected maternity leave with all normal contractual entitlements except pay. Women who have two years continuous service were entitled to an additional period of maternity leave, lasting from the end of statutory maternity leave until the 28th week after the baby was born. Women could start maternity leave at any time from 11 weeks before the date of expected childbirth (still 40 job protected weeks). Women who had been with their employers for 26 continuous weeks, ending the 15th week before the expected date of birth were entitled to receive Statutory Maternity Pay for 18 weeks, with six of these weeks at 90% of earnings and 12 weeks at an improved flat rate (Callendar et al., 1997).

From 15 December 1999, parents with children under five years of age were each entitled to up to 13 weeks of unpaid leave. Where individual employers had not chosen to negotiate their own arrangements with employees, leave allowed within one calendar year was limited to four weeks.

From 30 April 2000, all employed pregnant women, regardless of hours or length of service, were entitled to 18 weeks of job protected maternity leave with all normal contractual entitlements except pay, up from 14 weeks (Gregg et al., 2007). This period further rose to 26 weeks from 6 April 2003: all employed pregnant women, regardless of hours or length of service, were entitled to 26 weeks of job protected maternity leave with all normal contractual entitlements except pay. Also from 6 April 2003, maternity leave increased to one year. It was divided into ordinary and additional maternity leave. The first period and second period were both 26 weeks. The payment was at 90% of earnings for the first six weeks, but the flat rate paid period was now 20 weeks long (up from 12 weeks). The 26 weeks additional maternity leave was unpaid. Paternity leave was introduced around the birth of a child for two weeks at flat rate payment. Mothers and fathers still both have access to 13 weeks of statutory unpaid parental leave while the child was less than five years old.

From 1 April 2007, the flat rate payment period increased from 20 weeks to 33 weeks.

Data from the *SOCX Database* suggests that public expenditures on social policies in 2007 were considerable also in the United Kingdom: EUR 5 025 million. The share of expenditures devoted to BRL policies amounted to 11% of the total social policy spending in the same year.

According to Moss (2012), the mean length of maternity leave taken by women increased by approximately two months between 2006 and 2008. A large majority of fathers (91%) took some leave around childbirth. 49% of them took statutory paternity leaves – 50% of them for a duration of two weeks and 34% for a shorter duration. 25% took statutory paternity leave plus other paid leave. Moss (2012) also highlights that the probability of taking paternity leave was higher for men working in the public sector. Data on the take-up of parental leave in the United Kingdom are limited. Those available indicate that in 2008, only 5% of mothers used "fully" paid parental leave in their first-job after birth compared with 17% of employed fathers.

7. Results

7.1. Germany

Table 1 reports the result of the various regression models estimated: column 1 is the simple OLS, column 2 is a FE and column 3 is the FE IV.

Some of the OLS estimates are in contradiction with FE and IV-FE estimates, in particular, OLS cannot control for cohort effects independently of age. This may be related to the inadequacy of the OLS estimator to control for unobservable characteristics of individuals which are pooled over time. The OLS estimator also mixes up cohort and age effects whereas FE holds cohort constant. The estimates of household income and of the treatment are also bigger in the OLS than the corresponding FE and IV-FE estimates which could be related to an endogeneity problem.

In contrast, the estimates are very stable across the IV and FE specifications and suggest a number of interesting patterns. First, being on BRL positively and significantly affects life satisfaction. This effect is robust to all specifications estimated and may be related to the ability for mothers on BRL to preserve the positive effects of being employed while they stay at home and raise their children. The highly significant coefficient is 0.125, which is very close to the impact on life satisfaction of being in paid employment. Perhaps it is the job protected nature of the birth related leave which is driving the similarities between being at work and being at home but with a job guarantee. Note that because the regression controls for household income, the coefficient on BRL refers only to the non-monetary effects of leave. However, this non-monetary effect will over-estimate the positive influence of leave on life satisfaction when BRL status is associated with less than full replacement of market wages and leads then to a reduction in household income.[12]

Both the fact of being employed and having a higher income increase life satisfaction. The results also suggest that life satisfaction declines up to the age 30-34 and then starts increasing with age. Older women with children seem more satisfied compared with younger ones. There are many reasons for this, but it may be that the ability to cope with family and professional life has worsened in Germany. Younger women may also be less satisfied because of the lower number of children they have in recent decades compared with older women in the past. In fact, the coefficient associated with the number of children suggests that life satisfaction increases with additional children. Married women also display higher satisfaction than non-married, divorced, separated women and

Table 1. **Estimation results**

	OLS	FE	FE-IV
		Lifesat	
Treatment effect	0.2652**	0.1272**	0.1250**
	(0.0468)	(0.0369)	(0.0369)
Other employment status (ref.)			
Employed	0.1401**	0.1492**	0.1505**
	(0.0248)	(0.0206)	(0.0206)
Married (ref.)			
Separated	-0.4825**	-0.3395**	-0.3428**
	(0.0783)	(0.0754)	(0.0754)
Single	-0.1404**	-0.1896**	-0.1897**
	(0.0519)	(0.0693)	(0.0689)
Divorced	-0.2114**	-0.0119	-0.0147
	(0.0595)	(0.0665)	(0.0666)
Widows	-0.0922	-0.1983*	-0.2007*
	(0.0872)	(0.0931)	(0.0935)
Partner abroad	-0.8798**	-0.3356	-0.3340
	(0.2668)	(0.2946)	(0.2948)
Below 20 (ref.)			
20-24	-0.2212**	0.0183	0.0173
	(0.0476)	(0.0497)	(0.0497)
25-29	-0.4197**	0.0151	0.0143
	(0.0529)	(0.0589)	(0.0590)
30-34	-0.6402**	-0.0149	-0.0134
	(0.0582)	(0.0684)	(0.0685)
35-39	-0.6135**	0.1134	0.1141
	(0.0691)	(0.0829)	(0.0830)
40-44	-0.4634**	0.2997**	0.3009**
	(0.1015)	(0.1110)	(0.1112)
> 44	-0.7145**	0.3901**	0.3919**
	(0.1405)	(0.1481)	(0.1483)
Net (ln) household income	0.6962**	0.4093**	0.4104**
	(0.0282)	(0.0266)	(0.0265)
Number of children	-0.0557**	1.7126**	1.7122**
	(0.0167)	(0.3905)	(0.3906)
9 to 2 month before birth	0.6123**	0.4183**	0.4168**
	(0.1418)	(0.1073)	(0.1076)
2 to 0 month before birth	0.3596**	0.4025**	0.4035**
	(0.0935)	(0.0835)	(0.0838)
2 to 6 month after birth	0.2072*	0.1984*	0.2021*
	(0.1048)	(0.0941)	(0.0939)
6 to 10 month after birth	0.0671	0.0705	0.0708
	(0.1127)	(0.0966)	(0.0970)
10 to 12 month after birth	0.2470	0.0637	0.0760
	(0.1676)	(0.1463)	(0.1450)
12 to 15 month after birth	0.1082	0.0297	0.0271
	(0.1714)	(0.1479)	(0.1480)
15 to 18 month after birth	0.2796	0.0867	0.0912
	(0.1878)	(0.1531)	(0.1528)
18 to 24 month after birth	0.0314	-0.0162	-0.0183
	(0.1652)	(0.1434)	(0.1430)
18 to 36 month after birth	0.1290	0.0155	0.0131
	(0.1222)	(0.1078)	(0.1077)
Years of education	0.0284**	-0.0116	-0.0117
	(0.0059)	(0.0109)	(0.0109)

Table 1. **Estimation results** *(cont.)*

	OLS	FE	FE-IV
		Lifesat	
Constant	2.0957**	1.2840	1.6009*
	(0.2274)	(0.8207)	(0.7827)
Time	Dummies	Dummies	Trends
Länder effects	Yes	Yes	Yes
Observations	81 828	81 828	81 828
Number of never changing person ID		10 277	10 277

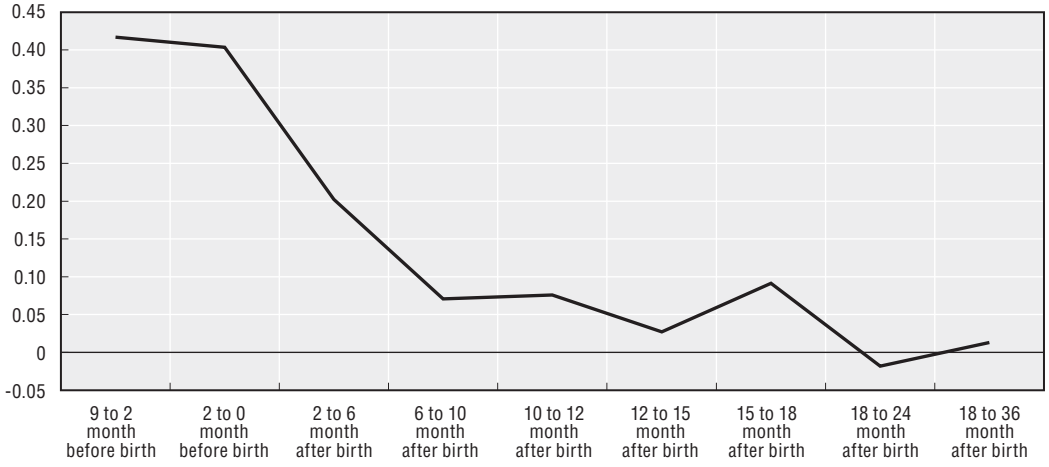
Robust standard errors in parentheses.
* Significant at 5%; ** significant at 1%.

widowers. Education does not significantly affect life satisfaction, but its effect is probably captured by (the natural log of) net household income. The *Länder* dummies indicate significantly life satisfaction compared with Bayern (the reference category), in particular living in Bremen gives the largest satisfaction premium (13%) compared with Bayern.

The coefficient estimates associated with the control variables related to marital status and income are in line with those found in the literature on life satisfaction. Similarly, the estimates associated with the age coefficient show the expected relation with life satisfaction found elsewhere in the literature.

The analysis of the leads and lags around childbirth confirm some of the evidence reported in the literature for the baby honeymoon (Figure 8): the largest positive effect on life satisfaction comes before and just after birth – with a size larger than more than twice the positive effect of being employed. However, this is a very temporary effect, which disappears as the infant grows older. After another two months, the effect is more than halved, and by about one year of age, any positive effect on life satisfaction related to child-age has gone. Moreover, the effect is only significant up to six months after birth. One might interpret this finding as some mild evidence in favour of a policy of BRL for a period of up to six months, if the policy goal is women's life satisfaction.

Figure 8. **The effects of child age and being on BRL and child age on life satisfaction in Germany**



Source: Based on column 3 of Table 1.

Table 2 reports the results of the difference-in-differences estimation, conditional on a range of covariates as in the previous models. These are the conditional double difference, i.e. the difference in life satisfaction before and after the policy changes between the treated and the untreated. All the estimates are positive suggesting a potential welfare effect of BRL policy. However, only four policy changes seem to have a significant effect on life satisfaction. The 1986 policy introducing eight months of paid parental leave, the 1992 and 1993 policy changes extending the duration of unpaid and paid parental leave and the 2001 introduction of flexibility in taking parental leave all had a statistically significant effect on life satisfaction. The effect tends also to become more positive over time.[13]

Table 2. **Estimation results, Germany, BRL and life satisfaction (DiD)**

Policy change	Difference-in-difference
1986 introduction of 8 months of paid parental leave	**0.353**
1988 rise in paid parental leave from 8 to 10 months	0.080
1989 rise in paid parental leave from 10 to 13 months	0.046
1990 rise in paid parental leave from 13 to 16 months	0.055
1992 rise in unpaid parental leave to three years	**0.185**
1993 rise in paid parental leave from 16 to 22 months	**0.195**
2001 introduction of flexibility in taking parental leave	**0.199**
2007 introduction of earnings related paid parental leave	0.132

Note: Figures in bold are statistically significant at 5%.

7.2. Results from the British Household Panel Survey

Table 3 presents the results from the regression models using FE and FE-IV to estimate the effect of being on BRL on life satisfaction of women interviewed in the BHPS. The specifications include a set of standard controls, as similar as possible to those used in the analysis of German data and individual fixed effects as well. The results highlight patterns similar to those found in the literature on subjective well-being.

In terms of the first column, being on BRL has a statistically significant impact on life satisfaction. Its magnitude is one-quarter of a life satisfaction point, or an effect size of 0.2 of a standard deviation. The effect obtained with FE in the United Kingdom is very close to the comparable German statistic (Table 1). The fact of being employed and going to work raises life satisfaction by 0.16 of a life satisfaction point (effect size of about 0.1).

The BHPS does not contain questions on net income. Income is calculated separately, as described in Bardasi et al. (1999). The release of the income measures always lags a couple of years behind the data releases, which means that net income data are currently only available for the rounds 1-16. In terms of the results, the log of household income has a statistically significant though economically small impact on life satisfaction: a 10% increase in income only increases life satisfaction by 0.005 points on a 7-point scale.

Even though fixed effect are being controlled for, there is still a concern that there are unobserved time-varying factors which are influencing both the decision of going on parental leave and life satisfaction. Therefore, column 2 of Table 2 shows a regression model in which the BRL dummy is instrumented with the dummies for policy change 2, policy change 3, and policy change 4. Note that the first and last policy changes cannot be used since life satisfaction is only measured from 1996 onwards, and income data is not available

Table 3. **Subjective well-being and BRL in the *British Household Panel Survey* (FE and IV)**

Variables	FE	FE IV
	Life satisfaction	
	(1)	(2)
Treatment effect	0.2426***	1.0042
	(0.052)	(4.872)
Employed	0.0794***	0.1339
	(0.016)	(0.349)
Youngest child < 1 year	0.0514*	-0.0726
	(0.026)	(0.764)
Youngest child 1-2 years	-0.0054	0.0071
	(0.026)	(0.071)
Log Household income	0.0511***	0.0405
	(0.011)	(0.065)
Education – high	0.0283	0.0241
	(0.057)	(0.061)
Education – medium	-0.0058	-0.0026
	(0.057)	(0.062)
Married	0.0373	0.0247
	(0.033)	(0.076)
Widowed	-0.2911***	-0.3025***
	(0.054)	(0.078)
Divorced	-0.0244	-0.0320
	(0.043)	(0.059)
Disabled	-0.1962***	-0.2026***
	(0.024)	(0.029)
Time dummies	Yes	Yes
Time trend	No	No
Age dummies	Yes	Yes
Observations	59 577	59 577

yet for the years 2007 and 2008. This specification includes fixed effects as well. The policy change dummies are, however, collinear with time dummies. It is thus necessary to assume that well-being evolves smoothly over time and the time dummies are replaced with a time trend as previously. This specification does not show us a significant relationship with education, and widowed individuals are significantly unhappier than the others. Being disabled is also negatively related to subjective well-being.

Having a child of less than one year of age has a positive and statistically significant though small impact on one's life satisfaction. The coefficient on its lag is negative, though, which suggests adaptation. However, the lag is not statistically significant.[14]

Results from the IV regressions are rather disappointing. The parental leave dummy has an ill-defined, very large though statistically insignificant, coefficient in column 2 which may be related to the weakness of the instruments used – as indicated by the Stock-Yogo tests.

Given the poor performance of the IV-FE specifications, a treatment-effect regression model was also estimated where the (endogenous) variable is being on BRL. Policy changes 2, 3, and 4 were used as exclusion restrictions to identify the effect of interest in the life satisfaction regression. The coefficient estimate of being on BRL is positive (0.69) and statistically significant, suggesting again an increasing effect of BRL on life satisfaction.

8. Conclusions

The theoretical and empirical academic literature considering subjective measures of well-being and their socio-economic determinants has increased substantially in recent decades. Concurrently to this burgeoning academic literature, a significant amount of data on subjective well-being has been collected.

The academic literature is beginning to have an influence in the public and policy arenas. Academics have begun writing article and books addressing policy issues and implications (for example, Layard, 2005; Diener et al., 2008; Greve [ed.], 2010). The Stiglitz-Sen-Fitoussi report, commissioned by the French government, has added further impetus to this movement. Both international and national reports were increasingly drawing on subjective measures as part of an indicator dashboard to describe the evolution of social outcomes, both between countries and across time (e.g. OECD, 2010; OECD, 2011b, c).[15]

This growth has contributed substantially to the knowledge of how to measure subjective well-being, the relationship between subjective measures and traditional economic indicators like GDP, and the background factors such as family life, health and income which may influence an individual's subjective well-being.

Despite this impetus a central issue remains. It is about how to use these tools to evaluate the welfare effects of social policies. A larger body of empirical work which directly addresses policy choices and their impact on subjective well-being needs to be built. This is even more important to shed light on the main mechanisms behind welfare effects in the absence of randomised experiments.

The aim of this paper was to fill some of these gaps by addressing the following question: Does being on BRL enhance life satisfaction? In doing so, the paper uses policy changes as the assignment rule into treatment and control groups, respectively, to build a quasi-experiment.

Differently from Pezzini (2005) the results from a variety of methods suggest that BRL policies tend to have a significant positive impact on life satisfaction of women eligible for this kind of policy both in Germany and in the United Kingdom. Differences between the findings of the present paper and the studies of Pezzini may be related to a variety of factors.

In particular, quality and precision of the data collected in the GSOEP and BHPS are bigger compared with Eurobarometer. The data collected in the GSOEP and BHPS are representative of populations in Germany and the United Kingdom. Sample sizes are much smaller in Eurobarometer. BHPS and GSOEP collect individual panel data and thus information on the same individuals over time which implies a greater stability over time than the Eurobarometer, both with respect to the question asked and the people surveyed which affect once again the robustness and validity of the data.

The result we find in our paper is robust to using a variety of approaches. Overall, the results for the United Kingdom offer cautious support for the German results. The effect of BRL on life satisfaction is relatively large, very close and even higher in the case of the United Kingdom to the effect on life satisfaction deriving from being in employment. It may well be the combination of time off *and* the right to return to work that result in higher life satisfaction. The effect seems to be causal, running from leave to satisfaction.

This result is useful to inform the debate of welfare effects of social policy. However, it is also important to highlight that that long leave periods tend to reduce women's attachment to the paid labour market. The key issue then becomes the optimal design of

BRL, a design which would allow the combination of high level of individuals' life satisfaction with a high probability of returning to work. For example, the analysis of German data suggests that life satisfaction for women on BRL is highest just before and up to six months after childbirth. Other studies and more precise data on BRL are necessary to see if this effect is robust over time and across countries.

The analysis presented here has also some limitations which should be acknowledged. *First*, the benefits of leave in terms of subjective well-being are likely to be shared across the family unit. The inability to assess the impact of maternity leave on fathers' well-being is a major issue. For example, data from the German Federal Statistics Office show that the proportion of fathers taking leave has increased substantially between 2006 and 2009. Future waves of the GSOEP might help to shed some light on the benefit of the 2007 reforms also on fathers.

Second, there are distributional aspects which the paper cannot address because of the lack of a natural experiment. Those who experience the positive effect on life satisfaction are likely to be different from those who do not. The methods used in the paper represent one possible attempt to deal with this issue. However enhanced data on the relevant outcomes might shed further light on these issues.

Notes

1. The definition of birth-related leave used in the paper covers both maternity and parental leave. See *www.oecd.org/els/soc/PF2.1_Parental_leave_systems%20-%20updated%20%2018_July_2012.pdf*.
2. In this paper, we often refer to "treated" and "untreated" (or "control"). These terms generally identify two different groups of individuals that have and have not, respectively, participated in a given experiment. Broadly speaking the group which is treated has undergone the treatment, while the control has not. Pure experimental methods allow identification of the effect of a given treatment by eliminating all "confounding" factors and, in particular, the endogeneity bias. The literature has extended these concepts to quasi-experimental methods (or natural experiments) which also comprise policy changes designed in such a way that individuals cannot select themselves into or out of the measure. In that case, the treated group is formed by those individuals that are affected by the policy change while the control group is formed by those individuals that are not affected by the policy change.
3. See also Burchardt (2006).
4. The measure of mental health was hospital admissions for mental health reasons.
5. Similarly there are very large variations in expenditures on BRL. See OECD (2011b) for details.
6. See for example Nielsen (2009).
7. It is worth noting that sampling/non-sampling errors may affect the construction of the treatment and control groups. For example, non-sampling errors are likely to be large in Eurobarometer because of the changes in the questionnaire and resulting impact on framing effects on responses to SWB questions.
8. We are indebted to one of the referees for this suggestion.
9. Correction of the selection bias with Heckman's treatment model is also possible and this approach will be used to estimate welfare effects of BRL in the United Kingdom.
10. Originally the analysis also included Belgium. However, for the small sample size, because of the different measure of subjective well-being recorded in the survey and of the unique policy change that occurred, we have decided to exclude it. Results are available for interested readers.
11. In 2011, the participation of fathers in paid parental leave benefits increased substantially to 27.3% but more than three-quarters of them requested benefits for no longer than two months. The reform of parental leave introduced in 2007 was explicitly aimed at increasing the take-up of leave by fathers. According to very recent data published by the Federal Statistics Office, the proportion

of fathers taking leave has increased steadily: from around 3% in 2006 to around 26% for children born in the third quarter of 2010 (see Blum and Erler, 2012).

12. Unfortunately, we cannot control for the generosity of leave benefits because of the unavailability of longitudinal data.

13. It is interesting to note that the average effect obtained with the DiD estimator is 0.152 which is not far from the one obtained with the IV estimator.

14. Ideally, as in the case of Germany, it would be desirable to break down these dummies in even finer sub-annual categories, to control for fluctuations of well-being during the months before and after giving birth – the so-called "baby honeymoon effects" examined with German data. The use of the additional information contained in the fertility calendar in the BHPS might allow undertaking such an exercise.

15. In 2011, for example, the OECD launched a new indicator (the Your Better Life Index) which allows comparison across countries in terms of average well-being along 11 dimensions. See *www.oecd.org/statistics/betterlifeinitiativeyourbetterlifeindex.htm*.

References

Baker, M. and K. Milligan (2005), "How Does Job-Protected Maternity Leave Affect Mothers' Employment and Infant Health", *NBER Working Papers*, No. 1135.

Baker, M. and K. Milligan (2008), "Evidence From Maternity Leave Expansions of the Impact of Maternal Care on Early Childhood Development", *NBER Working Papers*, No. 13826, February.

Baldi, S. and S. Chapple (2010), "A Panel Data Set of Changes in Birth-Related Leave Across the OECD, 1969-2008", mimeo.

Baldi, S., S. Chapple and J. Zapata (2011), "A Panel Data Set of Changes in Birth-Related Leave Across the OECD, 1969-2010", *OECD Social Employment and Migration Working Paper*.

Bardasi, E., S.P. Jenkins and J.A. Rigg (1999), "Documentation for Derived Current and Annual Net Household Income Variables, BHPS Waves 1-7", Institute for Social and Economic Research, Colchester: University of Essex (data available from the DataArchive, University of Essex.).

BayLErzGG (1989), "Gesetz zur Gewährung eines Landeserziehungsgeldes und zur Ausführung des Bundeserziehungsgeldgesetzes".

BayLErzGG (1995), "Gesetz zur Gewährung eines Landeserziehungsgeldes und zur Ausführung des Bundeserziehungsgeldgesetzes".

BayLErzGG (2001), "Bayrisches Gesetz zur Zahlung eines Landeserziehungsgeldes und zur Ausführung des Bundeserziehungsgeldgesetzes".

BayLErzGG (2007), "Gesetz zur Neuordnung des Bayrischen Landeserziehungsgeldes".

Bird, K. (2001), "Parental Leave in Germany – An Institution With Two Faces?", *in* Leisering, Lutz/Müller, Rainer and Schuhmann, Karl F. (ed.) *Institutionen und Lebensläufe im Wandel. Institutionelle Regulierungen von Lebensläufen*, Bonn, Juventa, pp. 55-90.

Burchardt, T. (2006), "Happiness and Social Policy: Barking Up the Right Tree in the Wrong Neck of the Woods", *Social Policy Review*, Vol. 18, June, pp. 145-164.

Brodeur, A. and M. Connolly (2013), "Do Higher Child Care Subsidies Improve Parental Well-Being? Evidence from Québec's Family Policies", *Journal of Economic Behavior and Organization,* 93, September, pp. 1-16.

Callender, C., N. Millward, S. Lissenburgh and J. Forth (1997), "Maternity Rights and Benefits in Britain, 1996", *Department for Social Security Research Report,* No. 67, The Stationery Office, Norwich.

Caneiro, P., K. Loken and K. Salvanes (2010), "A Flying Start? Maternity Leave and Long-Term Outcomes for Mother and Child", *www.econ.yale.edu/seminars/labor/lap08/carneiro-081031.pdf*, April 2010, version supplied by Katrine Loken.

Chatterji, P. and S. Markowitz (2004), "Does the Length of Maternity Leave Affect Maternal Health?", *NBER Working Papers*, No. 10206, January.

Chatterji, P. and S. Markowitz (2008), "Family Leave After Childbirth and the Health of New Mothers", *NBER Working Papers*, No. 14156, July.

Clark, A and A. Oswald (2002) "A Simple Statistical Model for Measuring How Life Events Affect Happiness", *International Journal of Epidemiology*, 31, December, pp. 1139-1144.

Clark, A., E. Diener, Y. Georgellis and R. Lucas (2008), "Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis", *The Economic Journal*, Vol. 118, June, pp. F222-F245.

Deven, F. and P. Moss (eds.) (1999), *Parental Leave: Progress or Pitfall*, Vlamme Gemeenschap, Brussels.

Deven, F. and P. Moss (eds.) (2005), *Leave Policies and Research: Reviews and Country Notes*, Brussels, CBGS – WERKDOCUMENT 2005/3.

D-Statis (2012), "Parental Allowance: Employment of Women leading to Higher Participation of Fathers", *Press Release,* 221, 27 June.

Diener, E. (1984), "Subjective Well-Being", *Psychological Bulletin* 95(3), pp. 542-575.

Diener, E., P. Kesebir and R. Lucas (2008). "Benefits of Accounts of Well-Being – For Societies and for Psychological Science", *Applied Psychology: An International Review*, 57(1), pp. 37-53.

Dyrdal, G.M. and R.E. Lucas (2013), "Reaction and Adaptation to the Birth of a Child: A Couple-Level Analysis", *Developmental Psychology*, Vol. 49(4), April, pp. 749-761, *http://dx.doi.org/10.1787/a0028335*.

Dustmann, C. and U. Schönberg (2008), "The Effect of Expansions in Maternity Leave Coverage on Children's Long-Term Outcomes", *CSGE Research Paper*, Anglo-German Foundation, July.

Escobedo, A. (2008), "Leave Policies and Public Systems of Care for Children under Three Years Old and their Families in the European Union", Bellaterra: Universitat Autònoma de Barcelona Departament de Sociologia, Facultat de Ciències Polítiques i de Sociologia.

Evenson, R.J. and R.W. Simon (2005), "Clarifying the Relationship Between Parenthood and Depression", *Journal of Health and Social Behaviour*, Vol. 46, December, pp. 341-358.

Ferrer-i-Carbonell, A. and P. Frijters (2004), "How Important Was Methodology for the Estimates of the Determinants of Happiness?", *Economic Journal*, Vol. 114, pp. 641-659.

Fonda, N (1980), "Statutory Maternity Leave in the UK: A Case Study", *in* P. Moss and N. Fonda (eds.), *Work and the Family*, London: Temple Smith.

Frey, B.S. and A. Stutzer (2003), "Reported Subjective Well-Being: A Challenge for Economic Theory and Economic Policy", *CREMA Working Papers*, No. 2003-7.

Gauthier, A.H. and A. Bortnick (2001), *Comparative Maternity, Parental, and Childcare Database, Version 2*, University of Calgary, *http://soci.ucalgary.ca/fypp/family_policy_databases.htm*.

Gjerdingen, D. and K. Chaloner (1994), "The Relationship of Women's Postpartum Mental Health to Employment, Childbirth, and Social Support", *Journal of Family Practice*, Vol. 38, pp. 465-472.

Gjerdingen, D., D. Froberg, K. Chaloner and P. McGovern (1993), "Changes in Women's Physical Health During the First Postpartum Year", *Archives of Family Medicine*, Vol. 2, pp. 277-83.

Gjerdingen, D., P. McGovern, K. Chaloner and H. Street (1995), "Women's Postpartum Maternity Benefits and Work Experience", *Family Medicine*, pp. 592-8.

Glenn, N.D. and S. McLanahan (1982), "The Effects of Children on the Psychological Well-Being of Older Adults", *Journal of Marriage and the Family*, 43, pp. 409-421.

Gregg, P., M. Gutierrez-Domenech and J. Waldfogel (2007), "The Employment of Married Mothers in Great Britain, 1974-2000", *Economica*, 74, pp. 842-864.

Greve, B. (ed.) (2010), *Happiness and Social Policy in Europe*. Cheltenham, Edward Elgar.

Hyde, J., M. Klein, M. Essex and R. Clark (1995), "Maternity Leave and Women's Mental Health", *Psychology of Women Quarterly,* Vol. 19, pp. 257-285.

Jordan, L. (1999), "Background Information on European and Canadian Parental Leave Laws", accessed 11 February 2009 at: *www.cga.state.ct.us/ps99/rpt/01r/htm/99-r-1214.htm*.

Kamerman, S. and A. Kahn (1991), *Child Care, Parental Leave and the Under 3s: Policy Innovation in Europe*, New York, Auburn House.

Khanam, R., H.S. Nghiem and L. Connelly (2009), "Does Maternity Leave Affect Child Health? Evidence from Parental Leave in Australia Survey", in 31st Australian Conference of Health Economists (AHES 2009), 1-2 October, Tasmania, Australia.

Lalive, R. and J. Zweimuller (2005), "Does Parental Leave Affect Fertility and Return-to-Work? Evidence from a 'True Natural Experiment'", *IZA Discussion Papers*, No. 1613.

Layard, R., (2005), *Happiness: Lessons from a New Science*, Penguin UK, 2005.

Liu, Q. and O. Skans (2009), "Paid Parental Leave, Mothers' Education and Children's Scholastic Performance", *IZA Discussion Papers*, No. 4244.

McGovern, P., B. Dowd, D. Gjerdingen, I. Moscovice, L. Kochevar and W. Lohman (1997), "Time Off Work and the Postpartum Health of Employed Women", *Medical Care*, Vol. 35, pp. 507-21.

Merz, M. (2004), "Women's Hours of Market Work in Germany: The Role of Parental Leave", *IZA Discussion Papers*, No. 1288.

Moss, P. and M. O'Brien (eds.) (2006), *International Review of Leave Policies and Related Research*, Department for Business Enterprise and Regulatory Reform, Employment Relations Research, Series No. 57.

Moss, P. and K. Wall (eds.) (2007), *International Review of Leave Policies and Related Research*, Department for Business Enterprise and Regulatory Reform Employment Relations Research, Series No. 80, July.

Moss, P. and M. Korintus (eds.) (2008), *International Review of Leave Policies and Related Research*, Department for Business Enterprise and Regulatory Reform, Employment Relations Research, Series No. 100, July.

Moss, P. (2012), *International Review of Leave Policies and Related Research 2012*, Institute of Education, University of London.

Nielsen, H. Skyt (2009), "Causes and Consequences of a Father's Child Leave: Evidence from a Reform of Leave Schemes", *School of Economics and Management Aarhus University Economics Working Papers*, No. 2009-8.

OECD (2010), *Subjective Well-Being and Social Policy*, first report to the European Union, September.

OECD (2011a), *Subjective Well-Being and Social Policy*, second report to the European Union, June.

OECD (2011b), *Doing Better for Families*, OECD Publishing.

OECD (2011c), *How's Life? Measuring Well-Being*, OECD Publishing.

Pezzini, S. (2005), "The Effect of Women's Rights on Women's Welfare: Evidence from a Natural Experiment", *The Economic Journal*, 115, March, pp. C208-C227.

Rasmussen, A.W. (2010), "Increasing the Length of Parents' Birth-Related Leave: The Effect on Children's Long-Term Educational Outcomes", *Labour Economics 17* (1), pp. 91-100.

Ruhm, C. (2000), "Parental Leave and Child Health", *Journal of Health Economics*, 19 (6), pp. 931-960.

SächsLErzGG (1992), "Gesetz über die Gewährung von Landeserziehungsgeld im Freistaat Sachsen und zur Änderung des Gesetzes über die Kindertageseinrichtungen".

SächsLErzGG (1996), "Neufassung des Gesetzes über die Gewährung von Landeserziehungsgeld im Freistaat Sachsen".

SächsLErzGG (2001), "Neufassung des Sächsischen Langeserziehungsgeldgesetzes".

SächsLErzGG (2008), "Gesetz über die Gewährung von Landeserziehungsgeldim Freistaat Sachsen", (Sächsisches Landeserziehungsgeldgesetz – SächsLErzGG). *http://amt24.sachsen.de/ZFinder/verfahren.do?action=showdetail&modul=VB&id=1422!0*.

Sargeant, M. and D. Lewis (2008), *Employment Law*, Longman, 3rd edition.

Staehelin, K., P. Bertea and E. Stutz (2007), "Length of Maternity Leave and Health of Mother and Child – A Review", *International Journal of Public Health*, 52, pp. 202-209.

Stiglitz, J., A. Sen and J.-P. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, *www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf*.

Tanaka, S. (2005), "Parental Leave and Child Health across OECD Countries", *The Economic Journal*, Vol. 115, pp. F7-F28.

Umberson, D. and K. Williams (1999), "Family Status and Mental Health", *in* C. Aneshensel and J. C.Phelan (eds.), *Handbook of the Sociology of Mental Health*, New York: KluwerAcademic/Plenum, pp. 225-254.

Wooldridge, J. (2007), *What's New in Econometrics?*, NBER Summer Institute 2007.

Würtz, A. (2007), "The Long-Term Effect on Children of Increasing the Length of Parents Birth-Related Leave", *Aarhus School of Business, Department of Economics Working Papers*, No. 07-11.

Zabel, C. (2009), "Eligibility for Maternity Leave and First Birth Timing", *Population Research and Policy Review*, Vol. 28, pp. 251-270.

Zarrabi, M.(2009), "Short- and Medium-Term Effects of Maternity Leave on Child Outcomes: Evidence from the NLSCY", May, downloaded on 2 May 2010 from *http://economics.uwaterloo.ca/chesg2009/CHESG 2009Program_000.htm*.

ZBFS (2010), "Bayrisches Landeserziehungsgeld", *www.zbfs.bayern.de/imperia/md/content/blvf/erziehungs geld/infolandeserziehungsgeld0110.pdf*.

APPENDIX A

Eurobarometer survey detail

Eurobarometer	Year	Fieldwork months	Question number
ECS73	1973	September, October	Q22
3	1975	May	Q61
4	1975	October	Q68
5	1976	May, June	Q149
6	1976	November	Q112
7	1977	April, May	Q112
8	1977	October, November	Q112
9	1978	May, June	Q112
10	1978	October, November	Q112
11	1979	April	Q124
13	1980	April	Q112
15	1981	April	Q120
17	1982	March, April	Q142
18	1982	October	Q232
19	1983	March, April	Q121
20	1983	October	Q238
21	1984	March, April	Q212
22	1984	October, November	Q226
23	1985	March, April	Q132
24	1985	October, November	Q155
25	1986	March, April	Q122
26	1986	October, November	Q145
27	1987	April	Q231
28	1987	October, November	Q135
29	1988	March, April	Q125
31	1989	March, April	Q124
31A	1989	July	Q123
32	1989	October, November	Q3
33	1990	March, April	Q2
34.0	1990	October, November	Q7
34.1	1990	October, November	Q2
34.2	1990	October, November	Q2
35.0	1991	March, April	Q2
36.0	1991	October, November	Q2
37.0	1992	March, April	Q2
37.1	1992	April, May	Q2
37.2	1992	April, May	Q2
37.0+1	1992	March, April, May	Q2
38.0	1992	September, October	Q2
38.1	1992	November	Q2

Eurobarometer	Year	Fieldwork months	Question number
39.0	1993	March, April	Q2
40.0	1993	October, November	Q2
41.0	1994	April, May	Q2
42.0	1994	November, December	Q2
43.1	1995	April, May	Q2
44.2bis	1996	February	Q2
47.1	1997	March, April	Q36
49	1998	April, May	Q5
52.0	1999	October, November	Q7
53	2000	April, May	Q4
54.1	2000	November, December	Q4
55.1	2001	April, May	Q7
56.1	2001	September, October	Q46
56.2	2001	October, November	Q4
57.1	2002	March, April, May	Q4
58.1	2002	October, November	Q4
60.1	2003	October, November	Q4
62.0	2004	October, November	Q4
63.4	2005	May, June	QA3
64.2	2005	October, November	QA3
65.2	2006	March, April, May	QA3
66.1	2006	September, October	QA3
67.2	2007	April, May	QA3
68.1	2007	September, October, November	QA3
69.2	2008	March, April, May	QA3
70.1	2008	November, December	QA3
71.1	2009	January, February	QA2
71.2	2009	May, June	QA1
71.3	2009	June, July	QA1
72.4	2009	October, November	QA1

Source: www.gesis.org/en/services/data/survey-data/eurobarometer-data-service/eb-trends-trend-files/list-of-trends/life-satisf.

APPENDIX B

Birth-related provision in the German Länder

A number of *Länder* – seven out of the 16, which covers approximately half the German population – have additional payments (*Landeserziehungsgeld*) which effectively extend the paid leave provided at a federal level. The *Länder* providing these benefits include Baden-Württemberg, Bavaria, Berlin, Mecklenburg-Vorpommern, Rhineland-Palatinate, Saxony and Thuringia. In most cases, these appear to be employment-conditioned young child benefits. See Map 1 for details of *Länder* locations, their absolute and relative populations (share of the current German population) and a summary of the policy shifts.

Baden-Württemberg

At the same time as the federal reform of 1 January 1986 was introduced, Baden-Württemburg introduced a DM 400 (EUR 205) per month payment lasting for one year, beginning after the child reached 10 months of age, when federal payments for maternity and parental leave ran out (RL-LErzG, 1986).

Conditions include a Länder residency condition, being an EU national, not being in full-time employment and being under a relatively low monthly net family income ceiling (less than DM 2 000 until 1/1/1994; DM 2 200 until 1/1/1996; DM 2 450 until 1/1/2001, and from 2001 EUR 1 380 and EUR 1 125 for single parents). The income limit rises by DM 300 (EUR 230 since 2001) for every additional child (RL-LErzG, 1986; 1988; 1995; 2001).

From 3 July 1995, the one year of payment commenced from the child's second birthday up until the third birthday (RL-LErzG, 1995). Thus people living in Baden-Württemburg were covered up to three years for the unpaid but job-protected leave introduced by the federal government in 1992. From the 1 January 2001 federal reform, the additional year of paid *Länder* leave could be used up until the child was eight years old. Additionally, a higher payment was introduced for the third child: The first two children receive EUR 205 a month, the third and additional child EUR 307 per month. On 1 January 2007, the payment for the third child was cut to EUR 240 (VwV-LErzG Mehrlinge, 2007).

Bayern (Bavaria)

On 1 July 1989, Bayern introduced a monthly parental leave payment of DM 500 (EUR 256) for six additional months after the legal end of the federal parental payment (13 months at that point) (BayLErzGG, 1989). On 1 July 1993, this was further extended to one year's payment for children born from 8 December 1994 onwards (BayLErzGG, 1995). Thus parents could be covered for the entire period of the three years job-protected leave, created federally in 1992 (BayLErzGG, 1995). From 1 January 2001, parents received EUR 307

Map 1. **Länder population, population shares and employment-conditioned young child benefits**



instead of EUR 256 for the third and any higher order children. Children born from 1 July 2002 receive EUR 200 per month for six months for a first child, EUR 250 for 12 months for a second child and EUR 350 for 12 months for the third or higher order children (BayLErzGG, 2001). From 1 January 2007, these payment rates were cut by EUR 50 to EUR 150, EUR 200 and EUR 300, respectively, (BayLErzGG, 2007).

Conditions include a *Länder* residency condition, being an EU national, and not being in full-time employment. Families were required to be under a relatively low annual income ceiling, similar to that calculated by annualising the monthly threshold in Baden-Württemburg (Annual net family income for a couple not exceeding EUR 15 050 and for a single parent EUR 12 100. Each additional child increased the annual limit by EUR 2 150).

Berlin

Berlin introduced a payment for children born from 1 January 1983. The Berlin system paid a universal base-payment of DM 3 000 (EUR 1 533) during the first year of a child's life. The payment increased to DM 4 800 (EUR 2 454) per child if the parent was not working

full-time (non-full-time bonus) (GesSozFam VB1, 1983). The payment was abolished from 1 January 1993.

Conditions included a *Länder* residency condition, and being an EU national. There was an annual gross income ceiling above which families were not eligible for payment of EUR 33 745, extended by EUR 3 375 for each additional child.

Mecklenburg-Vorpommern

From 1 July 1995, families in Mecklenburg-Vorpommern received EUR 307 per month for a year after the legal end of the federal *Erziehungeld* for children born after 31 December 1993 (LErzGG M-V, 1995). From 1 April 1997, the duration of the payment was reduced to six months (LErzGÄndG M-V, 1997). Eligibility was reduced to children with mothers in education and training courses from 25 June 1999 (LErzGG M-V, 1999), and finally abolished on 1 May 2005, with final payments for some families up until 14 October 2005 (HRG, 2004/05).

Conditions included a *Länder* residency condition, not being full-time employed and being an EU national. There was an annual gross income ceiling above which families were not eligible for payment of EUR 15 032 for married couples and EUR 12 118 for everybody else. The threshold was extended by EUR 2 147 for each additional child.

Rheinland-Palatinate

From 1 July 1984, EUR 102 per month was paid for six months for a third or higher order child. The rate increased to EUR 153 for the second six months. There was an income ceiling. From 1 January 1986, parents of three or more children received EUR 153 per month for 18 months after federal pay finished. The income limits were similar to the limits of the federal scheme. From 1 January 1993, the duration of the payment introduced in 1986 was reduced to 12 months because the federal Erziehungsgeld was extended to 24 month. Thus the maximum period of 36 months remained unchanged. The payment was abolished for children born after 15 April 1995.

Saxony

From 1 September 1992, a payment was introduced, of EUR 205 per month for six months duration following the end of parental leave if the child was born between 1 January 1992 and 31 December 1993, and for 12 months if born after 1 January 1994. There is no payment if the child is in a public kindergarten (SächsLErzGG, 1992). From 1 January 2001, payment duration fell to nine months (SächsLErzGG, 2001). The regular pay of EUR 205 increased to EUR 307 for third or higher order children if the child was the third or more or if the entitled person was a student and when the child was born after 31/12/1994 (SächsLErzGG, 1996).

From 25 November 2007, payments were EUR 200 for the first child and EUR 250 for the second, both for nine months, and EUR 300 for the third and higher order child for one year. The annual net income-ceiling for married persons changed to EUR 17 100 and for others EUR 14 100. Each additional child increased the limit by EUR 3 140 (SächsLErzGG, 2008).

Conditions included a *Länder* residency condition, not being full-time employed and being an EU national.

Thuringia

From 1 January 1994, a young child payment was introduced. The amount and condition of the pay are linked to the federal *Erziehungsgeld* legislation. For every change in federal legislation, there was one for Thurigia. Parents receive the payments for six months following the end of the federal payment at the same monthly payment rate. The income thresholds were also as with the federal system (Thür LErzG, 1993).

There was a further change from 3 February 2006. The payment period was extended. Parents received EUR 150 for 12 months for the first child, EUR 200 for the second, EUR 250 for the third and EUR 300 for the fourth and additional child. If the child attended a kindergarten, parents could claim up to EUR 150 of this payment (Thür LErzG, 2006). From 14 May 2010, payment ceased if a child was in a kindergarten for more than five hours per week (Thür KitaG, 2010).

OECD Journal: Economic Studies
Volume 2013
© OECD 2014

Grade repetition: A comparative study of academic and non-academic consequences

by
Miyako Ikeda *and* Emma García*

This article explores country-by-country differences in academic performance and attitudes towards school between students who repeated a grade in primary school, in secondary school or never repeated a grade. The analyses use PISA 2009 for 30 countries in which a relatively high proportion of students repeated a grade before the age of 15. The comparisons across countries and the examination of models of both academic and non-academic performance contribute to shed some light on the consequences of repeating a grade for students. The estimated associations suggest that in most countries examined, at the age of 15, students who repeated a grade in secondary school tend to perform better academically than do students who repeated a grade in primary school, but worse than non-repeaters. In terms of the measure of behavioural performance chosen for this analysis, attitudes towards school, in the majority of countries, non-repeaters tend to report more positive attitudes towards schools than primary and secondary-school repeaters, but the comparison between repeaters in primary and secondary schools shows less consistent patterns across countries. These differences are observed after accounting for background characteristics of the students and exploring some differential relationships between grade repetition and education outcomes according to student characteristics. The achievement and behavioural gaps among groups of repeaters may reflect differences in the development of academic and behavioural skills over the school years, as well as differences in the way these groups of students are treated across different educational systems.

JEL classification: I21, J13, H52

Keywords: Grade repetition, academic performance, attitudes towards school, PISA

* The authors are from the OECD Directorate for Education and Skills (*miyako.ikeda@oecd.org*) and the Teachers College of Columbia University (*meg2173@tc.columbia.edu*), respectively. They would like to thank Marion Goussé and Emily Farchy for their valuable comments and suggestions. The views expressed in this article are the authors' and are not necessarily shared by the OECD or its member countries.
The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

The empirical literature assessing the consequences of grade repetition provides inconclusive evidence of the academic and socio-emotional effects of retention on students. In fact, the existing empirical literature is divided into two camps (as summarised by Thompson and Cunningham, 2000). The first supports the contention that grade repetition is beneficial for the student. The second suggests that it is detrimental for the students. These mixed findings make the assessment of this educational intervention difficult and its utilisation by education policymakers controversial, even putting aside the recognition that grade repetition is costly for the systems (West, 2012). Despite attempts to explain the differences in findings as the result of the disparity of the contexts analysed – in terms of students' characteristics, educational system features, country differences, etc. – a great amount of confusion about retention policies still remains. The puzzle calls for further analysis to help disentangle the positive and negative effects of grade retention.

This paper explores whether the two opposite evaluations of grade retention – i.e. whether its effects are positive or negative, on average – can be better understood when analysing the following questions. First, this paper examines how students' performance in reading at the age of 15 differs according to whether and when they have repeated a grade. Second, the paper focuses on students' attitudes towards school and analyses the relationships between what 15-year-old students think about their schools and whether and when they have repeated a grade. This paper also examines the role played by students' individual characteristics in the impact of retention on educational outcomes (i.e. socio-economic background, gender, migrant backgrounds, and age). These characteristics are incorporated to explore how the relationships between grade repetition and students' performance and attitudes towards school vary according to their background characteristics. The analyses build on the idea that academic and socio-emotional developments of students evolve differently over the school years, and accordingly, retention at different stages of their educational career is likely to have different implications and meaning for students. Also, the relationships between repetition and performance/attitudes towards school may differ across countries, and the comparison of these relationships across countries may uncover relevant information for our understanding of the policy. For these purposes, the analyses of this paper are based on the OECD PISA 2009 dataset.

1. Literature review

As mentioned above, previous research on the consequences of retention has noted the dissimilar impacts of the policy and its economic costs for the educational systems. A set of papers that are observational or meta-analysis studies reinforces these findings. For example, Thompson and Cunningham (2000) and West (2012) summarise the results of the assessment of retention policies. On the one hand, the negative consequences of retention arise from the fact that it discourages students with low motivation, confidence and social promotion, and forces retained students to repeat the same curriculum while their advancing peers keep learning more advanced topics. On the other hand, retention signals

to the students no tolerance for weak efforts and low performance, and gives lagging students an opportunity to get serious and get ready for the next grade. Regarding the economic pressure imposed on educational systems, West (2012) estimates that the cost to society of retaining 2.3% of the total student population in the United States is higher than $12 billion per year (p. 3). Similar evidence for some other countries (OECD, 2011) reinforces the idea that financial resources devoted to repetition are substantial. This cross-country comparison of total annual costs – including direct costs and opportunity costs[1] – suggests that, relative to total expenditure on primary and secondary education, the costs of grade repetition may represent a fraction up to 10 to 12% in Belgium, Spain, the Netherlands and Portugal (and 5 to 10% in Brazil, Germany and Italy).

Rumberger and Lim (2008) review some empirical evidence on the relationship between retention and probability of graduating from high school. The majority of the studies reviewed suggest that retention in elementary and/or middle school was associated with an increase in the odds of dropping out of high school. High school repetition, however, did not seem to have any significant effects on dropout probability. In Colombia, over-age students, who are often grade repeaters, are more likely to drop out of school even at the very early stage of their educational career in primary school (Garcia-Jaramillo et al., 2011).

Xia and Kirby (2009) provide a detailed review of the literature that analyses the impact of retention in various educational outcomes (covering 91 publications). Although the paper notes the existence of mixed results when evaluating the effects of grade retention, the authors conclude that grade retention does not seem "to benefit" academic achievement (p. ix and x). On average, papers identifying positive effects on academic achievement also find that these have dissipated over time. Outcomes negatively associated with retention are the probabilities of graduating and pursuing post-secondary education, and the quality of employment (lower-paid jobs). In terms of non-academic outcomes, the summarised findings are mixed as well. The paper also provides a detailed overview of the characteristics of retained students. In general, retained students tend to have more disadvantaged socio-economic backgrounds, which are measured by proxy variables such as living in poverty, lower parental educational attainment and also lower measures of intelligence.

Allen, Chen, Willson and Hughes (2009) examine the effect of grade retention on academic outcomes. The authors review various studies using quasi-experimental designs published between 1990 and 2007 and conduct a meta-analysis of over 200 effect sizes, using multilevel analysis. The mean Hedges effect size is –0.11 (standard deviation was 0.01). They conclude that there is a negative effect of grade repetition on academic outcomes.

As noted by most of these observational papers, the estimated coefficients reported by some of the existing empirical evidence cannot be interpreted as the statistical impact of retention on achievement and other non-cognitive outcomes. Observational studies are affected by the selection of students in the policy. Consequently, they may represent both the effect of the policy and that of several unobserved variables that would also affect achievement. In the last decades, some papers have utilised quasi-experimental approaches to study the effect of grade repetition on performance, with particular emphasis on regression discontinuities, empirical strategies or instrumental variable estimation. These methods are expected to approximate the ideal conditions of a

randomised experiment, which is the "gold-standard" design for research to make unbiased causal claims (Murnane and Willett, 2011). Most of these strategies find ways to ensure exogenous treatment variability that helps overcome selection on observable and unobservable biases in the causal estimates of the effect of a treatment. For instance, Eide and Showalter (2001) model repetition as an endogenous family decision based on expected costs and benefits. Their paper shows that retention did not exert an independent influence on dropping out or earnings (for white males). They use the variation in age of entry into kindergarten across states as an instrument for grade repetition. The estimated coefficients for white males suggest that retention would lower the probability of dropping out and would increase earnings (for white males, the estimated coefficients associated with the impact of being retained on the log of earnings range from 0.178 to 0.316). The estimated coefficients for black males were statistically indistinguishable from zero.

Manacorda (2012) exploits the discontinuity induced by a regulation in Uruguay, according to which automatic grade failure is established for junior high school students who fail more than three subjects in the same academic year. Adjusting potential correlations between failure and other unobserved variables and reverse causality between failure and school outcomes, the paper shows that grade failure leads to substantial drop-out and lower educational attainment both in the immediate term and also four to five years after grade failure first occurred.

Jacob and Lefgren (2004) evaluate the impact on performance of attending summer school and being retained, as determined by the accountability policy in Chicago Public Schools. This policy was implemented in Chicago in 1996, and entitled students who scored below a certain threshold to summer school participation and potential grade retention. Using a (fuzzy) regression discontinuity design to overcome the effects of selection into the programme, this paper showed that retention increased achievement in mathematics and reading for third grade students, although it did not have any impact on math achievement and had a negative impact on reading achievement for sixth grade students. The estimated effects, reported in logits, suggest that retention in the third grade increased performance in reading by 0.18 and in math by 0.23 logits one year after retention. For students retained in sixth grade, the effect on academic performance one year later is not statistically significantly different from zero, while it is negative in reading performance two years later (estimated coefficient is 0.15 logits).

A more recent paper by Jacob and Lefgren (2009) provides interesting information about the impact of retention in different grade levels, using follow-up information from the previous paper and the same identification strategy. Their estimates suggest that grade retention among students in sixth grade did not affect the probability of completing high school. However, retention of students in eighth grade increased the probability of dropping out of high school, with a coefficient of -9.9 percentage points (the effect was stronger for some minorities, such as African-American girls).

An important paper in terms of its contributions to the hypotheses in our article is Schwerd and West (2012). This paper examines the long-term effects of retention in third grade in Florida. In particular, the study uses a (fuzzy) regression discontinuity to explore the probability of grade repetition that results from Florida's test-based promotion policy. This policy required that schools would retain students in third grade who failed to demonstrate basic proficiency on the state reading test (some exceptions applied). Using

the vertically scaled test scores in reading and math, the authors compare the achievement of students in the different grades. Thus, the counterfactual for students who were retained is those who had been promoted to the next grade. This between-grade comparison would avoid the traditional same-grade comparison bias, in which the test scores' comparisons cannot disentangle the effect of retention from the effect of being a year older at the time the test was administered.[2] This paper concludes that the effects (for students who comply with the policy, i.e. the local average treatment effect) of retention suggest large short-term gains in achievement, which fade out over the grades, and a statistically significant reduction in the probability of retention in subsequent years. The magnitude of the estimated effects is as follows: after one year, retained students outperformed their non-retained peers by 0.61 standard deviations in reading (0.46 standard deviations in math). The results were 0.42 standard deviations in reading after two years (0.22 standard deviations in math); 0.25 standard deviations in reading after three years (0.20 standard deviations in math); and 0.14 standard deviations in reading after four years (statistically insignificant in math). Regarding the effects of retention on the reduction in the probability of being retained in later grades, the estimated effects are 11% less probability of being retained two years later, 3% less probability of being retained three years later; and 4% less probability of being retained four or five years later. The authors mention that the effects of retention on achievement and educational attainment could vary according to the grade at which the students are retained. They mention that retention in earlier grades could be beneficial to students compared with later retention, although the main argument behind this statement seems to be the fact that earlier retention would reduce the probability of being retained in later years (plus a reference to the papers by Jacob and Lefgren, 2004 and 2009, summarised earlier).[3] Importantly, as acknowledged by the authors, these estimates do not distinguish which part of the effect on performance and later retention is a result of the policy *per se*, and which part is due to some complementary interventions. Florida's test-based promotion policy also provided retained students with other measures such as the opportunity to attend a summer reading programme before the beginning of the next academic year; they were also assigned to a high-performing teacher, and received intensive reading support during the year.[4]

Very few publications have analysed retention policies from a cross-country perspective, from a microeconomic perspective and using quantitative methods.[5] But, there are some exceptions. OECD (2011) provides a comparative international description of retention across countries. This report uses the information from PISA 2009 and identifies some of the associations between retention, student characteristics (socio-economic background) and the educational system (ability of schools to transfer students based on performance, behavioural problems or special learning needs). In terms of the proportion of students who had repeated at least one grade, the report situated the PISA 2009 OECD average at 13%. The between-country variation was high, running from 0% of students retained in Japan, Korea and Norway, to over 25% in countries such as Belgium, France, Luxembourg, the Netherlands, Portugal and Spain (and other partner countries around the world, with Macao-China, Tunisia and Brazil being the countries with the highest retention rates of over 40%, as shown later in Figure 1). This study suggests that in countries with a higher proportion of retained students, overall performance tends to be lower and social background has a stronger impact on learning outcomes than is the case in countries where fewer students are retained – regardless of the country's wealth. This

study also shows that about 15% of the variation in performance across countries was explained by differences in grade repetition rates.

A very recent work by Goos, Schreier, Eduard-Knipprath, De Fraine, Van Damme, and Trautwein (2013) examines the predictive value of different national educational policy characteristics in explaining the probability of students' repeating a grade. This paper, like OECD (2011) and our current research, builds on the PISA 2009 data. The authors specify a three-level model for the likelihood of being retained in primary or secondary school as a function of individual, school and country characteristics for 34 countries (OECD's *Education at a Glance* indicators and Teaching and Learning International Survey [TALIS] indicators are also used). Their findings suggest that a fifth to a quarter of the variation in the likelihood of being retained lay at the country level, and that some specific national educational factors[6] are statistically significant predictors of such a probability (p. 76).

2. Data

The analyses in this paper are based on the data of the *OECD Programme for International Student Assessment* (PISA). PISA has been assessing key competencies of 15-year-old students since 2000. The survey has been conducted every three years to assess students' competencies to apply their knowledge in real life contexts. The PISA 2009 dataset is the latest PISA dataset available to the public at the time of analysis for this paper. The data were collected in 2009 from around 470 000 students in 34 OECD countries and 31 non-OECD countries and economies. Each participating student was given two hours to take a pencil-and-paper cognitive test, and about 30 minutes to complete a questionnaire. Every PISA assessment covers three main subjects – reading, mathematics and science – with emphasis on one of these three subjects. As the main focus of PISA 2009 was reading, the analyses of this paper are based on student performance in reading.

2.1. Grade repetition

In the PISA 2009 Student Questionnaire,[7] students were asked whether they have repeated a grade at the primary education level, at the lower secondary education level, or at the upper secondary education level. The response categories were "No, never", "Yes, once" or "Yes, twice or more" for each of these three education levels.

Table A.1 in Appendix A shows the percentage of students who reported having repeated a grade at each education level. Table A.1 summarises students' responses to the three questions related to grade repetition (i.e. grade repetition at the primary, low-secondary, and upper-secondary levels): students who have never repeated a grade, students who have repeated a grade(s) at the primary education level only, students who have repeated a grade(s) at the secondary education level only, and students who have repeated grades at both the primary and secondary education levels. These are the descriptive statistics of the grade repetition variables that we use for the empirical estimation.

In some countries such as Japan, Korea and Norway, no 15-year-old students have repeated a grade. In about a half of the PISA 2009 participating countries, only one out of ten students or fewer has repeated a grade. Therefore, this paper focuses on the countries and economies where over 10% of 15-year-old students have ever repeated a grade. These include 16 OECD countries (Austria, Belgium, Chile, France, Germany, Hungary, Ireland, Italy, Luxembourg, Mexico, the Netherlands, Portugal, Spain, Switzerland, Turkey and the

United States) and 14 non-OECD countries and economies (Argentina, Brazil, Colombia, Dubai (UAE), Hong Kong-China, Indonesia, Latvia, Macao-China, Panama, Peru, Qatar, Trinidad and Tobago, Tunisia and Uruguay).

2.2. Two outcome variables

This article examines the following two outcome variables: student achievement in reading and the index of students' attitudes towards school. These variables are used as prime measures of a variable capturing the cognitive performance and a non-cognitive indicator of the student, linkable to the student's motivation or utility of school in his/her life. Table A.2 shows the descriptive statistics of countries' average performance. The metric for this reading scale is based on a mean for OECD countries set at 500, with a standard deviation of 100, in PISA 2000 when reading was the main focus for the first time.[8] In the PISA 2009 dataset, five plausible values (PVs) are included for reading achievement (the variables are ***PV1READ, PV2READ, PV3READ, PV4READ*** and ***PV5READ***). These five PVs are used for the analyses in this paper. The detailed use of PVs in the analysis is described in the appendix.

Table A.3 shows the descriptive statistics of the index of attitudes toward school (the variable name is ***ATSCHL***). This index was derived from students' responses (i.e. strongly agree, agree, disagree, or strongly disagree) to the following four statements: school has done little to prepare me for adult life when I leave school; school has been a waste of time; school has helped give me confidence to make decisions; and school has taught me things which could be useful in a job.

As explained in the PISA 2009 Technical report (OECD, 2012), students' responses to these items were scaled using Item Response Theory. Warm Likelihood Estimates (WLEs) in logits were computed for the latent construct of individual students. Then, WLEs were transformed to a scale with an OECD average of 0 and a standard deviation of 1. The negatively phrased questions were inverted before scaling so that higher values on this index indicate more positive attitudes towards school. For further details, please refer to PISA 2009 Technical Report (OECD, 2012).

2.3. Student characteristics

Students' gender, their migration status, their socio-economic backgrounds, and their age are examined in this article.

In the PISA 2009 Student Questionnaire, students were asked whether they are female or male. Based on this question, a dichotomous variable "**MALE**" (1 = boys and 0 = girls) is utilised in this paper. Table A.4 presents the proportion of boys and girls. The percentages of students who are boys range from 46% in Argentina to 53% in Hong Kong-China.

Students were asked the country of birth of themselves and their parents. This information is summarised and the index of immigrant background (IMMIG) is included in the *PISA 2009 Database*. This index has three categories: 1) native students (those students born in the country of assessment with at least one parent born in that country; and students who were born abroad with at least one parent born in the country of assessment); 2) second-generation students (those born in the country of assessment but whose parents were born in another country); and 3) first-generation students (those born outside the country of assessment and whose parents were also born in another country). Students were also asked to indicate the language they usually speak at home. Based on

this question, a variable of language at home (ST19Q01) is developed and included in the PISA 2009 database: 1) language at home is the same as the language of assessment; and 2) language at home is different language than the language of assessment. Based on these two variables in the PISA 2009 database, a dichotomous variable "***MIGLAN***" is developed for the analyses in this paper. MIGLAN is 1 when either IMMIG is 2 or 3, or ST19Q01 is 2; and MIGLAN is 0 for all other cases. In other words, MIGLAN is 1 when a student has a migrant background (either first- or second-generation) or speaks a language at home that is different from the language of assessment (either as a result of being an immigrant to the country or as a consequence of the coexistence of various languages in a given country). The percentages of students with migrant backgrounds or with a language at home that is different from the language of assessment range from 0.4% in Tunisia to 94% in Luxembourg.

Students' socio-economic backgrounds are measured with the index of economic, social and cultural status (***ESCS***) included in the PISA 2009 database. This index was derived from students' responses to questions about parents' occupational status, parents' education level, and home possessions.[9] The values on this index have an OECD mean of 0 and a standard deviation of 1. As seen in Table A.4, the country averages of the index of economic, social and cultural status vary from -1.6 in Indonesia to 0.5 in Qatar. The standard deviation of ESCS varies from 0.80 in Dubai (UAE) to 1.33 in Panama.

Finally, students provided their birthdays. The variable ***AGE*** included in the PISA 2009 dataset was calculated as the difference between the middle month and the year in which students were assessed and their month and year of birth, expressed in years and months. This variable varies from 15.25 to 16.33 at the student level (or 15 years and 3 months and 16 years and 4 months, respectively). For the analysis in this paper, the variable of student age (***AGE***) is recoded into a dichotomous variable to facilitate the interpretation of the results. The dichotomous variable ***rAGE*** is zero when a student is at or above the country average age and 1 when a student is below the country average age. Table A.4 presents the average age of students by country. As expected, about half of the students are at or above the country average age and the other half are below the country average age.

3. Method

This section introduces important statistical components linked to the design of PISA (sampling weights, replication weights, and plausible values) and the regression equations that are estimated in this study.

3.1. Sampling weights

The PISA target population is "15-year-old students attending educational institutions in grades 7 and higher" (OECD, 2012). PISA applies two-stage sampling. Schools are first selected with probabilities proportional to their size (PPS). Then, in general, 35 students are randomly sampled from each of the participating schools. In a two-stage sample design with a PPS sample of schools, all students should have the same probability of selection and consequently they should have the same weights. However, in PISA, student final weights (***W_FSTUWT***) vary because they are adjusted for non-responses of students or schools; the size measure for schools on the school sampling frame prepared at the early stage of the project is adjusted or updated after collecting data, and over-sampling or under-sampling of some strata of population is adjusted for. Thus, it is important to apply the student final weights when conducting analyses.

3.2. Replicate weights

As PISA applies two-stage sampling but not simple random sampling, it is recommended to compute standard errors based on the Balanced Repeated Replication (BRR) method (OECD, 2009). The *PISA 2009 Database* contains 80 replicate weights (***W_FSTR0*** to ***W_FSTR80***). The statistics of interest are computed with the student final weights as well as with 80 replicate weights (i.e. repeat the same computation 81 times). Then, the sampling variance is derived by aggregating the differences between each of the estimates with 80 replicate weights and the estimate computed with the student final weights. As PISA uses the Fay method with a factor of 0.5, the following formula is used to compute sampling variance:

$$\sigma^2_{(\hat{\theta})} = \frac{1}{80(1-0.5)^2}\sum_{i=1}^{80}(\hat{\theta}_{(i)} - \hat{\theta})^2$$

3.3. Plausible Values (PVs)

Students' responses to the cognitive items are scaled, based on classical Item Response Theory (IRT). As student proficiencies, which cannot be directly observed, have to be inferred based on the observed students' responses to cognitive items (reading, for the purposes of our analyses, but also for the other two areas of assessment, mathematics and science), PISA applies one of the inference approaches known as Plausible Values (PVs). See Appendix B for a detailed description of PVs.

3.4. Regressions

The results in Tables A.6, A.7, A.8 and A.9 in Appendix A are based on Ordinary Least Squares (OLS) regression models.

For grade repetition, three dummy variables are derived for the regression models:

- ***DREPN:*** A dummy variable indicating that a student has never repeated a grade.
- ***DREP1:*** A dummy variable indicating that a student has repeated a grade(s) at the primary education level only.
- ***DREP12:*** A dummy variable indicating that a student has repeated grades both at the primary and secondary education levels.

These three grade repetition dummy variables are included in the regression models. The reference group is students who have repeated a grade(s) at the secondary education level only.

In the regression models, the index of economic, social and cultural status (***ESCS***) is standardised to have the OECD mean of zero and a standard deviation of one across OECD countries.

Other students' characteristics variables – ***MALE, MIGLAN*** and ***rAGE*** – are all dichotomous variables (1 and 0) and defined in detail in the previous section.

- ***MALE:*** 1 = Boys and 0 = girls.
- ***MIGLAN:*** 1 = Students with migrant backgrounds or for whom the language spoken at home is different from the language of assessment; and 0 = others.
- ***rAGE:*** 1 = Students below the country average age; and 0 = students at or above the country average age.

Table A.6 presents the results of the following three models. Model 1 examines the relationship between student performance in reading and grade repetition before accounting for any of the students' characteristics. The intercept β_0 can be interpreted as the average score for the reference group, which is students who have repeated a grade(s) at the secondary education level only. The regression coefficient β_1 indicates the performance difference between students who have never repeated a grade and students who have repeated a grade(s) at the secondary education level only. The regression coefficient β_2 indicates the performance difference between students who have repeated a grade(s) at the primary education level only and students who have repeated a grade(s) at the secondary education level only. The regression coefficient β_3 indicates the performance difference between students who have repeated grades both at the primary and secondary education levels and students who have repeated a grade(s) at the secondary education level only.

Model 2 examines the relationship between student performance in reading and grade repetition after accounting for student socio-economic background. Model 3 examines this relationship after accounting for student socio-economic background, gender, migrant background and age.

Model 1(Equation 1):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12}$$

Model 2 (Equation 2):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12} + \beta_4 \text{ESCS}$$

Model 3 (Equation 3):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12} + \beta_4 \text{ESCS} + \beta_5 \text{MALE} + \beta_6 \text{MIGLAN} + \beta_7 \text{rAGE}$$

Four models in Table A.7 are built on Model 3 in Table A.6. In each of these four models, the interactions between grade repetition dummies and each of the student characteristics are included. The main focus is the performance difference between students who have never repeated a grade and students who have repeated a grade(s) at the secondary education level only; and the performance difference between students who have repeated a grade(s) at the primary education level only and students who have repeated a grade(s) at the secondary education level only. Thus, only two of the three grade repetition dummy variables are interacted with each of the students' characteristics in order to analyse how these performance differences by grade repetition vary depending on students' characteristics.

Model 4 (Equation 4):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12} + \beta_4 \text{ESCS} + \beta_5 \text{MALE} + \beta_6 \text{MIGLAN} + \beta_7 \text{rAGE} + \beta_8 \text{DREPN}*\text{ESCS} + \beta_9 \text{DREP1}*\text{ESCS}$$

Model 5 (Equation 5):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12} + \beta_4 \text{ESCS} + \beta_5 \text{MALE} + \beta_6 \text{MIGLAN} + \beta_7 \text{rAGE} + \beta_8 \text{DREPN}*\text{MALE} + \beta_9 \text{DREP1}*\text{MALE}$$

Model 6 (Equation 6):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12} + \beta_4 \text{ESCS} + \beta_5 \text{MALE} + \beta_6 \text{MIGLAN} + \beta_7 \text{rAGE} + \beta_8 \text{DREPN}*\text{MIGLAN} + \beta_9 \text{DREP1}*\text{MIGLAN}$$

Model 7 (Equation 7):

$$\text{Read} = \beta_0 + \beta_1 \text{DREPN} + \beta_2 \text{DREP1} + \beta_3 \text{DREP12} + \beta_4 \text{ESCS} + \beta_5 \text{MALE} + \beta_6 \text{MIGLAN} + \beta_7 \text{rAGE} + \beta_8 \text{DREPN}*\text{rAGE} + \beta_9 \text{DREP1}*\text{rAGE}$$

For all models in Tables A.6 and A.7, five Plausible Values, the student final weights and 80 replicate weights are used.

Similarly, three models (see Table A.8) and four models with interactions (see Table A.9) are developed with the index of student attitudes towards school (***ATSCHL***) as an outcome variable. All of these models in Tables A.8 and A.9 are based on the cases without any missing values in the index of student attitudes towards school, grade repetition, student socio-economic background, gender, migrant background and age. For all models in Tables A.8 and A.9, the student final weights and 80 replicate weights are used.

Table A.8 presents the results of Models 8 to 10.

Model 8 (Equation 8):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12$$

Model 9 (Equation 9):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12 + \beta_4 ESCS$$

Model 10 (Equation 10):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12 + \beta_4 ESCS + \beta_5 MALE + \beta_6 MIGLAN + \beta_7 rAGE$$

Table A.9 presents the results of Models 11 to 14.

Model 11 (Equation 11):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12 + \beta_4 ESCS + \beta_5 MALE + \beta_6 MIGLAN + \beta_7 rAGE + \beta_8 DREPN{*}ESCS + \beta_9 DREP1{*}ESCS$$

Model 12 (Equation 12):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12 + \beta_4 ESCS + \beta_5 MALE + \beta_6 MIGLAN + \beta_7 rAGE + \beta_8 DREPN{*}MALE + \beta_9 DREP1{*}MALE$$

Model 13 (Equation 13):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12 + \beta_4 ESCS + \beta_5 MALE + \beta_6 MIGLAN + \beta_7 rAGE + \beta_8 DREPN{*}MIGLAN + \beta_9 DREP1{*}MIGLAN$$

Model 14 (Equation 14):

$$ATSCHL = \beta_0 + \beta_1 DREPN + \beta_2 DREP1 + \beta_3 DREP12 + \beta_4 ESCS + \beta_5 MALE + \beta_6 MIGLAN + \beta_7 rAGE + \beta_8 DREPN{*}rAGE + \beta_9 DREP1{*}rAGE$$

3.5. Counterfactual for retained students and causality

Two final considerations about our empirical strategy are discussed to conclude this section. Some of the papers estimating the effects of retention face an important limitation when constructing the counterfactual for the performance of a retained student. A proper counterfactual is not possible in the absence of vertically scaled test scores that make it possible to compare the achievement of students tested in different grades (Schwerd and West, 2012): "same grade comparisons conflate any effect of retention on achievement with the effect of being a year older when you take the test." As in their paper, PISA provides us with the right performance counterfactual due to the fact that PISA's goal is to test the ability to apply the "contents" of education to real life, and not the ability to "master" the contents. As described in Section 2, the student population in PISA is composed of 15-year-olds, who, as a result of retention or other factors, may not be attending the grade level that is appropriate for their age (the modal grade). Thus, as in the paper by Schwerd and West, our counterfactual condition for retained students is to have been immediately promoted to the next grade, and we compare the ability to apply educational contents to real life of retained versus non-retained students. Regarding

another educational outcome variable used in this paper (i.e. the index of attitudes towards school), a single index was developed in a way that students' attitudes towards school can be directly compared among students sampled in the PISA 2009 regardless of their grade levels.

On causality, the characteristics of the data source do not allow us to utilise a quasi-experimental method that would be more appropriate to address the causal impact of retention on the students' achievement and socio-emotional behaviour. To overcome the selection into the retention treatment this paper sets out an observational study with observable individual level controls. Since some of these controls are measured after the treatment, and although it is unlikely that the controls, as measured by PISA, are affected by the students' outcomes, the results should be interpreted cautiously acknowledging the limitations. The concerns with unobservable characteristics, including whether they are affected by the students' outcomes that are norm in this type of identification, also apply to the analysis in this paper.

4. Empirical results

4.1. Grade repetition

Figure 1 shows 16 OECD countries and 14 non-OECD countries and economies where over 10% of 15-year-old students reported that they had repeated a grade at least once. In the Netherlands, Mexico, Ireland, the United States, Switzerland, Trinidad and Tobago, Indonesia, Peru and Hong Kong-China, students are more likely to repeat a grade at the level of primary education compared with secondary education. On the other hand, in Spain, Italy, Turkey, Chile, France, Germany, Belgium, Austria, Tunisia, Argentina, Macao-China, Brazil, Panama and Uruguay, students are more likely to repeat a grade at the level of secondary education compared with primary education. In Hungary, Portugal, Luxembourg, Qatar, Latvia, Dubai (UAE) and Colombia, there is not much difference in the percentage of students who repeated a grade in primary versus secondary education. In all of these 30 countries and economies except Spain, Brazil, Uruguay, Macao-China and Tunisia, 5% or less of students repeated grades both at the primary and secondary education levels (Table A.1).

4.2. Grade repetition and student performance in reading

Figure 2 shows that, when the bivariate relationship between grade repetition and performance is examined, with almost no exception, performance of students who never repeated a grade, compared with students who were retained in secondary school, is higher (Model 1 in Table A.6). The achievement gaps between these groups range from 35 score points (in Austria and Colombia) to close to 80-100 score points in a wide set of countries (France, Argentina, Spain, Hungary, Ireland, Italy, Portugal, Dubai, Qatar, Uruguay and the United States) and 152 score points in Trinidad and Tobago. The difference is not statistically significant in the Netherlands (where the proportion of retained students was one of the smallest in the sample).

Another important finding is that students who repeated a grade at the secondary education level only tend to perform better than students who repeated a grade at the primary education level only in 23 out of 30 countries and economies that were examined (Model 1 in Table A.6). The score-point differences range from 12 points in Macao-China to around 74 points in Austria and Luxembourg. The only two exceptions are Qatar and

Figure 1. **Grade repetition**

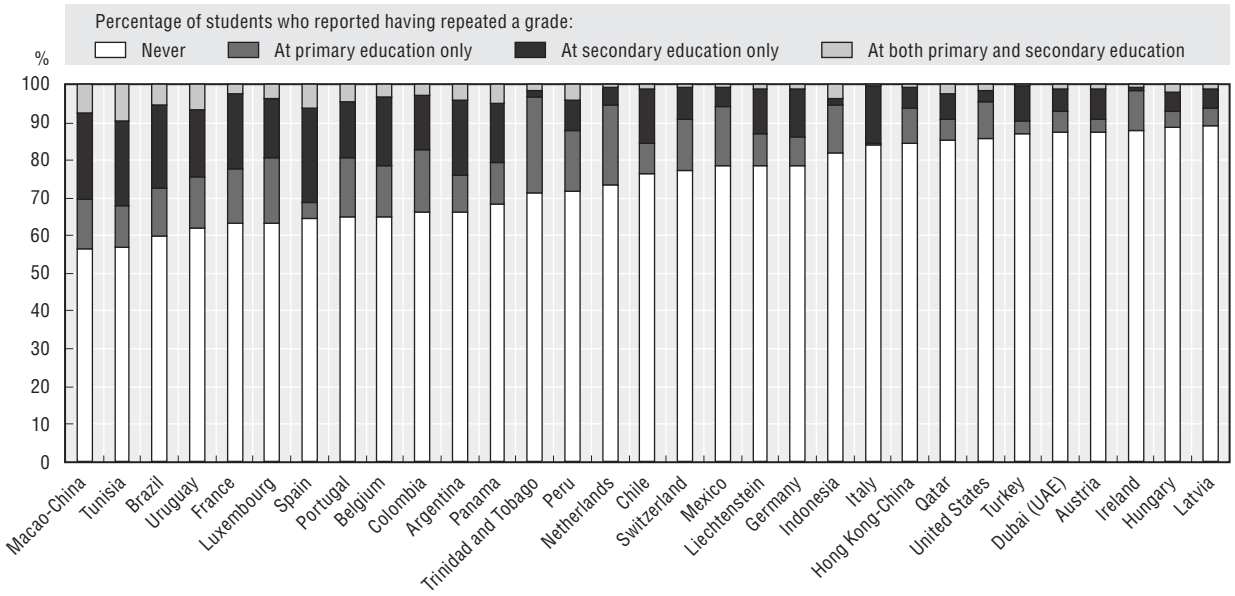
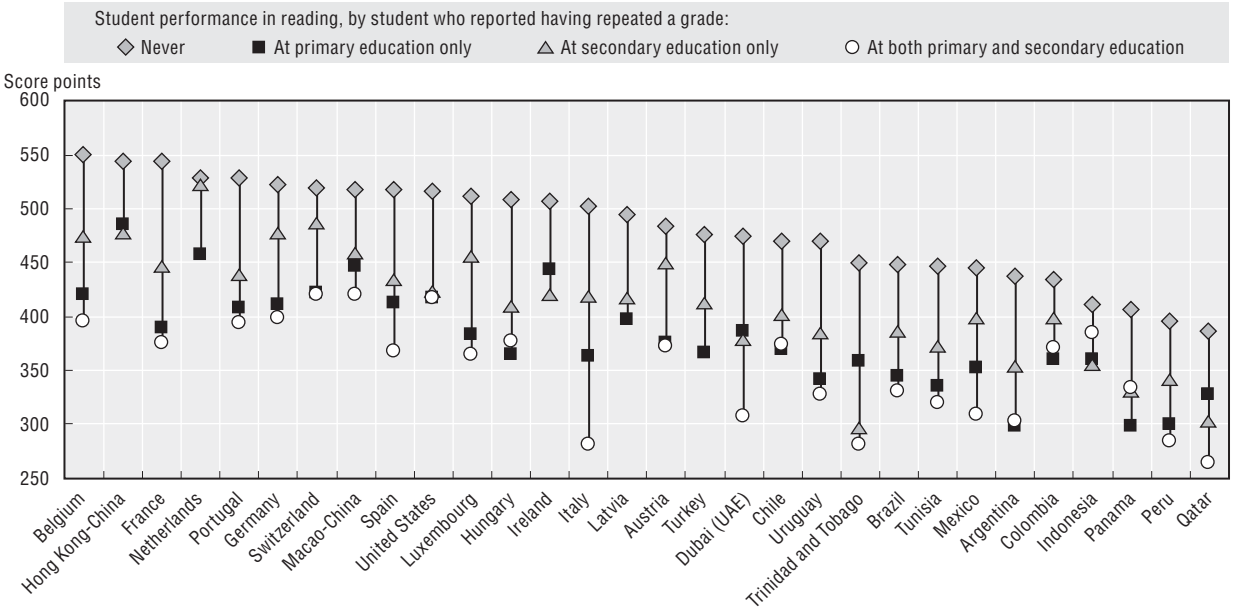


Figure 2. **Reading performance by grade repetition**



Trinidad and Tobago, where the gap is favourable to students who were retained in primary school, compared with students retained in secondary school. In Hong Kong-China, Indonesia, Ireland, Dubai (UAE) and the United States, the difference in performance between these groups is not statistically significant. In these five countries and economies, there is no difference in reading performance between these two groups of students and a small percentage of students repeats a grade at the secondary education level (1.0% in Ireland, 3.2% in the United States, 5.3% in Hong Kong-China, 5.8% in Dubai (UAE) and 1.7% in Indonesia).

4.3. Relationship between performance and grade repetition after accounting for student background

Model 2 in Table A.6 examines the achievement gaps among students who repeated a grade after accounting for student socio-economic background. The results suggest that the achievement gaps between students who never repeated a grade and students who were retained in secondary school decrease after accounting for student socio-economic background, while the statistical significance of the coefficient remains unchanged. The largest drops in performance difference among these students are found in Hungary (32 point change), the United States (24 point change) and Dubai (16 point change). The gap between non-repeaters and students who repeated a grade in secondary school only increases in Mexico and Indonesia (very modestly, by 2 score points). Similarly, the gap between students who repeated in primary school and students who repeated in secondary school is also reduced after accounting for the students' socio-economic background, by a considerable number of score points. The largest reduction in the achievement gap between students who repeated a grade in primary school and in secondary school is observed in Luxembourg, Peru, Mexico, Austria, Germany, Italy and Turkey (between 15 and 20 score-point change). The difference among students who repeated in primary school and students who repeated in secondary school is not statistically significant after accounting for student socio-economic background in Latvia and Panama. Regarding the association between performance and student socio-economic background, the results in all countries and economies that are examined in this paper show a strong positive association (statistically significant).

After accounting for other individual characteristics in Model 3, the performance gaps shrink slightly, compared with the results of Model 2 where the relationship is examined only after accounting for student socio-economic background. The gap between students who never repeated a grade and students who repeated a grade in secondary school is reduced by 3 to 6 score points in most of the countries, and by 11 to 15 score points in Dubai (UAE), Qatar, Trinidad and Tobago and Latvia. Regarding the gap between students who were retained in primary and secondary school, the achievement gap stays the same or increases by around 2 score points in more than half of the countries. This gap decreases in Austria, France, Hungary, Italy, Luxembourg, Mexico, Portugal, Switzerland, Argentina, Brazil, Qatar, Dubai and Peru, and increases in Indonesia and Trinidad and Tobago. The association between reading performance and student socio-economic background is still positive, strong and statistically significant across all the countries. In contrast, the performance gap between boys and girls, identified by the coefficients associated with the male dummy, suggests that boys underscore girls in all the countries in the sample. The gender gaps vary between 10 score points in Colombia and 51 score points (which is equivalent to half of a standard deviation) in Trinidad and Tobago. Among native and immigrant students, immigrant students significantly outperform native students in only three countries (22 point difference in Luxembourg, 52 point difference in Qatar and 72 point difference in Dubai), while in most countries native students score higher in the reading test than do students with an immigrant background (40 point difference and more in Colombia, Brazil, Mexico, Peru, Chile and Panama). Finally, the association between reading achievement and the age of the students is examined. Overall, the estimated coefficients suggest that students who are below the average age of the country underperform the elder students. On average, the estimated coefficients vary between 4 and 20 fewer score points in Spain and Luxembourg, respectively.

Overall, even after accounting for students' background characteristics, similar relationships to the bivariate relationships found in Model 1 and Table A.6 can still be observed. In some countries, the performance differences between grade repetition and performance are reduced after accounting for these characteristics, which suggests that the student background characteristics are related to both grade repetition and performance. But, in general, even after accounting for students' background characteristics, non-repeaters perform better than repeaters; and secondary-education repeaters perform better than primary-education repeaters. This finding leads to the question of whether these relationships are common across different sub-groups of students. The following sections examine how the relationship between grade repetition and performance varies according to student background characteristics.

4.4. How the relationships between performance and grade repetition vary according to student background characteristics

In this section, the interaction between grade repetition and student characteristics is examined. In particular, the analyses focus on how the relationship between repetition and performance is different across some characteristics of the students such as student socio-economic background, gender, migrant/language background and age. The results are shown in Table A.7.

4.4.1. *Heterogeneous relationships across student socio-economic background*

Table 1 (i.e. Model 4 in Table A.7) examines the interaction of grade repetition and student socio-economic background in relation to performance. For students whose socio-economic backgrounds are around or at the level of the country-average, a student who never repeated a grade tends to perform better than a student who repeated during secondary school, in all countries and economies examined except the Netherlands (see the coefficients of drepN). Similar to the results of Model 3 in Table A.6, the difference ranges from 24 score points in Austria to 146 in Trinidad and Tobago. In 18 countries, this gap is even wider for students with more advantaged socio-economic backgrounds (see the coefficients of drepN*ESCS). In the Netherlands, the interaction is statistically significant, despite the fact that the difference between non-repeaters and secondary school repeaters was not significant for students whose socio-economic backgrounds are at the OECD average level.

Figure 3 presents the performance differences between non-repeaters and repeaters in secondary school for students with a disadvantaged socio-economic background (i.e. one standard deviation lower than the OECD average), and those for students with more advantaged socio-economic backgrounds (i.e. one standard deviation higher than the OECD average) for these 17 countries and economies with significant results. The former is shown with black bars and the latter with grey bars. The biggest differences between the black and grey bars are found in Trinidad and Tobago, where the difference is over 40 score points. This difference is over 30 score points in Peru, Turkey, Uruguay and Luxembourg.

Model 4 in Table A.7 also examines the performance advantage of students who repeated a grade during secondary school relative to students who repeated a grade during primary school, and how this performance advantage differs between students with country-average levels of socio-economic background and students with more advantaged socio-economic backgrounds (see the coefficients of drep1*ESCS). Only in Mexico, Argentina, Brazil and Panama does this performance advantage for secondary-school

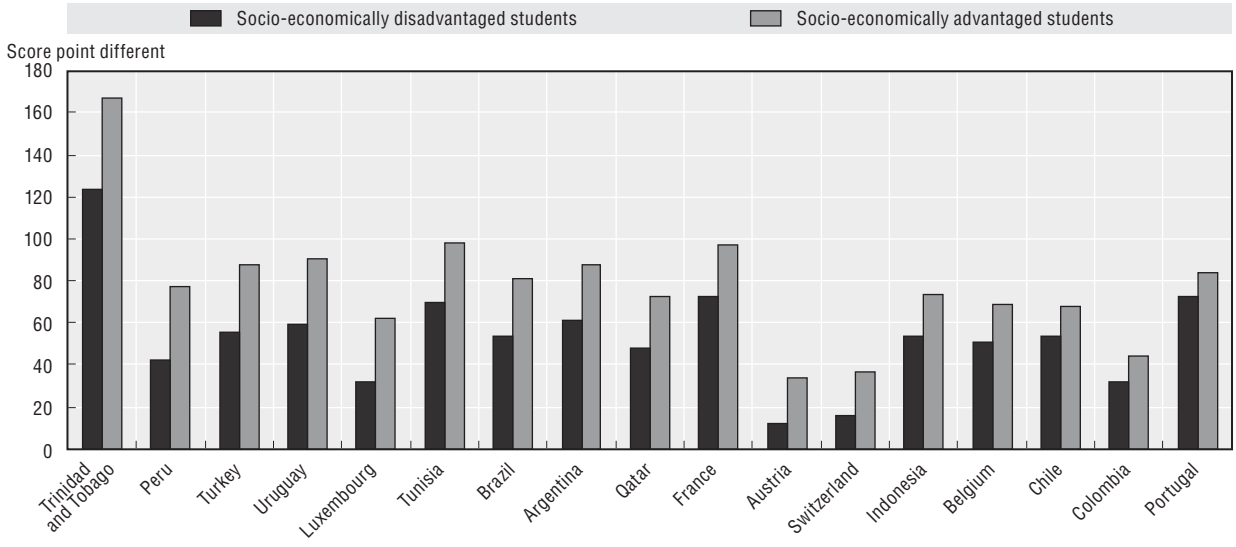
Table 1. **Grade repetition, reading performance and student socio-economic background**

	Model 4																				
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*ESCS		drep1*ESCS		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**487**	(6.0)	**24**	(5.5)	**-68**	(8.4)	**-64**	(13.3)	**31**	(4.7)	**-40**	(4.5)	**-28**	(5.5)	**-20**	(3.0)	**11**	(5.0)	-15	(9.3)	0.25
Belgium	**488**	(3.9)	**61**	(3.4)	**-47**	(4.3)	**-69**	(6.7)	**23**	(3.1)	**-20**	(2.7)	**-18**	(3.6)	3	(2.2)	**9**	(3.4)	-2	(4.2)	0.39
Chile	**433**	(3.9)	**61**	(3.7)	**-28**	(6.6)	-27	(13.9)	**21**	(2.8)	**-16**	(3.1)	**-55**	(12.1)	**-7**	(2.0)	**7**	(2.9)	-1	(4.7)	0.33
France	**473**	(7.8)	**85**	(8.1)	**-50**	(8.5)	**-62**	(14.5)	**21**	(5.7)	**-31**	(3.3)	**-27**	(5.7)	1	(2.7)	**12**	(6.3)	-2	(7.4)	0.45
Germany	**506**	(4.2)	**33**	(4.2)	**-57**	(5.5)	**-59**	(12.0)	**28**	(4.0)	**-36**	(3.3)	**-25**	(3.8)	**-15**	(2.8)	7	(4.6)	-7	(6.9)	0.31
Hungary	**475**	(7.6)	**62**	(7.0)	**-52**	(16.2)	**-34**	(10.8)	**40**	(5.6)	**-38**	(3.0)	-2	(9.8)	**-11**	(2.4)	1	(5.7)	-12	(10.2)	0.39
Ireland	**463**	(14.7)	**69**	(14.8)	11	(15.8)	c	c	**31**	(10.8)	**-40**	(3.7)	**-18**	(6.6)	**-13**	(2.7)	7	(11.1)	3	(11.7)	0.23
Italy	**460**	(2.7)	**69**	(2.6)	**-45**	(14.0)	**-99**	(19.5)	**22**	(3.1)	**-39**	(2.3)	**-27**	(2.5)	**-7**	(1.5)	5	(3.2)	-10	(12.5)	0.26
Luxembourg	**457**	(6.2)	**47**	(3.5)	**-57**	(5.5)	**-76**	(8.6)	**19**	(2.5)	**-35**	(2.2)	**21**	(5.6)	**-20**	(2.7)	**15**	(2.9)	4	(4.4)	0.40
Mexico	**436**	(6.7)	**48**	(6.3)	**-46**	(7.4)	**-72**	(10.7)	**16**	(2.7)	**-20**	(1.3)	**-61**	(5.5)	**-10**	(1.5)	4	(2.9)	**-8**	(3.6)	0.31
Netherlands	**531**	(9.0)	0	(8.2)	**-60**	(8.3)	c	c	**17**	(6.5)	**-22**	(2.6)	-12	(7.8)	-3	(3.0)	**16**	(6.6)	9	(7.4)	0.22
Portugal	**465**	(3.6)	**79**	(4.0)	**-25**	(4.3)	**-39**	(5.2)	**14**	(2.5)	**-28**	(2.0)	**-16**	(3.9)	-3	(1.9)	**6**	(2.8)	2	(3.7)	0.46
Spain	**462**	(2.7)	**71**	(2.1)	**-23**	(5.4)	**-61**	(4.8)	**14**	(2.0)	**-21**	(1.6)	**-14**	(2.9)	**-4**	(1.7)	3	(1.8)	-3	(4.2)	0.38
Switzerland	**517**	(3.9)	**27**	(4.2)	**-54**	(4.7)	**-55**	(12.6)	**23**	(3.4)	**-37**	(2.2)	**-22**	(2.5)	**-13**	(2.4)	**10**	(4.1)	3	(4.7)	0.29
Turkey	**451**	(8.7)	**72**	(8.3)	**-52**	(24.3)	c	c	**12**	(3.9)	**-32**	(3.1)	-14	(8.6)	-3	(2.3)	**16**	(3.9)	-8	(11.1)	0.31
United States	**454**	(8.8)	**65**	(8.6)	-17	(10.7)	-15	(14.7)	**40**	(6.6)	**-22**	(2.7)	6	(3.5)	**-7**	(2.5)	-2	(7.1)	-9	(12.3)	0.26
Partners																					
Argentina	**391**	(6.2)	**75**	(6.4)	**-60**	(11.3)	**-43**	(9.0)	**24**	(3.1)	**-29**	(3.1)	-18	(9.8)	-4	(2.9)	**13**	(4.0)	**-12**	(6.1)	0.38
Brazil	**422**	(3.0)	**68**	(3.8)	**-48**	(5.3)	**-49**	(4.5)	**15**	(1.8)	**-25**	(1.8)	**-64**	(9.2)	**-9**	(2.1)	**14**	(2.5)	**-8**	(2.8)	0.32
Colombia	**431**	(5.7)	**39**	(5.7)	**-31**	(7.5)	**-22**	(8.2)	**21**	(2.8)	**-10**	(3.0)	**-68**	(14.2)	**-6**	(2.9)	**6**	(3.2)	-1	(3.9)	0.25
Dubai (UAE)	**363**	(5.5)	**63**	(5.1)	13	(6.8)	**-44**	(11.5)	**37**	(4.7)	**-48**	(2.5)	**72**	(2.9)	**-11**	(2.6)	0	(5.1)	0	(7.4)	0.32
Hong Kong-China	**512**	(8.4)	**60**	(8.3)	7	(10.4)	c	c	**16**	(5.8)	**-30**	(3.6)	3	(4.3)	**-6**	(2.4)	-1	(5.9)	-3	(6.0)	0.15
Indonesia	**387**	(13.4)	**64**	(12.1)	7	(14.0)	**28**	(8.7)	8	(5.3)	**-34**	(3.0)	7	(4.0)	-3	(2.4)	**10**	(4.9)	0	(5.8)	0.21
Latvia	**458**	(7.0)	**60**	(6.7)	-21	(12.4)	c	c	**17**	(7.1)	**-43**	(2.6)	-7	(5.5)	-3	(2.7)	10	(6.8)	-2	(12.2)	0.26
Macao-China	**483**	(3.4)	**57**	(2.9)	-7	(5.2)	**-38**	(3.8)	4	(2.1)	**-25**	(1.7)	**-5**	(2.0)	**-6**	(2.2)	3	(2.3)	4	(4.4)	0.25
Panama	**371**	(8.9)	**70**	(7.5)	**-52**	(12.2)	-1	(9.7)	**23**	(4.4)	**-19**	(6.6)	**-37**	(14.1)	-6	(5.1)	8	(4.5)	**-21**	(5.3)	0.35
Peru	**387**	(7.0)	**60**	(7.3)	**-23**	(8.4)	**-50**	(7.6)	**19**	(3.4)	**-16**	(2.9)	**-60**	(7.3)	-4	(2.6)	**17**	(4.0)	3	(4.3)	0.38
Qatar	**301**	(4.1)	**61**	(3.8)	**19**	(5.8)	**-32**	(6.8)	**8**	(3.7)	**-39**	(1.7)	**52**	(2.0)	**-6**	(1.7)	**13**	(3.7)	**11**	(5.2)	0.18
Trinidad and Tobago	**345**	(17.9)	**146**	(17.4)	**62**	(18.3)	-24	(17.5)	15	(9.8)	**-50**	(2.9)	**-25**	(9.5)	-2	(2.9)	**21**	(10.0)	11	(10.4)	0.29
Tunisia	**386**	(5.3)	**84**	(6.0)	**-29**	(7.6)	**-52**	(4.5)	0	(2.1)	**-20**	(2.1)	c	c	**-6**	(2.2)	**14**	(2.7)	5	(3.6)	0.39
Uruguay	**418**	(5.2)	**76**	(5.1)	**-32**	(6.1)	**-52**	(5.2)	**12**	(2.4)	**-33**	(2.7)	**-22**	(8.7)	-3	(2.2)	**15**	(2.6)	5	(3.2)	0.41

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Figure 3. **Grade repetition and reading performance, by student socio-economic background**

Difference in reading performance between non-repeaters and repeaters in secondary education



repeaters significantly increase for socio-economically advantaged students compared with students with average socio-economic backgrounds. In Qatar, the opposite relationship is observed: primary-school repeaters tend to perform at higher scores than secondary-school repeaters for students with average socio-economic background; and this difference increases for students with more advantaged backgrounds. The interaction is not statistically significant in the rest of the countries.

4.4.2. *Heterogeneous relationships across gender*

In Model 5 in Table A.7, the relationship between grade repetition and reading performance is examined according to student gender. In most countries and economies examined, the relationships between grade repetition and performance seem to be similar for boys and for girls (see the coefficients of drepN*MALE and drep1*MALE). Only in a few countries do the relationships differ between boys and girls. In Mexico, girls who have never repeated a grade outperform girls who repeated in secondary school by 55 score points and girls who repeated in secondary school outperform girls who repeated in primary school by 13 score points, while boys who have never repeated a grade outperform boys who repeated in secondary school by 38 score points and boys who repeated in secondary school outperform boys who repeated in primary school by 37 score points. In Latvia, Qatar, and Tunisia, girls who never repeated outperform girls who repeated in secondary school, and this performance gap is wider for boys than girls. The performance gap between secondary-school repeaters and primary-school repeaters (i.e. the secondary-school repeaters outperform the primary-school repeaters) is wider for boys than girls in Luxembourg (40 points difference for girls and 62 points difference for boys), but this gap is narrower for boys than girls in Spain (32 points difference for girls and 12 points difference for boys). In the rest of the countries, the difference in performance between repeaters and non-repeaters is not statistically significantly different across boys and girls.

4.4.3. *Heterogeneous relationships across migrant/language background*

According to Model 6 in Table A.7, the differences in performance between non-repeaters and secondary-school repeaters vary depending on immigrant/language backgrounds of students in eight countries and economies. For native students, the difference is in favour of students who have never repeated a grade in all countries examined except the Netherlands. Similarly, among immigrant students, the performance difference between non-repeaters and secondary-school repeaters stays the same, except for the following six countries and economies: in Germany, Luxembourg, Portugal, Colombia and Macao-China, the performance advantage of non-repeaters is smaller for students with migrant backgrounds than native students. In Colombia, for native students, non-repeaters outperform secondary-school repeaters, while for students with migrant backgrounds, secondary-school repeaters outperform non-repeaters. Only in Qatar is the performance advantage for non-repeaters even greater for students with migrant backgrounds than native students.

In the majority of countries, the achievement gap between students who repeated in secondary school and students who repeated in primary school for native students is in favour of the former group. On average, in OECD countries, this difference is between 21 and 68 score points. A similar relationship is observed in non-OECD countries, although the gaps are smaller (16 to 38 score points). The gap in Qatar and Trinidad and Tobago is positive, which suggests that primary school-native repeaters outperform (18 and 47 score points) secondary school-native repeaters. The gaps are very similar for students with immigrant backgrounds. Only in Belgium and Mexico is this performance gap between secondary-school repeaters and primary-school repeaters smaller for students with migrant backgrounds. In Mexico, for students with migrant backgrounds, on average, primary-school repeaters outperform secondary-school repeaters.

4.4.4. *Heterogeneous relationships across age*

Finally, Model 7 tests if the association between performance and repetition would be different for younger students or for older students. In general, the coefficients estimated for the different countries suggest that the achievement gaps between non-repeaters and students who repeated in secondary school are not different for students who are younger or older than students at the country-average age. Only in Brazil is the performance advantage of non-repeaters relative to secondary-school repeaters smaller for younger students than older students. In Austria, among students whose ages are at or above the country average, secondary-school repeaters outperform primary-school repeaters, and this difference is smaller among younger students. In Ireland, among students whose ages are at or above the country average, there is no performance difference between primary- and secondary-school repeaters, while primary-school repeaters outperform secondary-school repeaters among younger students.

4.5. *Grade repetition and student attitude towards school*

As anticipated earlier, this research examines the associations of grade repetition with an indicator of the socio-emotional or non-cognitive outcomes of the student, represented in this case by the index of student attitude towards school. When the bivariate relationship between the index of student attitude towards school and grade repetition is examined, the results show that, in general, in many countries secondary-school repeaters tend to have lower values of the index (i.e. more negative attitudes towards school) than

non-repeaters. But, across countries, the results are less consistent in terms of the difference in student attitude between primary-school repeaters and non-repeaters as well the difference between primary- and secondary-school repeaters.

Table 2 shows that non-repeaters tend to have higher values on the index of student attitudes towards school than do those who repeated a grade in secondary education in 11 OECD countries and 10 non-OECD countries and economies. Primary-school repeaters tend to report more positive attitudes towards school than do secondary-school repeaters in 11 countries and economies, while the opposite is observed in two countries.

Table 2. **Students' attitudes towards school and grade repetition**

	Model 8								
	Intercept		drepN		drep1		drep12		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD									
Austria	**-0.17**	(0.07)	**0.26**	(0.06)	**0.32**	(0.10)	0.01	(0.26)	0.01
Belgium	**-0.14**	(0.03)	0.02	(0.03)	**0.15**	(0.04)	-0.05	(0.08)	0.00
Chile	**0.15**	(0.04)	**0.19**	(0.05)	-0.07	(0.07)	-0.06	(0.18)	0.01
France	**-0.13**	(0.04)	**0.27**	(0.05)	**0.23**	(0.07)	0.13	(0.16)	0.01
Germany	**-0.25**	(0.05)	**0.16**	(0.05)	**0.21**	(0.07)	-0.22	(0.20)	0.00
Hungary	**-0.19**	(0.09)	0.16	(0.10)	0.05	(0.10)	0.08	(0.16)	0.00
Ireland	0.23	(0.17)	-0.08	(0.17)	-0.17	(0.18)	c	c	0.00
Italy	**-0.12**	(0.02)	**0.17**	(0.03)	0.20	(0.16)	-0.33	(0.26)	0.00
Luxembourg	**-0.28**	(0.04)	**0.16**	(0.05)	**0.25**	(0.06)	0.16	(0.13)	0.00
Mexico	**0.25**	(0.05)	**0.28**	(0.05)	**-0.19**	(0.05)	**-0.50**	(0.10)	0.04
Netherlands	**-0.29**	(0.06)	0.06	(0.06)	0.10	(0.07)	c	c	0.00
Portugal	**0.12**	(0.05)	**0.32**	(0.05)	**0.15**	(0.06)	0.08	(0.08)	0.02
Spain	-0.01	(0.03)	**0.25**	(0.03)	**0.19**	(0.05)	0.02	(0.05)	0.01
Switzerland	**-0.11**	(0.05)	**0.11**	(0.04)	**0.27**	(0.05)	0.13	(0.18)	0.00
Turkey	0.08	(0.05)	**0.13**	(0.05)	0.21	(0.19)	c	c	0.00
United States	0.08	(0.08)	0.15	(0.09)	0.13	(0.11)	-0.15	(0.18)	0.00
Partners									
Argentina	**0.08**	(0.03)	0.05	(0.05)	0.10	(0.07)	-0.14	(0.11)	0.00
Brazil	**0.37**	(0.03)	**0.12**	(0.03)	-0.02	(0.04)	**-0.16**	(0.05)	0.01
Colombia	**0.38**	(0.05)	**0.15**	(0.05)	-0.11	(0.06)	0.02	(0.10)	0.01
Dubai (UAE)	**-0.12**	(0.05)	**0.21**	(0.05)	0.02	(0.07)	-0.14	(0.14)	0.01
Hong Kong-China	**-0.59**	(0.05)	**0.11**	(0.05)	**0.13**	(0.05)	c	c	0.00
Indonesia	0.05	(0.09)	**0.48**	(0.09)	**0.27**	(0.11)	**0.42**	(0.12)	0.01
Latvia	0.10	(0.08)	0.11	(0.08)	-0.19	(0.12)	c	c	0.01
Macao-China	**-0.49**	(0.02)	**0.06**	(0.03)	0.05	(0.03)	-0.01	(0.04)	0.00
Panama	**0.23**	(0.08)	**0.26**	(0.08)	-0.05	(0.11)	-0.02	(0.12)	0.02
Peru	**0.16**	(0.05)	0.08	(0.06)	**-0.17**	(0.06)	**-0.33**	(0.10)	0.01
Qatar	**-0.24**	(0.04)	**0.13**	(0.04)	0.01	(0.06)	-0.14	(0.08)	0.00
Trinidad and Tobago	-0.17	(0.13)	**0.71**	(0.13)	**0.55**	(0.14)	-0.33	(0.18)	0.02
Tunisia	**0.43**	(0.04)	**0.18**	(0.05)	0.06	(0.05)	**-0.20**	(0.07)	0.01
Uruguay	**0.11**	(0.03)	0.08	(0.04)	0.09	(0.05)	-0.11	(0.06)	0.00

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

4.6. Relationship between attitudes towards school and grade repetition after accounting for student background

The results of Model 9 and Model 10 in Table A.8 indicate that, even after accounting for student socio-economic background and other student background characteristics, the patterns found with Model 8 mainly remain unchanged. The only differences in the coefficient for grade repetition dummies between Model 8 and Models 9 and 10 are found in Macao-China and Qatar, where there is no difference in students' attitudes towards school between non-repeaters and secondary-school repeaters in Models 9 and/or 10, and in Uruguay, where non-repeaters tend to report more positive attitudes than secondary-school repeaters in Models 9 and 10.

Although the coefficients associated with the repetition variables remain stable after the inclusion of the covariates of student background characteristics, it is important to note some statistically significant associations between the index of attitudes towards school and some of the background characteristics. For instance, boys tend to report more negative attitudes towards schools than girls, on average. The gaps between boys and girls vary, on average, between -0.05 index points in Belgium and -0.20 index points in Portugal. The only country in which boys report more positive attitudes towards school than girls is Chile (0.08 index point difference). The difference in the index between boys and girls is not statistically significant in Austria, Ireland, the Netherlands, the United States, Argentina, Colombia, Hong Kong-China, Macao-China, Panama, Qatar and Uruguay. Regarding the differences between native students and students with migrant backgrounds, the direction and the size of difference vary across countries. Native students tend to report more negative attitudes toward school than do students with migrant backgrounds in Austria, Belgium, Switzerland, Germany, France, the Netherlands, the United States, Indonesia and Qatar. In contrast, native students tend to report more positive attitudes towards school than do students with migrant backgrounds in Chile, Spain, Mexico, Brazil, Panama and Peru. In the rest of the countries, the differences are not statistically significant. Regarding student age, no significant relationship between age of the students and their attitudes towards school is found in most countries. The exceptions are the Netherlands and Switzerland, where younger students tend to report more positive attitudes towards school than do older students and Mexico, where older students tend to report, on average, more positive attitudes than do younger students. Finally, the relationship between socio-economic status and the index of attitudes towards school remains similar between Models 9 and 10. In half of the countries (according to Model 10 in Table A.8), the association between attitudes towards school and socio-economic status is positive, including Belgium, France, Ireland, Mexico, the Netherlands, the United States, Hong Kong-China, Indonesia, Macao-China, Panama, Peru and Qatar. In contrast, the relationship is negative in Germany, Luxembourg and Turkey. The relationship is not statistically significant in the remaining countries in the sample.

4.7. How the relationship between attitudes towards school and grade repetition vary according to student background characteristics

4.7.1. Heterogeneous relationships across student socio-economic background

Table A.9 presents the results of Models 11 to 14. Model 11 adds an interaction between grade repetition and socio-economic background to Model 10. This model examines how the relationship between grade repetition and students' attitudes toward school varies according to student socio-economic background. The interaction terms (see

the columns of drepN*ESCS and drep1*ESCS) suggest that the relationship between grade repetition and student attitudes towards school are similar across different levels of student socio-economic background in most countries and economies examined.

The only exceptions are in Italy, Mexico, Turkey, Brazil, Macao-China, Panama and Peru. In Italy and Brazil, for students with the OECD average level of socio-economic background, there is no difference in attitudes towards school between primary- and secondary-school repeaters, while for students with more advantaged socio-economic backgrounds (i.e. one standard deviation higher on the ESCS index) primary-school repeaters tend to report more negative attitudes than do secondary-school repeaters by 0.34 index points and 0.06 index points respectively. In Mexico, for students with the OECD average level of socio-economic background, non-repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.18 points and secondary-school repeaters tend to report more positive attitudes than do primary-school repeaters by 0.32 points, while for students with more advantaged socio-economic backgrounds the difference in attitudes between non-repeaters and secondary-school repeaters decreases, and secondary-school repeaters tend to report more positive attitudes than do primary-school repeaters by 0.44 points. In Turkey, for students with the average level of socio-economic background for the OECD countries, there is no difference in attitude between non-repeaters and secondary-school repeaters, while for students with more advantaged socio-economic backgrounds, secondary-school repeaters tend to report more positive attitudes than do non-repeaters by 0.20 index points. In Macao-China, for students with the OECD average level of socio-economic background non-repeaters report more positive attitudes than secondary-school repeaters by 0.09 index points, and for students with more advantaged backgrounds non-repeaters tend to report even more positive attitudes than do secondary-school repeaters by 0.16 index points. In Panama, for students with the average level of socio-economic background for the OECD countries there is no difference in attitude between primary- and secondary-school repeaters, while for students with more advantaged backgrounds, primary-school repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.45 index points. In Peru, for students with the OECD average level of socio-economic background, there is no difference in attitudes between non-repeaters and secondary-school repeaters as well as between primary- and secondary-school repeaters, while for students with more advantaged backgrounds secondary-school repeaters tend to report more negative attitudes than do non-repeaters by 0.26 index points and secondary-school repeaters tend to report more negative attitudes than do primary-school repeaters by 0.13 index points.

4.7.2. Heterogeneous relationships across gender

The results of Model 12 show that the relationships between students' attitudes towards school and grade repetition do not differ between boys and girls in almost all countries. The gender difference is observed only in the comparison between students who repeated in primary school and those who repeated in secondary school in Luxembourg, Tunisia and Trinidad and Tobago. In these countries, girls who repeated in primary school tend to report more positive attitudes than do girls who repeated in secondary school. In Luxembourg, for boys there is almost no difference between primary- and secondary-school repeaters (0.10 index point difference). In Trinidad and Tobago, the difference in attitudes between primary- and secondary-school repeaters is smaller for

boys than for girls. In Tunisia, for boys, secondary-school repeaters have more positive attitudes towards schools than primary-school repeaters.

4.7.3. Heterogeneous relationships across immigrant status

Model 13 in Table A.9 examines the relationships between grade repetition and students' attitudes towards school and how these relationships vary between native students and students with migrant backgrounds. The results show that the relationships are similar between native students and students with migrant backgrounds in most countries. There are four exceptions. In Austria, for native students, non-repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.35 index points, while for students with migrant backgrounds there is almost no difference in attitudes (0.06 index points) between those two groups of students. In Indonesia, for native students, non-repeaters tend to report more positive attitudes than secondary-school repeaters by 0.59 index points and primary-school repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.41 index points, while for students with migrant backgrounds both of these differences are smaller (0.32 index points and 0.13 index points, respectively). In Panama, for native students, non-repeaters tend to have more positive attitudes than do secondary-school repeaters with a difference of 0.21 index points and there is no difference in attitudes between primary- and secondary-school repeaters, while for students with migrant backgrounds non-repeaters tend to have better attitudes towards school than do secondary-school repeaters with a difference of 0.61 index points, and secondary-school repeaters tend to have better attitudes than do primary-school repeaters with a difference of 0.85 index points. In Qatar, for native students there is no difference in attitudes between non-repeaters and secondary-school repeaters, while for students with migrant backgrounds secondary-school repeaters tend to report more positive attitudes than do non-repeaters with a difference of 0.14 index points.

4.7.4. Heterogeneous relationships across age

Model 14 examines the relationships between attitudes towards school and repetition and how these relationships differ depending on the student's age. In general, the coefficients estimated for the different countries suggest that these relationships are similar between students who are at or above the country-average age and students who are below the country-average age. The exceptions are observed in six countries and economies. In Ireland, for older students secondary-repeaters tend to report more positive attitudes than do non-repeaters and primary-school repeaters (by 0.57 and 0.61 index points, respectively), while for younger students these differences are smaller (by 0.17 and 0.07 score points, respectively). In Argentina, for younger students there is no difference in attitudes between non-repeaters and secondary-school repeaters and between primary- and secondary-school repeaters, while for older students non-repeaters and primary-school repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.15 and 0.26 index points, respectively. In Dubai (UAE), for younger students, there is no significant relationship between grade repetition and students' attitudes towards school, while for older students, non-repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.32 index points. In Indonesia, for younger students primary-school repeaters tend to have more positive attitudes than do secondary-school repeaters with the difference of 0.39 index points, while for older students this difference is smaller (with the difference of 0.15 index points). In Macao-China, for younger students

there is no difference in attitudes between non-repeaters and secondary-school repeaters, while for older students non-repeaters tend to report more positive attitudes than do secondary-school repeaters by 0.11 index points. Finally, in Qatar, for younger students there is no difference in attitudes between primary-school repeaters and secondary-school repeaters, while for older students secondary-school repeaters tend to report more positive attitudes than do primary-school repeaters by 0.14 index points.

5. Conclusions

Using the OECD PISA 2009 dataset and a sample of 30 countries and economies with over 10% of 15-year-old students having repeated a grade at least once, this article analyses the relationships between grade retention and educational outcomes – performance and students' attitudes towards school. In this article, students are grouped according to the educational levels (primary or secondary school) at which they have repeated a grade. The results show that this distinction between grade repetitions at primary and at secondary education levels is important, because the extent of the relationship between grade repetition and educational outcomes differs according to whether students have repeated a grade in primary or secondary school. In most countries examined, students who repeated a grade in secondary school tend to perform better than do students who repeated a grade in primary school, and non-repeaters tend to perform even better than secondary-school repeaters. These findings remain unchanged, in general, even after controlling for student socio-economic background and a set of individual characteristics.

The results of this article contribute to shedding some light on the literature that has emerged with the availability of better (longitudinal) databases and some more appropriate quasi-experimental methods that allow for identifying the causal effect of retention. Our results are difficult to compare with other empirical findings because of existing differences in the statistical information available to all these analyses. For instance, Schwerdt and West (2012) conclude that there are some large short-term gains in achievement, which fade out over the grades, for students retained in third grade. While the timing (and the context and identification strategy) of their paper is very different from ours, our speculation is that, if measured in later grades, the faded-out initial positive effects could follow the negative trend and become negative. In addition to this, the somewhat narrower performance gaps that exist between secondary-school repeaters and non-repeaters in comparison to those between primary-school repeaters and non-repeaters found in this paper may indicate that either the profile of the repeaters is different between primary- and secondary-school repeaters or that repeaters at the later stage of their education career are still experiencing the "immediate" gains (or smaller losses) in achievement as identified by Schwerdt and West (2012), which may or may not fade out in later years.

In terms of the association between grade repetition and a measure of non-cognitive skills provided by PISA (the attitudes towards school index), the results show a more complex picture and the relationships vary across countries. In about one-third of countries and economies, primary-school repeaters tend to report more positive attitudes towards schools than do secondary-school repeaters. In 13 countries and economies, non-repeaters tend to report more positive attitudes than do primary-school repeaters. In about two-thirds of countries and economies examined, non-repeaters report more positive attitudes towards school than do secondary-school repeaters. These inconsistent patterns across countries might be related to the differences in how repeaters are treated across

different educational systems. Only in Belgium and Switzerland do primary-school repeaters tend to report more positive attitudes towards school than non-repeaters. Both these countries have highly selective education systems, where students are selected and sorted into different programmes at the age of 12, and four distinct educational programmes are available to 15-year-old students, while the OECD average is 14 years old for the first selection with 2.5 programmes available for 15-year-old students on average (OECD, 2010). Further examination is required to disentangle how the system-level characteristics affect the relationships between grade repetitions and students' attitudes towards school.

The analyses in this article incorporate some interaction terms of the grade repetition dummies and the students' characteristics. In the models with reading performance as an outcome, the interaction between grade repetition and student socio-economic background suggests differences in the relationships between repeaters and non-repeaters depending on their socio-economic background. In 18 countries and economies, the performance difference between non-repeaters and secondary-school repeaters is greater for socio-economically advantaged students than for socio-economically disadvantaged students. This may imply that for socio-economically advantaged students, whether a student repeated a grade in secondary school or not is mainly determined by their academic ability, while for socio-economically disadvantaged students, it is determined by other factors in addition to their academic ability. Some of these findings associated with these interactions can be considered good reasons to further investigate the consequences of grade retention within and across countries.

Analyses in this article are built on the hypothesis that academic and socio-emotional developments of students evolve differently over the school years, and accordingly, students' retention is likely to have different implications and meanings at different ages. The article also shows how these relationships differ across countries. In the future, the analyses in this paper can be extended to include some country-level variables in order to explain the variation of these relationships with country-level characteristics (similar to the framework used by Goos, Schreier, Eduard-Knipprath, De Fraine, Van Damme, & Trautwein, 2013). Instead of drawing on country-by-country regressions as done in this article, a (hierarchical) model that incorporates some variables describing the countries' repetition policies or their educational systems can be developed to examine the between-countries variation and to better assess when and for whom grade repetition is beneficial. Simultaneously, it will be important to keep trying to disentangle the positive and negative effects of grade repetition, in order to identify what type of supplementary services should be provided to retained students to sustain their performance in school and improve their educational outcomes, a puzzle which also has repercussions for the distribution of economic resources available for education.

Notes

1. These include the cost of providing an additional year of schooling to students who repeated a grade and the opportunity cost in delaying their entrance into the labour market by one year (see OECD, 2011, point 2).
2. See below for a discussion of the suitability of PISA to overcome this problem, i.e. on how to use valid between-grade comparison to assess the impact of retention.
3. Indeed, the real reasons why this would be happening, other than reducing the probability of being retained later, are not discussed.

4. The authors conclude that they do not provide definitive evidence that early grade retention is beneficial for students in the long run.
5. Lee and Barro (2001) study the importance of family background and school qualities in grade repetition rates at the country level.
6. Among the available country-level variables, only a few of them are statistically-significant predictors of the probability of grade repetition. Because of this, the authors note that retention practices are "only slightly affected by any national educational policy" (p. 77).
7. Questionnaires are available at *www.pisa.oecd.org*.
8. For the purposes of measuring trends, PISA 2009 reading was linked to the PISA 2000 reading scale using a six-step equating approach. The detail of this equating approach is described in *PISA 2009 Technical Report* (OECD, 2012). In PISA 2009, the OECD mean is 493, with a standard deviation of 93.
9. *PISA 2009 Technical Report* (OECD, 2012).

APPENDIX A

Tables

Table A.1. **Grade repetition**

	Repeated a grade							
	Never		At primary education only		At secondary education only		At both primary and secondary education	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD								
Australia	91.6	(0.4)	7.0	(0.3)	1.1	(0.1)	0.3	(0.1)
Austria	87.4	(0.9)	3.5	(0.4)	7.9	(0.6)	1.2	(0.2)
Belgium	65.1	(0.7)	13.4	(0.5)	18.2	(0.6)	3.4	(0.3)
Canada	91.6	(0.3)	3.8	(0.2)	4.0	(0.2)	0.5	(0.1)
Chile	76.6	(0.9)	8.1	(0.5)	14.2	(0.7)	1.2	(0.2)
Czech Republic	96.0	(0.3)	1.7	(0.2)	1.9	(0.3)	0.4	(0.1)
Denmark	95.6	(0.3)	3.4	(0.3)	0.8	(0.2)	0.2	(0.1)
Estonia	94.4	(0.6)	3.2	(0.3)	1.8	(0.3)	0.6	(0.2)
Finland	97.2	(0.3)	2.4	(0.2)	0.4	(0.1)	0.0	(0.0)
France	63.1	(1.0)	14.5	(0.9)	20.2	(1.0)	2.2	(0.3)
Germany	78.6	(0.7)	7.7	(0.5)	12.6	(0.6)	1.2	(0.2)
Greece	94.3	(0.8)	1.5	(0.4)	3.7	(0.7)	0.4	(0.1)
Hungary	88.9	(1.2)	4.0	(0.6)	5.0	(0.7)	2.0	(0.4)
Iceland	99.1	(0.2)	0.5	(0.1)	0.2	(0.1)	0.2	(0.1)
Ireland	88.0	(0.6)	10.3	(0.5)	1.0	(0.2)	0.6	(0.1)
Israel	92.5	(0.5)	1.9	(0.3)	3.2	(0.4)	2.4	(0.3)
Italy	84.0	(0.4)	0.6	(0.1)	15.0	(0.4)	0.3	(0.1)
Luxembourg	63.5	(0.5)	17.4	(0.5)	15.7	(0.5)	3.4	(0.2)
Mexico	78.5	(0.7)	15.8	(0.7)	5.0	(0.2)	0.8	(0.1)
Netherlands	73.3	(1.1)	21.3	(1.2)	4.7	(0.4)	0.6	(0.2)
New Zealand	94.9	(0.3)	3.1	(0.2)	1.2	(0.2)	0.7	(0.1)
Poland	94.7	(0.4)	1.4	(0.2)	3.4	(0.3)	0.4	(0.1)
Portugal	65.0	(1.9)	15.7	(1.0)	15.0	(0.9)	4.4	(0.5)
Slovak Republic	96.2	(0.5)	1.7	(0.3)	1.9	(0.3)	0.2	(0.1)
Slovenia	98.5	(0.4)	0.0	c	1.5	(0.4)	0.0	c
Spain	64.7	(0.7)	4.2	(0.2)	24.9	(0.6)	6.3	(0.3)
Sweden	95.4	(0.4)	3.1	(0.3)	0.8	(0.1)	0.7	(0.1)
Switzerland	77.2	(0.9)	13.5	(0.7)	8.5	(0.5)	0.8	(0.1)
Turkey	87.0	(0.8)	3.4	(0.5)	9.6	(0.7)	0.1	(0.0)
United Kingdom	97.8	(0.2)	1.2	(0.1)	0.6	(0.1)	0.4	(0.1)
United States	85.8	(0.9)	9.6	(0.7)	3.2	(0.4)	1.3	(0.2)

Table A.1. **Grade repetition** *(cont.)*

	Repeated a grade							
	Never		At primary education only		At secondary education only		At both primary and secondary education	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Partners								
Albania	95.3	(0.5)	1.7	(0.3)	1.8	(0.3)	1.2	(0.2)
Argentina	66.2	(1.6)	9.7	(0.6)	20.0	(1.3)	4.0	(0.4)
Azerbaijan	98.3	(0.3)	0.6	(0.2)	0.9	(0.2)	0.2	(0.1)
Brazil	59.9	(0.9)	12.6	(0.5)	22.1	(0.6)	5.5	(0.3)
Bulgaria	94.4	(0.5)	1.2	(0.2)	2.9	(0.4)	1.5	(0.2)
Colombia	66.1	(1.1)	16.9	(1.0)	14.2	(0.7)	2.9	(0.3)
Croatia	97.2	(0.2)	0.1	(0.1)	1.6	(0.2)	1.1	(0.2)
Dubai (UAE)	87.4	(0.3)	5.8	(0.3)	5.8	(0.3)	1.1	(0.1)
Hong Kong-China	84.4	(0.7)	9.5	(0.5)	5.3	(0.5)	0.8	(0.1)
Indonesia	82.0	(1.2)	12.4	(1.1)	1.7	(0.2)	3.8	(0.4)
Jordan	93.4	(0.5)	1.7	(0.2)	2.7	(0.3)	2.1	(0.2)
Kazakhstan	98.3	(0.2)	1.0	(0.2)	0.3	(0.1)	0.4	(0.1)
Kyrgyzstan	95.7	(0.4)	1.7	(0.2)	1.4	(0.2)	1.2	(0.2)
Latvia	88.9	(1.0)	4.9	(0.5)	5.3	(0.8)	0.9	(0.2)
Liechtenstein	78.5	(2.0)	8.4	(1.5)	11.9	(1.7)	1.1	(0.6)
Lithuania	96.1	(0.4)	1.7	(0.2)	1.8	(0.2)	0.3	(0.1)
Macao-China	56.3	(0.5)	13.3	(0.4)	23.0	(0.5)	7.4	(0.3)
Montenegro	98.2	(0.5)	0.4	(0.2)	1.2	(0.4)	0.2	(0.0)
Panama	68.2	(2.7)	11.0	(1.3)	16.0	(1.5)	4.8	(0.7)
Peru	71.9	(1.2)	16.0	(0.8)	8.0	(0.6)	4.1	(0.3)
Qatar	85.2	(0.3)	5.8	(0.2)	6.6	(0.2)	2.5	(0.2)
Roumania	95.8	(0.7)	1.5	(0.4)	1.8	(0.4)	0.8	(0.2)
Russian Federation	96.8	(0.5)	2.0	(0.3)	1.0	(0.2)	0.3	(0.1)
Serbia	98.0	(0.5)	0.3	(0.1)	1.6	(0.4)	0.1	(0.0)
Shanghai-China	92.5	(0.7)	4.0	(0.5)	3.1	(0.5)	0.3	(0.1)
Singapore	94.6	(0.3)	1.9	(0.2)	3.1	(0.2)	0.4	(0.1)
Chinese Taipei	98.4	(0.2)	0.5	(0.1)	0.7	(0.1)	0.4	(0.1)
Thailand	96.5	(0.3)	2.1	(0.2)	1.2	(0.2)	0.2	(0.1)
Trinidad and Tobago	71.2	(0.6)	25.7	(0.6)	1.8	(0.2)	1.4	(0.2)
Tunisia	56.8	(1.2)	11.2	(0.8)	22.4	(0.9)	9.6	(0.5)
Uruguay	62.0	(1.1)	13.4	(0.6)	17.8	(0.7)	6.8	(0.4)

Table A.2. **Student performance in reading**

	All students				Gender differences						Percentiles											
	Mean score		Standard deviation		Boys		Girls		Difference (B – G)		5th		10th		25th		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Mean score	S.E.	Mean score	S.E.	Score dif.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
OECD																						
Austria	470	(2.9)	100	(2.0)	449	(3.8)	490	(4.0)	**-41**	(5.5)	299	(5.2)	334	(6.1)	399	(4.3)	545	(3.3)	596	(3.4)	625	(4.3)
Belgium	506	(2.3)	102	(1.7)	493	(3.4)	520	(2.9)	**-27**	(4.4)	326	(6.1)	368	(4.3)	436	(3.8)	583	(2.2)	631	(2.7)	657	(2.9)
Chile	449	(3.1)	83	(1.7)	439	(3.9)	461	(3.6)	**-22**	(4.1)	310	(5.1)	342	(5.0)	393	(4.1)	506	(3.3)	556	(3.6)	584	(5.1)
France	496	(3.4)	106	(2.8)	475	(4.3)	515	(3.4)	**-40**	(3.7)	305	(8.2)	352	(7.0)	429	(4.7)	572	(4.0)	624	(3.9)	651	(4.6)
Germany	497	(2.7)	95	(1.8)	478	(3.6)	518	(2.9)	**-40**	(3.9)	333	(4.8)	367	(5.1)	432	(4.5)	567	(2.8)	615	(3.2)	640	(3.1)
Hungary	494	(3.2)	90	(2.4)	475	(3.9)	513	(3.6)	**-38**	(4.0)	332	(7.4)	371	(6.9)	435	(4.3)	559	(3.6)	607	(3.5)	632	(4.0)
Ireland	496	(3.0)	95	(2.2)	476	(4.2)	515	(3.1)	**-39**	(4.7)	330	(7.8)	373	(4.7)	435	(3.9)	562	(2.8)	611	(2.8)	638	(3.2)
Italy	486	(1.6)	96	(1.4)	464	(2.3)	510	(1.9)	**-46**	(2.8)	320	(3.7)	358	(2.6)	422	(2.3)	556	(1.7)	604	(1.7)	631	(2.1)
Luxembourg	472	(1.3)	104	(0.9)	453	(1.9)	492	(1.5)	**-39**	(2.3)	288	(3.6)	332	(3.5)	403	(2.4)	547	(1.7)	600	(2.0)	630	(3.7)
Mexico	425	(2.0)	85	(1.2)	413	(2.1)	438	(2.1)	**-25**	(1.6)	281	(3.9)	314	(2.9)	370	(2.4)	485	(1.9)	531	(2.2)	557	(2.4)
Netherlands	508	(5.1)	89	(1.6)	496	(5.1)	521	(5.3)	**-24**	(2.4)	365	(4.7)	390	(5.0)	442	(6.1)	575	(5.4)	625	(4.6)	650	(4.0)
Portugal	489	(3.1)	87	(1.6)	470	(3.5)	508	(2.9)	**-38**	(2.4)	338	(4.8)	373	(4.9)	432	(4.4)	551	(3.4)	599	(3.5)	624	(3.6)
Spain	481	(2.0)	88	(1.1)	467	(2.2)	496	(2.2)	**-29**	(2.0)	326	(4.2)	364	(3.5)	426	(3.3)	543	(2.0)	588	(2.0)	613	(2.4)
Switzerland	501	(2.4)	93	(1.4)	481	(2.9)	520	(2.7)	**-39**	(2.5)	337	(4.1)	374	(4.0)	437	(3.6)	569	(3.0)	617	(3.3)	645	(4.4)
Turkey	464	(3.5)	82	(1.7)	443	(3.7)	486	(4.1)	**-43**	(3.7)	325	(5.1)	356	(4.3)	409	(3.8)	522	(4.5)	569	(5.2)	596	(5.4)
United States	500	(3.7)	97	(1.6)	488	(4.2)	513	(3.8)	**-25**	(3.4)	339	(4.2)	372	(3.9)	433	(4.0)	569	(4.6)	625	(5.0)	656	(5.8)
Partners																						
Argentina	398	(4.6)	108	(3.4)	379	(5.1)	415	(4.9)	**-37**	(3.8)	209	(11.3)	257	(8.3)	329	(5.8)	473	(6.3)	535	(7.1)	568	(6.7)
Brazil	412	(2.7)	94	(1.5)	397	(2.9)	425	(2.8)	**-29**	(1.7)	262	(3.0)	293	(3.2)	348	(2.7)	474	(3.9)	537	(4.2)	572	(4.6)
Colombia	413	(3.7)	87	(1.9)	408	(4.5)	418	(4.0)	**-9**	(3.8)	269	(6.4)	302	(5.2)	355	(4.4)	473	(3.9)	524	(4.1)	554	(4.0)
Dubai (UAE)	459	(1.1)	107	(0.9)	435	(1.7)	485	(1.5)	**-51**	(2.3)	277	(3.4)	317	(2.8)	386	(2.4)	536	(2.4)	596	(2.7)	628	(3.1)
Hong Kong-China	533	(2.1)	84	(1.7)	518	(3.3)	550	(2.8)	**-33**	(4.4)	380	(5.5)	418	(4.5)	482	(3.0)	592	(2.5)	634	(2.9)	659	(3.1)
Indonesia	402	(3.7)	66	(2.0)	383	(3.8)	420	(3.9)	**-37**	(3.3)	291	(5.8)	315	(5.0)	357	(4.1)	447	(4.6)	487	(5.0)	510	(5.8)
Latvia	484	(3.0)	80	(1.5)	460	(3.4)	507	(3.1)	**-47**	(3.2)	348	(6.3)	379	(4.2)	429	(3.8)	541	(3.3)	584	(3.2)	610	(4.3)
Macao-China	487	(0.9)	76	(0.8)	470	(1.3)	504	(1.2)	**-34**	(1.7)	357	(2.7)	388	(1.8)	437	(1.4)	540	(1.4)	582	(1.8)	608	(1.8)
Panama	371	(6.5)	99	(3.5)	354	(7.0)	387	(7.3)	**-33**	(6.7)	209	(12.0)	246	(10.0)	304	(7.4)	436	(7.7)	502	(9.3)	540	(10.0)
Peru	370	(4.0)	98	(2.4)	359	(4.2)	381	(4.9)	**-22**	(4.7)	209	(5.0)	241	(3.9)	302	(4.3)	437	(5.2)	496	(6.4)	530	(7.0)
Qatar	372	(0.8)	115	(0.8)	347	(1.3)	397	(1.0)	**-50**	(1.8)	196	(2.4)	228	(2.2)	288	(1.4)	450	(1.4)	529	(2.1)	573	(2.8)
Trinidad and Tobago	416	(1.2)	113	(1.3)	387	(1.9)	445	(1.6)	**-58**	(2.5)	220	(5.8)	265	(3.9)	339	(2.5)	496	(2.3)	559	(2.5)	594	(3.0)
Tunisia	404	(2.9)	85	(1.8)	387	(3.2)	418	(3.0)	**-31**	(2.2)	258	(4.4)	293	(3.8)	348	(3.4)	462	(3.4)	510	(4.8)	538	(5.2)
Uruguay	426	(2.6)	99	(1.9)	404	(3.2)	445	(2.8)	**-42**	(3.1)	257	(5.2)	297	(4.2)	359	(3.4)	495	(3.1)	552	(3.3)	584	(4.5)

Note: Values that are statistically significant are indicated in bold.
Source: PISA 2009 Results, Vol. 1 (OECD, 2010).

Table A.3. **Students' attitudes towards school**

	Index of attitude towards school		School has done little to prepare me for adult life when I leave school								School has been a waste of time							
			Strongly disagree		Disagree		Agree		Strongly agree		Strongly disagree		Disagree		Agree		Strongly agree	
	Index points	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD																		
Austria	0.06	(0.02)	34.9	(0.9)	39.1	(1.0)	20.7	(0.7)	5.3	(0.4)	60.4	(1.0)	29.9	(0.9)	5.8	(0.4)	3.9	(0.3)
Belgium	-0.11	(0.01)	28.8	(0.6)	52.6	(0.6)	15.0	(0.5)	3.7	(0.2)	39.3	(0.7)	52.1	(0.7)	6.1	(0.3)	2.5	(0.2)
Chile	0.29	(0.02)	36.2	(1.1)	39.9	(0.8)	17.7	(0.7)	6.2	(0.3)	64.8	(0.8)	30.7	(0.8)	2.5	(0.2)	1.9	(0.2)
France	0.07	(0.02)	36.5	(0.8)	44.1	(0.8)	15.4	(0.7)	4.0	(0.3)	47.2	(1.0)	44.3	(0.9)	5.9	(0.4)	2.5	(0.3)
Germany	-0.11	(0.02)	25.7	(0.8)	38.9	(0.8)	27.1	(0.7)	8.4	(0.5)	58.7	(0.8)	32.9	(0.7)	5.5	(0.3)	2.9	(0.2)
Hungary	-0.04	(0.02)	35.5	(1.1)	48.9	(1.1)	12.7	(0.6)	2.9	(0.3)	43.8	(0.9)	48.1	(0.9)	5.5	(0.4)	2.7	(0.3)
Ireland	0.14	(0.02)	35.4	(0.8)	45.3	(0.8)	14.1	(0.6)	5.1	(0.4)	47.2	(0.9)	45.5	(0.9)	5.4	(0.4)	2.0	(0.3)
Italy	0.03	(0.01)	27.3	(0.4)	52.2	(0.4)	16.1	(0.3)	4.3	(0.2)	47.5	(0.5)	45.1	(0.4)	4.9	(0.2)	2.5	(0.1)
Luxembourg	-0.13	(0.01)	31.4	(0.7)	36.4	(0.7)	23.8	(0.7)	8.4	(0.4)	54.2	(0.7)	36.4	(0.6)	5.8	(0.4)	3.6	(0.3)
Mexico	0.43	(0.01)	35.8	(0.4)	30.5	(0.4)	23.0	(0.5)	10.6	(0.3)	71.7	(0.5)	23.7	(0.4)	2.5	(0.2)	2.1	(0.1)
Netherlands	-0.22	(0.02)	17.1	(0.7)	61.5	(0.8)	18.1	(0.7)	3.3	(0.3)	29.1	(0.9)	57.4	(1.0)	9.1	(0.5)	4.4	(0.5)
Portugal	0.36	(0.02)	47.6	(1.1)	39.1	(0.8)	9.9	(0.5)	3.5	(0.3)	57.2	(0.9)	38.6	(0.8)	3.4	(0.3)	0.9	(0.1)
Spain	0.17	(0.02)	37.0	(0.9)	42.2	(0.7)	15.5	(0.4)	5.3	(0.3)	55.9	(0.8)	37.1	(0.8)	4.7	(0.2)	2.3	(0.2)
Switzerland	0.01	(0.02)	33.1	(0.7)	42.9	(0.5)	19.0	(0.6)	5.0	(0.4)	52.6	(0.9)	38.2	(0.8)	5.9	(0.3)	3.3	(0.3)
Turkey	0.20	(0.02)	33.1	(0.8)	35.2	(0.7)	21.2	(0.8)	10.5	(0.5)	60.6	(0.9)	32.8	(0.8)	4.5	(0.3)	2.0	(0.3)
United States	0.22	(0.02)	40.1	(0.9)	37.6	(0.8)	16.6	(0.7)	5.7	(0.3)	52.2	(1.0)	39.6	(0.9)	5.9	(0.4)	2.3	(0.2)
Partners																		
Argentina	0.11	(0.02)	27.6	(0.8)	37.2	(0.8)	25.5	(0.8)	9.6	(0.6)	54.5	(1.0)	38.2	(0.8)	4.2	(0.3)	3.1	(0.3)
Brazil	0.43	(0.02)	44.2	(0.8)	40.7	(0.6)	11.0	(0.4)	4.1	(0.3)	64.2	(0.8)	32.6	(0.7)	1.9	(0.1)	1.3	(0.2)
Colombia	0.46	(0.02)	47.2	(1.0)	31.6	(0.8)	13.1	(0.6)	8.1	(0.6)	68.9	(1.1)	27.0	(0.9)	2.2	(0.3)	1.9	(0.2)
Dubai (UAE)	0.07	(0.01)	26.0	(0.6)	34.3	(0.7)	28.3	(0.7)	11.4	(0.4)	56.9	(0.7)	34.7	(0.7)	5.5	(0.3)	2.8	(0.2)
Hong Kong-China	-0.49	(0.01)	6.6	(0.4)	52.3	(0.7)	35.2	(0.6)	5.8	(0.4)	34.5	(0.8)	55.9	(0.8)	7.2	(0.4)	2.4	(0.2)
Indonesia	0.50	(0.02)	32.6	(1.0)	45.3	(1.0)	18.0	(0.6)	4.1	(0.4)	69.0	(0.9)	28.5	(0.9)	1.7	(0.2)	0.8	(0.2)
Latvia	0.19	(0.02)	37.4	(1.1)	49.4	(1.0)	10.4	(0.6)	2.8	(0.3)	49.5	(1.1)	45.5	(1.0)	3.3	(0.3)	1.6	(0.2)
Macao-China	-0.45	(0.01)	9.5	(0.4)	41.3	(0.8)	39.0	(0.6)	10.3	(0.4)	34.7	(0.6)	56.6	(0.7)	6.5	(0.3)	2.2	(0.2)
Panama	0.38	(0.03)	38.7	(1.2)	30.6	(1.6)	19.0	(1.0)	11.8	(0.9)	67.1	(1.4)	23.8	(0.9)	5.6	(0.9)	3.6	(0.5)
Peru	0.18	(0.02)	25.3	(1.0)	41.7	(0.9)	26.3	(0.9)	6.7	(0.4)	56.8	(1.0)	33.9	(0.8)	5.8	(0.5)	3.6	(0.3)
Qatar	-0.13	(0.01)	26.2	(0.5)	28.5	(0.4)	33.6	(0.5)	11.7	(0.4)	45.9	(0.6)	37.5	(0.5)	11.5	(0.3)	5.2	(0.2)
Trinidad and Tobago	0.47	(0.01)	41.8	(0.7)	29.4	(0.8)	17.3	(0.6)	11.5	(0.5)	64.9	(0.7)	30.9	(0.6)	2.8	(0.3)	1.4	(0.2)
Tunisia	0.51	(0.02)	38.1	(1.0)	36.3	(0.8)	19.7	(0.9)	5.9	(0.5)	70.2	(0.8)	24.8	(0.8)	3.3	(0.3)	1.7	(0.2)
Uruguay	0.16	(0.02)	36.4	(0.8)	40.8	(0.8)	16.5	(0.6)	6.4	(0.4)	60.5	(0.8)	33.2	(0.7)	4.0	(0.3)	2.3	(0.2)

Table A.3. **Students' attitudes towards school** *(cont.)*

	School has helped give me confidence to make decisions								School has taught me things which could be useful in a job							
	Strongly disagree		Disagree		Agree		Strongly agree		Strongly disagree		Disagree		Agree		Strongly agree	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD																
Austria	10.6	(0.5)	24.2	(0.6)	49.3	(0.9)	15.9	(0.8)	5.3	(0.4)	7.3	(0.5)	32.8	(0.9)	54.6	(1.1)
Belgium	7.0	(0.3)	23.8	(0.5)	57.9	(0.6)	11.4	(0.4)	2.6	(0.2)	7.0	(0.3)	51.4	(0.7)	39.0	(0.7)
Chile	5.6	(0.4)	13.1	(0.6)	53.3	(0.8)	28.1	(0.8)	3.9	(0.3)	5.2	(0.3)	36.4	(0.8)	54.4	(0.8)
France	8.5	(0.4)	20.1	(0.7)	55.0	(0.9)	16.4	(0.6)	2.4	(0.2)	4.5	(0.4)	42.9	(0.9)	50.1	(0.9)
Germany	12.1	(0.5)	28.3	(0.6)	45.5	(0.9)	14.0	(0.6)	3.2	(0.2)	7.5	(0.4)	40.1	(0.8)	49.2	(0.8)
Hungary	6.3	(0.5)	22.5	(0.9)	59.9	(0.9)	11.3	(0.6)	2.9	(0.3)	7.6	(0.5)	53.6	(0.9)	35.8	(1.0)
Ireland	4.6	(0.4)	14.3	(0.6)	59.7	(0.8)	21.5	(0.7)	3.1	(0.3)	6.5	(0.4)	44.8	(1.0)	45.6	(0.9)
Italy	5.5	(0.2)	16.4	(0.3)	61.3	(0.3)	16.8	(0.3)	3.3	(0.1)	6.5	(0.2)	43.6	(0.4)	46.7	(0.5)
Luxembourg	13.9	(0.6)	27.8	(0.8)	44.5	(0.7)	13.8	(0.5)	4.2	(0.3)	8.8	(0.4)	42.4	(0.7)	44.6	(0.7)
Mexico	3.1	(0.2)	4.8	(0.2)	51.9	(0.4)	40.2	(0.5)	2.8	(0.1)	2.6	(0.2)	31.5	(0.4)	63.1	(0.4)
Netherlands	2.9	(0.3)	14.9	(0.7)	63.8	(0.9)	18.4	(0.7)	1.9	(0.3)	10.3	(0.6)	58.6	(0.9)	29.1	(1.0)
Portugal	2.3	(0.2)	9.1	(0.4)	66.2	(0.8)	22.3	(0.8)	1.3	(0.2)	5.4	(0.3)	47.1	(0.8)	46.3	(0.9)
Spain	5.3	(0.3)	14.3	(0.5)	60.6	(0.6)	19.8	(0.6)	3.3	(0.2)	6.3	(0.3)	43.7	(0.6)	46.6	(0.7)
Switzerland	10.6	(0.5)	24.9	(0.7)	47.5	(0.8)	17.0	(0.5)	4.7	(0.3)	8.1	(0.4)	40.5	(0.7)	46.8	(0.9)
Turkey	4.3	(0.3)	10.2	(0.5)	51.9	(0.7)	33.6	(0.9)	3.8	(0.3)	9.2	(0.5)	40.3	(0.8)	46.7	(1.1)
United States	4.3	(0.3)	12.9	(0.6)	56.0	(0.9)	26.7	(0.9)	3.0	(0.3)	5.8	(0.4)	45.1	(1.0)	46.1	(1.0)
Partners																
Argentina	4.4	(0.3)	12.2	(0.7)	59.1	(0.9)	24.4	(0.8)	4.2	(0.3)	3.4	(0.3)	36.7	(0.9)	55.7	(1.1)
Brazil	3.5	(0.3)	10.0	(0.4)	57.2	(0.6)	29.2	(0.6)	2.7	(0.3)	3.6	(0.2)	39.2	(0.6)	54.6	(0.7)
Colombia	5.7	(0.4)	8.7	(0.5)	54.0	(0.8)	31.6	(1.0)	3.4	(0.3)	3.9	(0.3)	32.8	(0.9)	59.9	(1.1)
Dubai (UAE)	4.7	(0.3)	11.0	(0.4)	53.0	(0.7)	31.3	(0.6)	4.5	(0.3)	7.3	(0.4)	42.4	(0.7)	45.9	(0.7)
Hong Kong-China	4.2	(0.3)	25.8	(0.6)	62.7	(0.7)	7.3	(0.5)	3.2	(0.3)	12.0	(0.5)	69.4	(0.8)	15.5	(0.7)
Indonesia	2.8	(0.3)	3.0	(0.3)	50.0	(1.1)	44.2	(1.1)	1.7	(0.2)	1.8	(0.2)	35.2	(0.9)	61.3	(0.9)
Latvia	4.0	(0.4)	15.1	(0.6)	66.3	(1.0)	14.6	(0.6)	3.0	(0.3)	4.4	(0.4)	39.7	(0.9)	52.9	(1.0)
Macao-China	4.4	(0.3)	22.4	(0.5)	61.9	(0.6)	11.3	(0.4)	3.2	(0.2)	9.8	(0.4)	63.8	(0.5)	23.2	(0.5)
Panama	5.7	(0.7)	8.1	(0.8)	49.7	(1.6)	36.4	(1.4)	3.9	(0.6)	3.8	(0.5)	29.3	(1.4)	63.1	(1.6)
Peru	3.2	(0.3)	7.2	(0.4)	57.0	(0.7)	32.7	(0.8)	3.4	(0.3)	4.4	(0.3)	42.8	(0.8)	49.4	(1.0)
Qatar	9.2	(0.3)	16.0	(0.4)	47.7	(0.5)	27.1	(0.5)	7.3	(0.3)	9.5	(0.3)	39.4	(0.5)	43.8	(0.5)
Trinidad and Tobago	3.0	(0.3)	6.2	(0.4)	48.4	(0.8)	42.5	(0.8)	2.6	(0.2)	3.0	(0.3)	32.4	(0.6)	62.0	(0.6)
Tunisia	3.6	(0.4)	5.5	(0.4)	41.1	(0.8)	49.8	(1.0)	3.0	(0.3)	4.0	(0.3)	31.0	(0.8)	62.0	(0.9)
Uruguay	4.9	(0.3)	13.6	(0.5)	61.2	(0.6)	20.4	(0.6)	3.2	(0.3)	4.7	(0.3)	48.8	(0.8)	43.3	(0.8)

Table A.4. **Descriptive statistics of students' background characteristics**

	PISA index of economic, social and cultural status (ESCS)				Gender (MALE)				Migrant background (MIGLAN)						Student age (rAGE)			
	Average ESCS		Standard Deviation		Boys (MALE = 1)		Girls (MALE = 0)		Students with migrant background or students who speak language at home different from the language of assessment (MIGLAN=1)		Others (MIGLAN = 0)		Student age (AGE)		At or above the country average age		Below the country average age	
	Mean index	S.E.	Std. Dev.	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	Mean age	S.E.	%	S.E.	%	S.E.
OECD																		
Austria	0.06	(0.02)	0.84	(0.01)	49.0	(1.83)	51.0	(1.83)	16.5	(1.24)	83.5	(1.24)	15.81	(0.01)	49.2	(0.82)	50.8	(0.82)
Belgium	0.20	(0.02)	0.93	(0.01)	51.1	(1.22)	48.9	(1.22)	28.4	(1.35)	71.6	(1.35)	15.85	(0.00)	53.0	(0.68)	47.0	(0.68)
Chile	-0.57	(0.04)	1.14	(0.02)	51.0	(1.12)	49.0	(1.12)	0.9	(0.14)	99.1	(0.14)	15.79	(0.00)	50.3	(0.62)	49.7	(0.62)
France	-0.13	(0.03)	0.84	(0.02)	48.7	(1.19)	51.3	(1.19)	15.3	(1.37)	84.7	(1.37)	15.86	(0.00)	51.9	(0.87)	48.1	(0.87)
Germany	0.18	(0.02)	0.90	(0.01)	51.1	(0.97)	48.9	(0.97)	19.0	(1.03)	81.0	(1.03)	15.83	(0.01)	47.0	(0.79)	53.0	(0.79)
Hungary	-0.20	(0.03)	0.97	(0.02)	50.4	(1.51)	49.6	(1.51)	2.9	(0.40)	97.1	(0.40)	15.72	(0.00)	50.4	(0.90)	49.6	(0.90)
Ireland	0.05	(0.03)	0.85	(0.01)	50.6	(1.08)	49.4	(1.08)	10.3	(0.97)	89.7	(0.97)	15.71	(0.00)	50.8	(0.83)	49.2	(0.83)
Italy	-0.12	(0.01)	1.02	(0.01)	51.4	(0.93)	48.6	(0.93)	15.1	(0.39)	84.9	(0.39)	15.73	(0.00)	48.5	(0.45)	51.5	(0.45)
Luxembourg	0.19	(0.01)	1.10	(0.01)	50.7	(0.16)	49.3	(0.16)	93.6	(0.47)	6.4	(0.47)	15.83	(0.00)	46.6	(0.58)	53.4	(0.58)
Mexico	-1.22	(0.03)	1.30	(0.01)	49.4	(0.44)	50.6	(0.44)	4.2	(0.32)	95.8	(0.32)	15.68	(0.00)	54.2	(0.44)	45.8	(0.44)
Netherlands	0.27	(0.03)	0.86	(0.02)	49.7	(0.70)	50.3	(0.70)	13.8	(1.53)	86.2	(1.53)	15.72	(0.00)	48.6	(0.67)	51.4	(0.67)
Portugal	-0.32	(0.04)	1.18	(0.02)	48.9	(0.62)	51.1	(0.62)	6.0	(0.46)	94.0	(0.46)	15.75	(0.01)	46.0	(0.82)	54.0	(0.82)
Spain	-0.31	(0.03)	1.09	(0.01)	50.8	(0.57)	49.2	(0.57)	23.3	(1.04)	76.7	(1.04)	15.85	(0.00)	52.8	(0.50)	47.2	(0.50)
Switzerland	0.08	(0.02)	0.88	(0.01)	50.8	(1.14)	49.2	(1.14)	26.5	(0.97)	73.5	(0.97)	15.85	(0.01)	53.6	(0.90)	46.4	(0.90)
Turkey	-1.16	(0.05)	1.22	(0.02)	51.6	(1.71)	48.4	(1.71)	4.4	(0.58)	95.6	(0.58)	15.82	(0.00)	47.4	(0.74)	52.6	(0.74)
United States	0.17	(0.04)	0.93	(0.02)	51.3	(0.75)	48.7	(0.75)	20.9	(1.40)	79.1	(1.40)	15.79	(0.01)	49.6	(1.05)	50.4	(1.05)
Partners																		
Argentina	-0.62	(0.05)	1.19	(0.03)	46.3	(1.06)	53.7	(1.06)	4.4	(0.50)	95.6	(0.50)	15.70	(0.01)	51.4	(0.94)	48.6	(0.94)
Brazil	-1.16	(0.03)	1.21	(0.01)	46.9	(0.38)	53.1	(0.38)	1.4	(0.16)	98.6	(0.16)	15.86	(0.00)	51.1	(0.61)	48.9	(0.61)
Colombia	-1.15	(0.05)	1.27	(0.02)	47.6	(1.20)	52.4	(1.20)	0.7	(0.10)	99.3	(0.10)	15.85	(0.01)	53.7	(1.10)	46.3	(1.10)
Dubai (UAE)	0.42	(0.01)	0.79	(0.01)	51.1	(0.14)	48.9	(0.14)	78.5	(0.40)	21.5	(0.40)	15.82	(0.00)	49.0	(0.58)	51.0	(0.58)
Hong Kong-China	-0.80	(0.04)	1.02	(0.02)	52.9	(1.76)	47.1	(1.76)	41.9	(1.53)	58.1	(1.53)	15.75	(0.00)	47.4	(0.66)	52.6	(0.66)
Indonesia	-1.55	(0.06)	1.10	(0.02)	49.5	(1.95)	50.5	(1.95)	64.0	(2.08)	36.0	(2.08)	15.76	(0.01)	54.0	(1.13)	46.0	(1.13)
Latvia	-0.13	(0.03)	0.88	(0.01)	49.3	(0.87)	50.7	(0.87)	12.7	(1.30)	87.3	(1.30)	15.77	(0.01)	51.0	(0.83)	49.0	(0.83)
Macao-China	-0.70	(0.01)	0.87	(0.01)	50.6	(0.09)	49.4	(0.09)	74.9	(0.62)	25.1	(0.62)	15.78	(0.00)	53.4	(0.60)	46.6	(0.60)
Panama	-0.81	(0.08)	1.33	(0.04)	49.6	(1.43)	50.4	(1.43)	7.8	(1.58)	92.2	(1.58)	15.84	(0.01)	56.0	(1.63)	44.0	(1.63)
Peru	-1.31	(0.05)	1.25	(0.03)	50.5	(1.21)	49.5	(1.21)	5.5	(0.83)	94.5	(0.83)	15.79	(0.00)	50.8	(0.71)	49.2	(0.71)
Qatar	0.51	(0.01)	0.91	(0.01)	50.9	(0.11)	49.1	(0.11)	63.8	(0.39)	36.2	(0.39)	15.74	(0.00)	47.2	(0.24)	52.8	(0.24)
Trinidad and Tobago	-0.58	(0.02)	0.93	(0.01)	49.5	(0.22)	50.5	(0.22)	4.6	(0.29)	95.4	(0.29)	15.68	(0.01)	54.3	(0.89)	45.7	(0.89)
Tunisia	-1.20	(0.05)	1.31	(0.02)	47.6	(0.48)	52.4	(0.48)	0.4	(0.11)	99.6	(0.11)	15.88	(0.00)	49.3	(0.83)	50.7	(0.83)
Uruguay	-0.70	(0.03)	1.22	(0.02)	47.0	(0.69)	53.0	(0.69)	2.7	(0.24)	97.3	(0.24)	15.86	(0.00)	51.0	(0.81)	49.0	(0.81)

Table A.5. **Student socio-economic background by grade repetition**

	Student socio-economic background, by repeated grade								Difference in student socio-economic background between students who have repeated a grade at:							
	Never		At primary education only		At secondary education only		At both primary and secondary education		Primary education only and Never (Primary education only – Never)		Secondary education only and Never (Secondary education only – Never)		Both primary and secondary education and Never (Both primary and secondary education – Never)		Primary education only and Secondary education only (Secondary education only – Primary education only)	
	Index points	S.E.	Index points	S.E.	Index points	S.E.	Index points	S.E.	Index dif.	S.E.	Index dif.	S.E.	Index dif.	S.E.	Index dif.	S.E.
OECD																
Austria	0.13	(0.02)	-0.38	(0.07)	0.01	(0.04)	-0.26	(0.12)	**-0.51**	(0.07)	**-0.12**	(0.05)	**-0.39**	(0.12)	**0.39**	(0.07)
Belgium	0.40	(0.02)	-0.31	(0.04)	0.08	(0.03)	-0.34	(0.06)	**-0.71**	(0.04)	**-0.31**	(0.03)	**-0.74**	(0.06)	**0.39**	(0.05)
Chile	-0.46	(0.03)	-1.11	(0.08)	-0.81	(0.06)	-0.98	(0.13)	**-0.65**	(0.07)	**-0.36**	(0.05)	**-0.52**	(0.14)	**0.29**	(0.08)
France	0.05	(0.04)	-0.66	(0.04)	-0.28	(0.05)	-0.57	(0.10)	**-0.71**	(0.05)	**-0.33**	(0.06)	**-0.62**	(0.10)	**0.38**	(0.06)
Germany	0.28	(0.02)	-0.32	(0.06)	0.12	(0.04)	-0.44	(0.17)	**-0.60**	(0.06)	**-0.16**	(0.05)	**-0.72**	(0.17)	**0.44**	(0.07)
Hungary	-0.09	(0.03)	-1.16	(0.12)	-0.92	(0.12)	-0.82	(0.16)	**-1.06**	(0.12)	**-0.82**	(0.12)	**-0.73**	(0.16)	0.24	(0.15)
Ireland	0.07	(0.03)	-0.09	(0.05)	-0.31	(0.12)	c	c	**-0.16**	(0.05)	**-0.38**	(0.12)	c	c	-0.22	(0.13)
Italy	-0.06	(0.01)	-0.97	(0.12)	-0.40	(0.03)	-1.27	(0.18)	**-0.91**	(0.11)	**-0.34**	(0.03)	**-1.21**	(0.18)	**0.57**	(0.12)
Luxembourg	0.38	(0.02)	-0.44	(0.04)	0.21	(0.04)	-0.25	(0.09)	**-0.83**	(0.04)	**-0.17**	(0.05)	**-0.63**	(0.10)	**0.66**	(0.05)
Mexico	-1.09	(0.03)	-1.89	(0.04)	-0.99	(0.06)	-1.52	(0.09)	**-0.80**	(0.05)	0.09	(0.06)	**-0.44**	(0.09)	**0.90**	(0.07)
Netherlands	0.35	(0.03)	0.09	(0.05)	0.39	(0.05)	c	c	**-0.26**	(0.04)	0.04	(0.05)	c	c	**0.30**	(0.07)
Portugal	-0.05	(0.04)	-0.91	(0.04)	-0.67	(0.05)	-0.95	(0.08)	**-0.86**	(0.06)	**-0.62**	(0.07)	**-0.90**	(0.09)	**0.25**	(0.06)
Spain	-0.05	(0.03)	-0.72	(0.06)	-0.73	(0.03)	-1.07	(0.06)	**-0.67**	(0.05)	**-0.68**	(0.03)	**-1.02**	(0.06)	-0.01	(0.06)
Switzerland	0.15	(0.03)	-0.30	(0.03)	0.09	(0.04)	-0.22	(0.10)	**-0.44**	(0.03)	-0.05	(0.04)	**-0.37**	(0.10)	**0.39**	(0.04)
Turkey	-1.11	(0.05)	-1.96	(0.15)	-1.35	(0.06)	c	c	**-0.85**	(0.14)	**-0.24**	(0.06)	c	c	**0.61**	(0.15)
United States	0.25	(0.04)	-0.23	(0.10)	-0.41	(0.07)	-0.18	(0.12)	**-0.48**	(0.10)	**-0.66**	(0.07)	**-0.42**	(0.12)	-0.18	(0.11)
Partners																
Argentina	-0.94	(0.04)	-1.62	(0.19)	-0.93	(0.13)	-1.06	(0.17)	**-0.69**	(0.19)	0.01	(0.14)	-0.12	(0.17)	**0.70**	(0.24)
Brazil	-0.98	(0.04)	-1.70	(0.04)	-1.27	(0.03)	-1.52	(0.08)	**-0.72**	(0.04)	**-0.29**	(0.04)	**-0.54**	(0.08)	**0.43**	(0.04)
Colombia	-1.04	(0.06)	-1.61	(0.07)	-1.09	(0.05)	-1.37	(0.13)	**-0.58**	(0.07)	-0.05	(0.06)	**-0.33**	(0.14)	**0.52**	(0.09)
Dubai (UAE)	0.48	(0.01)	0.00	(0.05)	0.13	(0.05)	-0.18	(0.11)	**-0.48**	(0.05)	**-0.35**	(0.05)	**-0.66**	(0.11)	0.13	(0.07)
Hong Kong-China	-0.72	(0.04)	-1.34	(0.05)	-1.02	(0.07)	-0.87	(0.31)	**-0.61**	(0.06)	**-0.30**	(0.07)	-0.14	(0.30)	**0.31**	(0.07)
Indonesia	-1.49	(0.06)	-1.90	(0.05)	-1.34	(0.18)	-1.76	(0.07)	**-0.41**	(0.07)	0.15	(0.18)	**-0.27**	(0.08)	**0.56**	(0.17)
Latvia	-0.08	(0.03)	-0.63	(0.08)	-0.46	(0.07)	-0.48	(0.21)	**-0.55**	(0.09)	**-0.39**	(0.07)	-0.40	(0.22)	0.16	(0.11)
Macao-China	-0.62	(0.02)	-0.99	(0.03)	-0.70	(0.02)	-0.81	(0.04)	**-0.37**	(0.04)	**-0.08**	(0.03)	**-0.19**	(0.04)	**0.29**	(0.04)
Panama	-0.65	(0.10)	-1.55	(0.10)	-1.00	(0.07)	-0.80	(0.20)	**-0.90**	(0.15)	**-0.35**	(0.14)	-0.16	(0.15)	**0.54**	(0.10)
Peru	-1.10	(0.06)	-2.07	(0.05)	-1.49	(0.06)	-1.80	(0.08)	**-0.97**	(0.07)	**-0.39**	(0.08)	**-0.70**	(0.11)	**0.58**	(0.07)
Qatar	0.56	(0.01)	0.15	(0.04)	0.28	(0.04)	0.25	(0.06)	**-0.41**	(0.04)	**-0.27**	(0.04)	**-0.30**	(0.06)	**0.14**	(0.06)
Trinidad and Tobago	-0.49	(0.02)	-0.80	(0.03)	-0.83	(0.11)	-0.69	(0.11)	**-0.31**	(0.04)	**-0.33**	(0.12)	-0.20	(0.10)	-0.03	(0.12)
Tunisia	-0.92	(0.08)	-1.87	(0.06)	-1.34	(0.06)	-1.73	(0.07)	**-0.95**	(0.09)	**-0.41**	(0.09)	**-0.81**	(0.10)	**0.53**	(0.07)
Uruguay	-0.33	(0.03)	-1.45	(0.04)	-1.09	(0.05)	-1.58	(0.06)	**-1.12**	(0.05)	**-0.76**	(0.05)	**-1.25**	(0.06)	**0.36**	(0.05)

Note: Values that are statistically significant are indicated in bold.

Table A.6. **Reading performance and grade repetition, Model 1, Model 2**

	Model 1									Model 2									
	Intercept		drepN		drep1		drep12		R^2	Intercept		drepN		drep1		drep12		ESCS		R^2
	Inter-cept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.		Inter-cept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																				
Austria	**448.2**	(6.5)	**35.4**	(6.7)	**-74.0**	(8.7)	**-76.0**	(14.6)	0.06	**447.9**	(5.9)	**30.2**	(6.0)	**-57.2**	(8.5)	**-64.5**	(13.3)	**42.9**	(2.3)	0.19
Belgium	**474.1**	(3.7)	**75.2**	(3.5)	**-53.8**	(4.1)	**-79.3**	(6.8)	0.29	**471.6**	(3.5)	**65.6**	(3.2)	**-41.8**	(4.1)	**-66.3**	(7.0)	**30.4**	(1.4)	0.37
Chile	**401.8**	(3.9)	**67.8**	(3.9)	**-32.9**	(5.1)	**-29.0**	(14.7)	0.18	**423.5**	(3.5)	**58.3**	(3.6)	**-25.2**	(4.6)	-25.5	(14.3)	**26.6**	(1.3)	0.31
France	**445.6**	(6.9)	**97.8**	(7.8)	**-57.1**	(7.7)	**-71.1**	(15.2)	0.36	**454.4**	(6.8)	**87.6**	(7.6)	**-45.6**	(7.7)	**-62.1**	(15.0)	**30.8**	(2.7)	0.41
Germany	**477.5**	(4.4)	**44.0**	(4.8)	**-67.5**	(5.3)	**-78.9**	(11.5)	0.13	**473.0**	(4.0)	**38.3**	(4.3)	**-52.0**	(5.3)	**-56.3**	(12.0)	**35.6**	(1.7)	0.25
Hungary	**409.5**	(7.6)	**98.6**	(7.9)	**-45.6**	(12.1)	**-32.9**	(13.2)	0.18	**445.5**	(5.4)	**66.3**	(5.3)	**-36.2**	(10.1)	**-36.6**	(10.9)	**39.1**	(2.0)	0.35
Ireland	**419.4**	(14.7)	**86.6**	(14.9)	23.3	(16.2)	c	c	0.06	**431.0**	(14.3)	**72.4**	(14.3)	15.0	(15.8)	c	c	**37.6**	(2.0)	0.17
Italy	**417.6**	(3.1)	**83.5**	(3.2)	**-54.4**	(10.8)	**-137.6**	(20.7)	0.12	**428.7**	(2.7)	**74.1**	(2.8)	**-38.5**	(11.3)	**-112.5**	(18.6)	**27.9**	(1.1)	0.21
Luxembourg	**455.2**	(3.3)	**55.3**	(3.6)	**-73.1**	(5.1)	**-91.7**	(8.9)	0.26	**449.1**	(3.0)	**50.4**	(3.3)	**-54.3**	(5.0)	**-78.2**	(8.3)	**28.5**	(1.3)	0.35
Mexico	**397.3**	(6.1)	**47.2**	(5.7)	**-46.1**	(5.4)	**-88.8**	(11.0)	0.18	**416.9**	(5.9)	**49.1**	(5.4)	**-28.3**	(5.2)	**-78.1**	(10.8)	**19.8**	(0.9)	0.27
Netherlands	**521.4**	(8.1)	7.2	(7.3)	**-63.9**	(7.9)	c	c	0.11	**509.0**	(8.3)	8.5	(7.2)	**-54.5**	(7.4)	c	c	**31.2**	(1.8)	0.20
Portugal	**437.8**	(3.8)	**90.7**	(4.4)	**-29.8**	(4.5)	**-44.4**	(5.2)	0.38	**449.3**	(3.5)	**80.1**	(3.8)	**-25.6**	(4.2)	**-39.5**	(5.0)	**17.3**	(1.2)	0.43
Spain	**432.9**	(2.6)	**83.4**	(2.4)	**-21.0**	(4.4)	**-66.0**	(4.9)	0.32	**444.9**	(2.4)	**72.3**	(2.0)	**-21.3**	(4.3)	**-60.5**	(4.9)	**16.3**	(1.5)	0.35
Switzerland	**486.2**	(4.2)	**32.6**	(4.6)	**-65.5**	(5.1)	**-65.9**	(13.0)	0.14	**482.8**	(4.1)	**30.9**	(4.5)	**-52.0**	(5.0)	**-55.1**	(13.0)	**33.8**	(2.0)	0.24
Turkey	**411.2**	(4.9)	**63.6**	(5.5)	**-45.5**	(8.6)	c	c	0.11	**447.7**	(5.2)	**57.1**	(5.1)	**-29.1**	(9.0)	c	c	**27.0**	(1.4)	0.26
United States	**422.9**	(7.7)	**92.8**	(7.8)	-6.7	(10.1)	-6.4	(14.8)	0.13	**437.9**	(7.8)	**68.8**	(7.7)	-13.2	(8.2)	-14.6	(14.6)	**36.2**	(2.6)	0.25
Partners																				
Argentina	**353.4**	(5.3)	**82.9**	(6.9)	**-55.6**	(7.4)	**-50.4**	(9.1)	0.24	**380.1**	(5.1)	**69.4**	(5.5)	**-41.7**	(7.3)	**-43.1**	(8.7)	**31.4**	(2.0)	0.35
Brazil	**384.9**	(2.5)	**61.7**	(3.2)	**-41.0**	(3.3)	**-54.6**	(4.2)	0.21	**412.3**	(2.7)	**55.4**	(3.1)	**-31.7**	(3.3)	**-49.1**	(4.3)	**21.6**	(1.4)	0.28
Colombia	**398.3**	(4.5)	**34.5**	(4.2)	**-39.1**	(5.4)	**-28.0**	(8.3)	0.11	**425.3**	(4.4)	**33.3**	(4.0)	**-26.0**	(4.7)	**-21.1**	(8.5)	**24.8**	(1.7)	0.24
Dubai (UAE)	**378.6**	(6.2)	**94.8**	(6.4)	7.2	(8.8)	**-72.5**	(13.6)	0.10	**373.2**	(5.6)	**78.7**	(5.8)	12.6	(7.7)	**-59.1**	(13.1)	**44.6**	(1.5)	0.20
Hong Kong-China	**477.1**	(6.5)	**66.8**	(6.6)	7.5	(8.3)	c	c	0.09	**491.6**	(6.7)	**62.6**	(6.6)	11.9	(8.1)	c	c	**14.1**	(1.9)	0.12
Indonesia	**354.4**	(8.5)	**56.0**	(8.8)	5.0	(8.6)	**29.3**	(9.3)	0.08	**374.8**	(9.4)	**58.3**	(8.4)	13.6	(8.3)	**35.8**	(9.1)	**15.2**	(2.3)	0.14
Latvia	**417.1**	(7.0)	**77.0**	(6.9)	**-20.5**	(8.9)	c	c	0.12	**428.6**	(7.1)	**67.3**	(6.9)	-16.4	(9.7)	c	c	**24.8**	(2.4)	0.19
Macao-China	**458.4**	(1.8)	**58.3**	(2.2)	**-11.8**	(3.3)	**-39.4**	(3.8)	0.21	**463.5**	(1.9)	**57.7**	(2.2)	**-9.7**	(3.4)	**-38.7**	(3.8)	**7.1**	(1.1)	0.22
Panama	**330.1**	(6.1)	**75.9**	(7.9)	**-33.3**	(9.7)	3.4	(13.0)	0.18	**357.2**	(6.1)	**66.4**	(5.7)	-18.3	(10.2)	-1.6	(9.2)	**27.1**	(3.1)	0.31
Peru	**341.4**	(5.1)	**53.7**	(6.1)	**-41.7**	(5.1)	**-58.6**	(7.8)	0.17	**393.1**	(5.3)	**40.0**	(4.7)	**-21.6**	(5.0)	**-47.7**	(8.0)	**34.7**	(2.0)	0.35
Qatar	**302.4**	(3.9)	**83.6**	(4.1)	**24.0**	(6.3)	**-38.7**	(6.7)	0.07	**296.2**	(3.9)	**77.7**	(4.1)	**27.0**	(6.4)	**-38.0**	(6.6)	**21.8**	(1.2)	0.10
Trinidad and Tobago	**295.9**	(12.8)	**152.3**	(13.0)	**61.2**	(13.3)	-15.4	(18.0)	0.17	**322.1**	(13.3)	**141.8**	(13.4)	**60.2**	(13.9)	-20.3	(18.4)	**31.7**	(1.8)	0.24
Tunisia	**371.2**	(3.5)	**74.7**	(4.8)	**-35.8**	(5.1)	**-52.3**	(4.5)	0.34	**383.6**	(3.9)	**70.9**	(4.4)	**-30.9**	(5.2)	**-48.7**	(4.6)	**9.3**	(1.3)	0.36
Uruguay	**384.4**	(4.0)	**84.0**	(4.7)	**-43.5**	(4.5)	**-57.2**	(5.6)	0.31	**408.7**	(4.3)	**67.1**	(4.4)	**-35.6**	(4.2)	**-46.4**	(5.5)	**22.2**	(1.4)	0.38

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.6. **Reading performance and grade repetition, Model 3** *(cont.)*

	Model 3																
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																	
Austria	**486.7**	(6.0)	**24.0**	(5.5)	**-58.7**	(8.4)	**-61.9**	(13.0)	**39.9**	(2.3)	**-40.0**	(4.5)	**-27.4**	(5.5)	**-19.6**	(3.0)	0.25
Belgium	**487.8**	(3.8)	**62.3**	(3.2)	**-44.0**	(4.0)	**-66.1**	(6.9)	**29.2**	(1.2)	**-20.3**	(2.7)	**-17.7**	(3.6)	3.1	(2.2)	0.39
Chile	**437.5**	(3.8)	**56.1**	(3.5)	**-25.3**	(4.5)	-26.2	(13.8)	**26.8**	(1.3)	**-16.5**	(3.1)	**-55.0**	(12.3)	**-6.7**	(2.0)	0.33
France	**475.4**	(7.4)	**83.4**	(7.7)	**-45.8**	(7.4)	**-60.1**	(14.4)	**29.0**	(2.4)	**-30.7**	(3.3)	**-27.5**	(5.8)	0.8	(2.7)	0.44
Germany	**505.8**	(4.1)	**34.7**	(4.2)	**-52.5**	(5.3)	**-55.7**	(12.2)	**33.0**	(1.7)	**-36.1**	(3.3)	**-24.4**	(3.8)	**-14.7**	(2.8)	0.30
Hungary	**474.5**	(5.7)	**61.9**	(5.4)	**-38.1**	(9.8)	**-34.3**	(10.8)	**40.0**	(2.0)	**-38.2**	(3.0)	-1.5	(10.0)	**-11.1**	(2.5)	0.39
Ireland	**465.2**	(14.3)	**67.0**	(14.4)	9.6	(15.6)	c	c	**38.3**	(1.9)	**-40.2**	(3.7)	**-18.5**	(6.6)	**-13.4**	(2.7)	0.23
Italy	**461.9**	(2.8)	**66.8**	(2.8)	**-33.4**	(11.0)	**-94.8**	(20.1)	**26.4**	(1.1)	**-39.0**	(2.3)	**-26.7**	(2.4)	**-6.9**	(1.5)	0.26
Luxembourg	**453.7**	(6.2)	**51.4**	(3.3)	**-51.6**	(5.0)	**-71.0**	(8.2)	**29.8**	(1.3)	**-35.1**	(2.2)	**22.0**	(5.7)	**-20.0**	(2.7)	0.39
Mexico	**439.0**	(5.9)	**43.3**	(5.4)	**-28.7**	(5.2)	**-70.9**	(10.6)	**19.2**	(0.8)	**-20.5**	(1.3)	**-60.8**	(5.5)	**-10.4**	(1.5)	0.31
Netherlands	**525.3**	(8.5)	6.1	(7.4)	**-54.9**	(7.6)	c	c	**30.5**	(2.1)	**-22.4**	(2.6)	-11.3	(7.6)	-3.3	(3.0)	0.22
Portugal	**468.5**	(3.3)	**75.4**	(3.7)	**-25.5**	(4.0)	**-37.5**	(5.0)	**18.3**	(1.2)	**-28.3**	(2.0)	**-16.1**	(3.9)	-2.7	(1.9)	0.46
Spain	**463.2**	(2.5)	**69.0**	(2.0)	**-21.3**	(4.4)	**-60.2**	(4.7)	**16.5**	(1.4)	**-20.9**	(1.6)	**-14.4**	(2.9)	**-4.1**	(1.7)	0.37
Switzerland	**516.2**	(3.9)	**28.1**	(4.2)	**-52.0**	(4.5)	**-52.4**	(12.5)	**31.3**	(2.0)	**-36.7**	(2.2)	**-22.1**	(2.5)	**-12.4**	(2.4)	0.29
Turkey	**470.5**	(5.7)	**51.1**	(5.1)	**-28.1**	(9.2)	c	c	**26.0**	(1.5)	**-32.4**	(3.1)	-14.9	(8.7)	-3.1	(2.3)	0.31
United States	**452.9**	(7.8)	**66.6**	(7.5)	-14.3	(8.0)	-14.4	(14.5)	**37.5**	(2.7)	**-22.5**	(2.7)	6.4	(3.5)	**-6.3**	(2.5)	0.26
Partners																	
Argentina	**398.7**	(5.4)	**65.8**	(5.4)	**-40.1**	(7.2)	**-40.8**	(8.9)	**32.1**	(1.9)	**-28.7**	(3.1)	**-20.8**	(9.7)	-3.7	(2.9)	0.37
Brazil	**432.5**	(2.9)	**51.8**	(3.1)	**-31.5**	(3.4)	**-47.3**	(4.4)	**22.9**	(1.3)	**-25.1**	(1.8)	**-65.5**	(8.9)	**-8.1**	(2.0)	0.31
Colombia	**435.3**	(4.6)	**32.0**	(4.1)	**-26.6**	(4.8)	**-20.9**	(8.5)	**25.1**	(1.7)	**-10.1**	(2.9)	**-66.1**	(14.4)	**-6.4**	(2.9)	0.25
Dubai (UAE)	**362.9**	(5.7)	**63.3**	(5.3)	12.6	(7.0)	**-44.2**	(11.5)	**37.0**	(1.5)	**-48.3**	(2.5)	**72.3**	(2.9)	**-11.0**	(2.6)	0.32
Hong Kong-China	**511.4**	(6.8)	**60.7**	(6.5)	10.9	(8.0)	c	c	**14.7**	(2.0)	**-30.3**	(3.6)	2.7	(4.3)	**-5.9**	(2.4)	0.15
Indonesia	**399.1**	(9.8)	**50.2**	(8.4)	11.5	(8.4)	**31.2**	(8.6)	**17.0**	(2.0)	**-34.2**	(3.0)	6.6	(4.1)	-3.2	(2.4)	0.21
Latvia	**462.3**	(7.0)	**56.0**	(6.6)	-17.7	(9.2)	c	c	**26.0**	(2.4)	**-42.6**	(2.6)	-7.6	(5.5)	-2.9	(2.8)	0.26
Macao-China	**484.6**	(3.0)	**54.9**	(2.2)	**-10.9**	(3.4)	**-37.7**	(3.8)	**6.5**	(1.1)	**-25.4**	(1.7)	**-5.4**	(2.0)	**-5.6**	(2.2)	0.25
Panama	**375.6**	(8.7)	**63.2**	(5.8)	-17.3	(9.3)	-1.3	(9.3)	**26.8**	(2.8)	**-20.6**	(6.7)	**-39.7**	(14.3)	-5.7	(5.2)	0.34
Peru	**407.9**	(5.8)	**35.4**	(4.6)	**-22.6**	(5.1)	**-46.1**	(7.7)	**33.0**	(1.9)	**-16.4**	(2.9)	**-59.1**	(7.3)	-4.2	(2.6)	0.37
Qatar	**297.9**	(4.1)	**64.6**	(3.8)	**22.3**	(6.0)	**-31.6**	(6.9)	**19.5**	(1.1)	**-39.1**	(1.7)	**51.9**	(2.0)	**-6.0**	(1.7)	0.18
Trinidad and Tobago	**360.7**	(12.5)	**129.1**	(12.3)	**52.8**	(12.7)	-27.1	(17.2)	**32.9**	(1.8)	**-50.5**	(2.9)	**-25.8**	(9.5)	-1.8	(2.9)	0.29
Tunisia	**399.1**	(4.5)	**67.3**	(4.4)	**-32.7**	(5.4)	**-48.3**	(4.6)	**9.9**	(1.3)	**-20.2**	(2.1)	c	c	**-6.0**	(2.1)	0.38
Uruguay	**431.0**	(4.6)	**61.5**	(4.3)	**-35.1**	(4.1)	**-46.0**	(5.2)	**23.6**	(1.4)	**-33.8**	(2.8)	**-20.8**	(8.7)	-3.0	(2.2)	0.41

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.7. **Reading performance, grade repetition and students' background characteristics, Model 4**

	Model 4																		
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*ESCS		drep1*ESCS		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**487.1**	(6.0)	**23.6**	(5.5)	**-68.0**	(8.4)	**-64.1**	(13.3)	**31.2**	(4.7)	**-40.0**	(4.5)	**-28.2**	(5.5)	**-19.9**	(3.0)	**10.7**	(5.0)	-15.4	(9.3)	0.25
Belgium	**488.1**	(3.9)	**60.6**	(3.4)	**-46.9**	(4.3)	**-68.6**	(6.7)	**23.4**	(3.1)	**-20.1**	(2.7)	**-17.6**	(3.6)	3.2	(2.2)	**9.1**	(3.4)	-1.9	(4.2)	0.39
Chile	**432.9**	(3.9)	**61.5**	(3.7)	**-27.7**	(6.6)	-27.0	(13.9)	**21.1**	(2.8)	**-16.4**	(3.1)	**-54.9**	(12.1)	**-6.8**	(2.0)	**7.2**	(2.9)	-0.7	(4.7)	0.33
France	**473.0**	(7.8)	**85.4**	(8.1)	**-49.9**	(8.5)	**-62.4**	(14.5)	**21.3**	(5.7)	**-30.6**	(3.3)	**-27.2**	(5.7)	1.0	(2.7)	**12.5**	(6.3)	-1.8	(7.4)	0.45
Germany	**506.5**	(4.2)	**33.5**	(4.2)	**-56.7**	(5.5)	**-58.7**	(12.0)	**28.2**	(4.0)	**-36.1**	(3.3)	**-24.8**	(3.8)	**-14.7**	(2.8)	6.8	(4.6)	-6.7	(6.9)	0.31
Hungary	**474.6**	(7.6)	**62.0**	(7.0)	**-52.5**	(16.2)	**-34.3**	(10.8)	**40.0**	(5.6)	**-38.3**	(3.0)	-1.6	(9.8)	**-11.2**	(2.4)	0.6	(5.7)	-12.4	(10.2)	0.39
Ireland	**463.1**	(14.7)	**69.1**	(14.8)	11.3	(15.8)	c	c	**31.5**	(10.8)	**-40.2**	(3.7)	**-18.4**	(6.6)	**-13.5**	(2.7)	7.3	(11.1)	2.9	(11.7)	0.23
Italy	**460.1**	(2.7)	**68.6**	(2.6)	**-45.1**	(14.0)	**-98.7**	(19.5)	**22.1**	(3.1)	**-38.9**	(2.3)	**-26.8**	(2.5)	**-7.0**	(1.5)	5.2	(3.2)	-9.6	(12.5)	0.26
Luxembourg	**456.6**	(6.2)	**47.5**	(3.5)	**-56.6**	(5.5)	**-75.9**	(8.6)	**19.5**	(2.5)	**-34.8**	(2.2)	**21.1**	(5.6)	**-19.9**	(2.7)	**15.0**	(2.9)	4.1	(4.4)	0.40
Mexico	**436.2**	(6.7)	**47.8**	(6.3)	**-46.5**	(7.4)	**-72.4**	(10.7)	**16.5**	(2.7)	**-20.5**	(1.3)	**-60.9**	(5.5)	**-10.3**	(1.5)	4.4	(2.9)	**-8.1**	(3.6)	0.31
Netherlands	**530.7**	(9.0)	-0.3	(8.2)	**-60.0**	(8.3)	c	c	**16.6**	(6.5)	**-22.3**	(2.6)	-11.9	(7.8)	-3.0	(3.0)	**16.1**	(6.6)	9.1	(7.4)	0.22
Portugal	**465.5**	(3.6)	**78.5**	(4.0)	**-25.0**	(4.3)	**-38.8**	(5.2)	**13.8**	(2.5)	**-28.3**	(2.0)	**-15.8**	(3.9)	-2.6	(1.9)	**5.8**	(2.8)	1.8	(3.7)	0.46
Spain	**461.6**	(2.7)	**70.6**	(2.1)	**-23.1**	(5.4)	**-60.9**	(4.8)	**14.3**	(2.0)	**-20.9**	(1.6)	**-14.3**	(2.9)	**-4.1**	(1.7)	3.2	(1.8)	-2.6	(4.2)	0.38
Switzerland	**517.1**	(3.9)	**27.0**	(4.2)	**-54.4**	(4.7)	**-55.1**	(12.6)	**22.7**	(3.4)	**-36.6**	(2.2)	**-22.4**	(2.5)	**-12.5**	(2.4)	**10.4**	(4.1)	3.2	(4.7)	0.29
Turkey	**451.0**	(8.7)	**72.4**	(8.3)	**-51.9**	(24.3)	c	c	**11.6**	(3.9)	**-32.3**	(3.1)	-14.2	(8.6)	-3.3	(2.3)	**16.0**	(3.9)	-7.7	(11.1)	0.31
United States	**454.0**	(8.8)	**65.4**	(8.6)	-16.9	(10.7)	-15.0	(14.7)	**40.0**	(6.6)	**-22.4**	(2.7)	6.3	(3.5)	**-6.5**	(2.5)	-1.9	(7.1)	-9.2	(12.3)	0.26
Partners																					
Argentina	**391.4**	(6.2)	**74.9**	(6.4)	**-60.0**	(11.3)	**-42.8**	(9.0)	**23.6**	(3.1)	**-28.7**	(3.1)	-18.2	(9.8)	-3.8	(2.9)	**12.8**	(4.0)	**-12.5**	(6.1)	0.38
Brazil	**422.2**	(3.0)	**67.8**	(3.8)	**-47.8**	(5.3)	**-49.4**	(4.5)	**14.7**	(1.8)	**-24.7**	(1.8)	**-64.2**	(9.2)	**-8.5**	(2.1)	**13.9**	(2.5)	**-7.6**	(2.8)	0.32
Colombia	**431.0**	(5.7)	**38.7**	(5.7)	**-31.2**	(7.5)	**-22.1**	(8.2)	**21.0**	(2.8)	**-10.3**	(3.0)	**-67.5**	(14.2)	**-6.4**	(2.9)	**6.3**	(3.2)	-1.5	(3.9)	0.25
Dubai (UAE)	**362.9**	(5.5)	**63.3**	(5.1)	12.6	(6.8)	**-44.2**	(11.5)	**36.9**	(4.7)	**-48.3**	(2.5)	**72.3**	(2.9)	**-11.0**	(2.6)	0.0	(5.1)	0.2	(7.4)	0.32
Hong Kong-China	**512.3**	(8.4)	**60.0**	(8.3)	7.1	(10.4)	c	c	**15.6**	(5.8)	**-30.3**	(3.6)	2.7	(4.3)	**-5.9**	(2.4)	-0.7	(5.9)	-3.0	(6.0)	0.15
Indonesia	**387.0**	(13.4)	**63.9**	(12.1)	7.2	(14.0)	**27.6**	(8.7)	8.4	(5.3)	**-33.7**	(3.0)	6.7	(4.0)	-3.0	(2.4)	**10.0**	(4.9)	0.3	(5.8)	0.21
Latvia	**458.1**	(7.0)	**60.3**	(6.7)	-20.8	(12.4)	c	c	**16.7**	(7.1)	**-42.7**	(2.6)	-7.2	(5.5)	-3.0	(2.7)	10.3	(6.8)	-2.5	(12.2)	0.26
Macao-China	**482.9**	(3.4)	**57.1**	(2.9)	-7.4	(5.2)	**-38.0**	(3.8)	4.0	(2.1)	**-25.4**	(1.7)	**-5.5**	(2.0)	**-5.6**	(2.2)	3.3	(2.3)	4.3	(4.4)	0.25
Panama	**371.2**	(8.9)	**69.7**	(7.5)	**-51.6**	(12.2)	-0.7	(9.7)	**23.1**	(4.4)	**-19.4**	(6.6)	**-36.5**	(14.1)	-6.1	(5.1)	7.7	(4.5)	**-20.7**	(5.3)	0.35
Peru	**386.8**	(7.0)	**60.0**	(7.3)	**-23.4**	(8.4)	**-50.4**	(7.6)	**19.0**	(3.4)	**-15.9**	(2.9)	**-60.4**	(7.3)	-4.0	(2.6)	**17.3**	(4.0)	3.5	(4.3)	0.38
Qatar	**301.0**	(4.1)	**60.7**	(3.8)	**19.2**	(5.8)	**-32.0**	(6.8)	**8.3**	(3.7)	**-39.0**	(1.7)	**51.8**	(2.0)	**-5.9**	(1.7)	**12.6**	(3.7)	**10.5**	(5.2)	0.18
Trinidad and Tobago	**345.3**	(17.9)	**145.7**	(17.4)	**62.1**	(18.3)	-24.2	(17.5)	14.6	(9.8)	**-50.4**	(2.9)	**-25.4**	(9.5)	-1.5	(2.9)	**21.4**	(10.0)	11.1	(10.4)	0.29
Tunisia	**385.9**	(5.3)	**84.3**	(6.0)	**-28.7**	(7.6)	**-52.1**	(4.5)	0.2	(2.1)	**-19.7**	(2.1)	c	c	**-6.0**	(2.2)	**14.0**	(2.7)	4.9	(3.6)	0.39
Uruguay	**418.0**	(5.2)	**75.6**	(5.1)	**-32.2**	(6.1)	**-51.7**	(5.2)	**11.9**	(2.4)	**-33.3**	(2.7)	**-21.8**	(8.7)	-3.3	(2.2)	**15.5**	(2.6)	4.9	(3.2)	0.41

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.7. **Reading performance, grade repetition and students' background characteristics, Model 5** *(cont.)*

	Model 5																			
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*MALE		drep1*MALE		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**488.5**	(7.7)	**22.5**	(8.3)	**-74.8**	(12.9)	**-61.8**	(13.0)	**40.0**	(2.3)	**-43.2**	(9.0)	**-27.2**	(5.4)	**-19.6**	(3.0)	2.5	(10.0)	29.6	(15.9)	0.25
Belgium	**483.9**	(4.3)	**67.1**	(4.3)	**-39.0**	(5.4)	**-66.2**	(6.9)	**29.2**	(1.2)	**-13.6**	(4.6)	**-17.7**	(3.6)	3.1	(2.2)	-8.4	(5.1)	-9.1	(6.8)	0.39
Chile	**436.7**	(5.2)	**56.8**	(6.0)	**-20.3**	(7.9)	-26.2	(13.9)	**26.9**	(1.3)	**-15.0**	(6.4)	**-54.9**	(12.3)	**-6.8**	(2.0)	-1.1	(7.4)	-8.4	(11.7)	0.33
France	**474.2**	(7.3)	**84.0**	(8.1)	**-39.8**	(8.9)	**-60.3**	(14.5)	**29.0**	(2.4)	**-28.6**	(6.6)	**-27.6**	(5.8)	0.8	(2.7)	-0.9	(7.8)	-11.2	(11.1)	0.45
Germany	**505.1**	(5.6)	**35.5**	(5.9)	**-51.9**	(6.8)	**-55.6**	(12.2)	**33.0**	(1.7)	**-34.9**	(6.8)	**-24.4**	(3.8)	**-14.7**	(2.8)	-1.4	(7.5)	-1.0	(11.4)	0.30
Hungary	**467.2**	(8.1)	**69.5**	(8.4)	-26.6	(14.0)	**-34.7**	(10.6)	**39.9**	(2.0)	**-26.8**	(8.3)	-1.1	(9.8)	**-11.2**	(2.5)	-11.8	(9.0)	-19.7	(17.9)	0.39
Ireland	**459.9**	(14.5)	**72.1**	(14.9)	17.5	(16.5)	c	c	**38.3**	(1.9)	-29.3	(22.5)	**-18.5**	(6.7)	**-13.5**	(2.7)	-10.5	(22.7)	-15.9	(24.3)	0.23
Italy	**461.9**	(4.2)	**66.7**	(4.4)	-15.7	(14.8)	**-94.7**	(20.2)	**26.4**	(1.1)	**-39.0**	(5.5)	**-26.8**	(2.4)	**-6.9**	(1.5)	0.2	(5.2)	-28.5	(20.4)	0.26
Luxembourg	**448.3**	(6.5)	**57.4**	(4.4)	**-40.4**	(6.6)	**-72.4**	(8.1)	**29.9**	(1.3)	**-24.1**	(5.8)	**21.6**	(5.7)	**-19.9**	(2.7)	-11.5	(6.8)	**-21.3**	(9.7)	0.39
Mexico	**426.8**	(6.4)	**55.4**	(6.1)	**-13.0**	(6.6)	**-71.0**	(10.6)	**19.2**	(0.8)	-2.9	(4.9)	**-60.6**	(5.5)	**-10.4**	(1.5)	**-17.6**	(5.0)	**-23.7**	(6.7)	0.31
Netherlands	**517.1**	(13.3)	14.2	(11.4)	**-45.0**	(12.7)	c	c	**30.6**	(2.1)	-8.7	(12.5)	-11.3	(7.6)	-3.3	(3.0)	-13.8	(12.6)	-16.9	(15.6)	0.22
Portugal	**469.0**	(3.6)	**73.9**	(3.9)	**-20.5**	(5.1)	**-37.5**	(5.1)	**18.3**	(1.2)	**-29.0**	(4.4)	**-16.2**	(3.9)	-2.7	(1.9)	3.3	(4.7)	-9.1	(6.6)	0.46
Spain	**463.4**	(3.1)	**69.3**	(3.0)	**-31.8**	(6.4)	**-60.2**	(4.7)	**16.5**	(1.4)	**-21.2**	(2.8)	**-14.4**	(2.8)	**-4.2**	(1.7)	-0.9	(3.4)	**20.0**	(9.2)	0.38
Switzerland	**516.0**	(4.8)	**28.7**	(5.3)	**-54.6**	(6.9)	**-52.4**	(12.5)	**31.3**	(2.0)	**-36.5**	(7.4)	**-22.1**	(2.5)	**-12.4**	(2.4)	-1.1	(8.4)	4.9	(9.9)	0.29
Turkey	**473.0**	(8.8)	**48.3**	(8.6)	**-26.2**	(13.3)	c	c	**26.0**	(1.5)	**-36.1**	(7.4)	-14.9	(8.7)	-3.2	(2.3)	4.2	(8.1)	-2.7	(15.7)	0.31
United States	**451.4**	(10.9)	**67.7**	(10.7)	-8.5	(12.2)	-14.5	(14.5)	**37.5**	(2.7)	-20.0	(14.5)	6.5	(3.5)	**-6.3**	(2.5)	-1.7	(14.9)	-10.3	(16.8)	0.26
Partners																					
Argentina	**395.3**	(6.0)	**71.4**	(6.0)	**-46.4**	(10.0)	**-41.3**	(8.9)	**32.1**	(2.0)	**-22.2**	(6.2)	**-20.7**	(9.7)	-3.8	(2.9)	-11.6	(7.1)	11.7	(14.4)	0.37
Brazil	**432.1**	(3.6)	**52.0**	(3.9)	**-29.4**	(4.4)	**-47.3**	(4.4)	**22.9**	(1.3)	**-24.4**	(3.9)	**-65.6**	(8.9)	**-8.1**	(2.0)	-0.3	(4.9)	-4.2	(6.0)	0.31
Colombia	**437.7**	(4.8)	**29.8**	(4.9)	**-32.4**	(5.8)	**-20.7**	(8.5)	**25.1**	(1.7)	**-14.5**	(7.2)	**-65.7**	(14.5)	**-6.3**	(2.9)	3.7	(7.8)	11.3	(9.9)	0.25
Dubai (UAE)	**374.7**	(8.1)	**50.2**	(8.0)	12.9	(10.4)	**-42.9**	(11.6)	**37.0**	(1.5)	**-65.7**	(10.2)	**72.3**	(2.9)	**-11.1**	(2.6)	20.1	(10.9)	-5.7	(14.3)	0.32
Hong Kong-China	**517.8**	(9.8)	**53.7**	(9.5)	7.1	(11.8)	c	c	**14.7**	(2.0)	**-41.1**	(11.9)	2.6	(4.3)	**-5.9**	(2.4)	12.0	(11.8)	6.3	(13.4)	0.15
Indonesia	**398.1**	(10.6)	**51.6**	(10.1)	7.6	(10.9)	**31.3**	(8.5)	**16.9**	(2.0)	**-32.9**	(9.6)	6.6	(4.1)	-3.2	(2.4)	-2.4	(9.7)	6.3	(11.2)	0.21
Latvia	**478.8**	(10.8)	**38.6**	(10.6)	-20.9	(12.9)	c	c	**25.9**	(2.4)	**-65.1**	(12.1)	-8.0	(5.5)	-3.0	(2.8)	**24.6**	(12.4)	2.8	(15.3)	0.26
Macao-China	**485.8**	(3.5)	**52.4**	(3.2)	-8.0	(4.6)	**-37.6**	(3.8)	**6.5**	(1.1)	**-27.4**	(3.4)	**-5.5**	(2.0)	**-5.6**	(2.2)	4.9	(4.1)	-5.7	(6.2)	0.25
Panama	**370.3**	(8.6)	**69.5**	(7.2)	-10.6	(12.1)	-1.3	(9.4)	**26.7**	(2.8)	-11.8	(8.0)	**-39.7**	(14.4)	-5.7	(5.2)	-10.9	(10.4)	-11.4	(14.6)	0.34
Peru	**407.8**	(7.1)	**36.4**	(6.8)	**-26.5**	(7.7)	**-46.1**	(7.7)	**33.0**	(1.9)	-16.2	(8.4)	**-58.8**	(7.3)	-4.2	(2.6)	-1.9	(9.7)	7.6	(11.0)	0.37
Qatar	**309.4**	(5.3)	**51.8**	(5.2)	15.2	(8.8)	**-29.5**	(7.0)	**19.6**	(1.1)	**-58.7**	(7.3)	**51.9**	(2.0)	**-6.0**	(1.7)	**22.1**	(7.6)	11.5	(11.6)	0.18
Trinidad and Tobago	**364.7**	(15.5)	**123.1**	(15.4)	**55.3**	(16.4)	-27.6	(16.9)	**32.9**	(1.8)	**-56.4**	(16.4)	**-25.9**	(9.5)	-1.9	(2.9)	10.2	(16.3)	-5.8	(18.5)	0.29
Tunisia	**402.2**	(4.8)	**61.9**	(4.6)	**-30.0**	(5.3)	**-48.3**	(4.6)	**9.9**	(1.3)	**-25.4**	(3.5)	c	c	**-5.9**	(2.1)	**10.7**	(4.2)	-7.3	(6.2)	0.38
Uruguay	**433.8**	(4.9)	**57.4**	(5.1)	**-35.6**	(5.5)	**-46.1**	(5.2)	**23.7**	(1.4)	**-38.8**	(4.7)	**-21.1**	(8.7)	-3.1	(2.2)	8.0	(5.9)	0.7	(7.9)	0.41

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.7. **Reading performance, grade repetition and students' background characteristics, Model 6** *(cont.)*

	Model 6																			
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*MIGLAN		drep1*MIGLAN		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**488.1**	(6.5)	**22.6**	(6.0)	**-68.0**	(9.0)	**-61.4**	(13.2)	**40.0**	(2.3)	**-39.8**	(4.5)	**-33.2**	(9.4)	**-19.6**	(3.0)	4.9	(11.2)	32.6	(16.6)	0.25
Belgium	**490.0**	(3.9)	**60.4**	(3.6)	**-51.6**	(4.3)	**-65.8**	(6.9)	**29.2**	(1.2)	**-20.4**	(2.7)	**-24.3**	(6.4)	3.2	(2.2)	5.5	(6.7)	**21.2**	(8.1)	0.39
Chile	**437.5**	(3.8)	**56.2**	(3.5)	**-24.7**	(4.6)	-26.1	(13.9)	**26.9**	(1.3)	**-16.5**	(3.1)	-43.7	(35.8)	**-6.7**	(2.0)	-8.4	(37.2)	-71.4	(40.2)	0.33
France	**474.9**	(7.9)	**84.8**	(8.2)	**-49.4**	(8.4)	**-60.2**	(14.4)	**28.8**	(2.4)	**-30.6**	(3.3)	**-25.5**	(9.3)	0.7	(2.7)	-9.6	(12.0)	14.8	(14.6)	0.45
Germany	**502.9**	(4.4)	**38.4**	(4.6)	**-55.7**	(6.5)	**-57.5**	(12.0)	**33.0**	(1.7)	**-36.0**	(3.3)	-11.9	(7.3)	**-14.6**	(2.8)	**-18.0**	(7.6)	5.5	(12.4)	0.31
Hungary	**475.0**	(5.7)	**61.4**	(5.3)	**-40.6**	(10.1)	**-32.7**	(10.9)	**40.0**	(2.0)	**-38.3**	(3.0)	-26.9	(34.6)	**-11.1**	(2.5)	25.1	(32.4)	53.2	(40.1)	0.39
Ireland	**460.1**	(15.6)	**72.4**	(15.9)	12.7	(16.5)	c	c	**38.2**	(1.9)	**-40.3**	(3.7)	-5.2	(24.4)	**-13.4**	(2.7)	-16.2	(24.9)	4.4	(27.3)	0.23
Italy	**459.9**	(3.2)	**69.1**	(3.3)	**-51.4**	(14.0)	**-97.6**	(19.8)	**26.4**	(1.1)	**-38.9**	(2.3)	**-18.1**	(5.4)	**-6.9**	(1.5)	-11.6	(6.0)	35.1	(19.3)	0.26
Luxembourg	**431.2**	(9.7)	**79.9**	(11.9)	-21.4	(17.3)	**-70.6**	(8.2)	**29.9**	(1.3)	**-35.3**	(2.2)	**46.4**	(10.5)	**-20.1**	(2.7)	**-30.6**	(12.9)	-32.5	(18.0)	0.39
Mexico	**440.7**	(5.9)	**41.7**	(5.4)	**-32.1**	(5.2)	**-66.6**	(10.6)	**19.1**	(0.8)	**-20.6**	(1.3)	**-95.1**	(23.0)	**-10.4**	(1.5)	29.5	(23.1)	**53.1**	(22.8)	0.31
Netherlands	**525.4**	(7.8)	6.2	(7.0)	**-56.7**	(7.4)	c	c	**30.6**	(2.1)	**-22.4**	(2.6)	-12.8	(18.7)	-3.3	(3.0)	-1.9	(20.2)	10.2	(20.1)	0.22
Portugal	**467.4**	(3.5)	**77.1**	(3.8)	**-25.3**	(4.2)	**-37.3**	(5.0)	**18.2**	(1.2)	**-28.2**	(2.0)	-2.9	(7.1)	-2.7	(1.9)	**-22.9**	(8.7)	-2.1	(10.9)	0.46
Spain	**462.8**	(2.5)	**69.5**	(2.2)	**-21.4**	(5.3)	**-60.1**	(4.7)	**16.5**	(1.4)	**-20.9**	(1.6)	**-13.0**	(3.6)	**-4.1**	(1.7)	-2.3	(3.9)	0.2	(9.3)	0.37
Switzerland	**517.3**	(4.1)	**27.2**	(4.4)	**-55.9**	(5.2)	**-51.4**	(12.5)	**31.3**	(2.0)	**-36.7**	(2.2)	**-25.8**	(6.3)	**-12.4**	(2.4)	2.5	(7.7)	11.1	(8.8)	0.29
Turkey	**470.1**	(5.9)	**51.5**	(5.4)	**-27.2**	(8.5)	c	c	**26.0**	(1.5)	**-32.4**	(3.1)	-5.4	(14.1)	-3.1	(2.3)	-10.2	(16.7)	-12.7	(35.7)	0.31
United States	**455.7**	(8.3)	**64.1**	(8.1)	**-21.5**	(9.5)	-14.3	(14.5)	**37.5**	(2.8)	**-22.5**	(2.7)	-4.5	(13.3)	**-6.3**	(2.5)	9.4	(13.6)	26.1	(16.6)	0.26
Partners																					
Argentina	**398.1**	(5.3)	**66.4**	(5.2)	**-37.9**	(7.0)	**-40.7**	(8.8)	**32.1**	(2.0)	**-28.7**	(3.1)	-7.7	(19.9)	-3.6	(2.9)	-13.1	(19.7)	-30.4	(23.8)	0.37
Brazil	**432.3**	(2.9)	**52.2**	(3.1)	**-31.5**	(3.4)	**-47.5**	(4.4)	**22.8**	(1.3)	**-25.1**	(1.8)	**-51.6**	(15.0)	**-8.1**	(2.0)	-32.7	(25.9)	-2.9	(23.7)	0.31
Colombia	**434.9**	(4.6)	**32.5**	(4.1)	**-26.2**	(4.8)	**-20.8**	(8.5)	**25.2**	(1.7)	**-10.0**	(2.9)	-28.9	(15.5)	**-6.4**	(2.9)	**-61.5**	(26.5)	-25.3	(44.0)	0.25
Dubai (UAE)	**357.8**	(7.2)	**68.9**	(7.7)	18.0	(10.3)	**-42.8**	(11.6)	**37.0**	(1.5)	**-48.2**	(2.5)	**79.8**	(8.7)	**-11.0**	(2.6)	-8.2	(9.6)	-8.1	(13.3)	0.32
Hong Kong-China	**504.5**	(8.6)	**68.6**	(8.7)	5.2	(11.9)	c	c	**14.6**	(2.0)	**-30.3**	(3.6)	16.9	(10.6)	**-5.7**	(2.4)	-17.3	(10.6)	4.3	(13.4)	0.15
Indonesia	**389.9**	(11.6)	**60.4**	(10.6)	16.1	(11.1)	**26.7**	(9.1)	**16.9**	(2.0)	**-34.1**	(3.1)	**23.9**	(10.3)	-3.2	(2.4)	-19.4	(11.1)	-10.8	(10.9)	0.21
Latvia	**462.0**	(7.2)	**56.2**	(6.8)	-15.8	(9.6)	c	c	**26.0**	(2.4)	**-42.7**	(2.6)	-4.3	(19.4)	-2.9	(2.8)	-2.8	(20.5)	-14.0	(27.3)	0.26
Macao-China	**482.4**	(3.5)	**61.1**	(3.5)	**-18.9**	(7.1)	**-37.6**	(3.8)	**6.4**	(1.1)	**-25.3**	(1.7)	-2.4	(3.2)	**-5.6**	(2.2)	**-8.3**	(4.0)	10.1	(7.9)	0.25
Panama	**376.7**	(8.8)	**61.5**	(5.9)	-15.6	(8.4)	-1.3	(9.4)	**26.9**	(2.8)	**-20.6**	(6.7)	**-51.3**	(17.2)	-5.7	(5.2)	21.2	(19.5)	-5.6	(29.4)	0.34
Peru	**408.1**	(5.9)	**35.5**	(4.7)	**-25.3**	(4.9)	**-45.9**	(7.7)	**32.9**	(1.9)	**-16.3**	(2.9)	**-64.6**	(14.6)	-4.3	(2.6)	-5.0	(16.2)	26.0	(14.5)	0.38
Qatar	**307.4**	(4.6)	**52.8**	(4.9)	**18.4**	(7.0)	**-32.4**	(6.7)	**19.5**	(1.1)	**-39.0**	(1.8)	**31.6**	(7.1)	**-5.8**	(1.7)	**23.4**	(7.4)	9.3	(10.6)	0.18
Trinidad and Tobago	**365.9**	(13.4)	**123.9**	(13.0)	**46.9**	(13.6)	-30.7	(17.3)	**32.9**	(1.8)	**-50.4**	(2.9)	**-69.4**	(25.6)	-1.8	(2.9)	42.0	(27.1)	54.3	(28.1)	0.29
Tunisia	**399.1**	(4.5)	**67.2**	(4.4)	**-32.7**	(5.4)	**-48.3**	(4.6)	**9.9**	(1.3)	**-20.2**	(2.1)	c	c	**-6.0**	(2.1)	c	c	c	c	0.38
Uruguay	**431.2**	(4.6)	**61.3**	(4.4)	**-35.5**	(4.2)	**-46.0**	(5.2)	**23.6**	(1.4)	**-33.8**	(2.7)	-28.5	(18.2)	-3.0	(2.2)	9.2	(22.1)	12.3	(24.6)	0.41

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.7. **Reading performance, grade repetition and students' background characteristics, Model 7** *(cont.)*

	Model 7																			
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*rAGE		drep1*rAGE		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**487.7**	(6.3)	**23.5**	(6.0)	**-77.8**	(9.9)	**-61.7**	(13.0)	**40.0**	(2.3)	**-39.9**	(4.5)	**-27.3**	(5.5)	**-22.1**	(7.7)	1.2	(8.4)	**42.2**	(14.5)	0.25
Belgium	**485.4**	(5.3)	**65.6**	(5.1)	**-42.0**	(6.2)	**-66.6**	(6.8)	**29.2**	(1.2)	**-20.4**	(2.7)	**-17.7**	(3.5)	7.7	(5.3)	-6.2	(5.8)	-3.9	(7.2)	0.39
Chile	**438.4**	(4.4)	**55.0**	(4.5)	**-24.2**	(6.6)	-26.3	(13.9)	**26.8**	(1.3)	**-16.5**	(3.1)	**-54.9**	(12.3)	-8.3	(4.9)	2.3	(5.4)	-1.9	(9.6)	0.33
France	**478.0**	(7.1)	**80.6**	(7.4)	**-53.1**	(9.2)	**-59.0**	(14.6)	**29.0**	(2.4)	**-30.6**	(3.3)	**-27.7**	(5.7)	-4.4	(7.9)	5.6	(8.3)	12.3	(10.3)	0.45
Germany	**506.1**	(4.9)	**34.4**	(5.5)	**-53.5**	(7.4)	**-55.5**	(12.2)	**33.0**	(1.7)	**-36.1**	(3.3)	**-24.4**	(3.8)	**-15.4**	(6.1)	0.7	(6.6)	2.8	(11.8)	0.30
Hungary	**477.5**	(6.2)	**58.9**	(6.3)	**-51.5**	(13.6)	**-33.4**	(10.4)	**40.0**	(2.0)	**-38.1**	(3.0)	-1.7	(10.0)	-20.1	(10.7)	8.9	(11.0)	27.0	(17.7)	0.39
Ireland	**482.8**	(18.8)	**50.1**	(19.2)	-16.0	(20.4)	c	c	**38.3**	(1.8)	**-40.1**	(3.7)	**-18.4**	(6.7)	**-45.9**	(21.2)	30.9	(20.6)	**50.9**	(21.8)	0.23
Italy	**462.4**	(3.2)	**66.1**	(3.2)	-26.0	(17.4)	**-94.6**	(19.9)	**26.4**	(1.1)	**-39.0**	(2.3)	**-26.7**	(2.5)	-8.1	(4.5)	1.5	(4.6)	-14.2	(23.5)	0.26
Luxembourg	**453.8**	(6.3)	**51.1**	(4.0)	**-51.3**	(5.3)	**-70.9**	(8.3)	**29.8**	(1.3)	**-35.1**	(2.2)	**22.1**	(5.7)	**-20.5**	(6.6)	0.9	(7.1)	-0.7	(10.2)	0.39
Mexico	**437.7**	(6.2)	**45.4**	(5.8)	**-30.6**	(5.8)	**-70.9**	(10.6)	**19.2**	(0.8)	**-20.5**	(1.3)	**-60.8**	(5.5)	-7.9	(6.0)	-3.9	(6.3)	3.6	(7.5)	0.31
Netherlands	**521.9**	(7.9)	11.7	(7.9)	**-58.6**	(8.3)	c	c	**30.3**	(2.1)	**-22.4**	(2.6)	-11.3	(7.6)	7.6	(13.9)	-15.5	(14.1)	2.3	(14.3)	0.22
Portugal	**469.5**	(3.8)	**74.9**	(4.5)	**-30.0**	(4.5)	**-37.3**	(5.0)	**18.3**	(1.2)	**-28.3**	(2.0)	**-16.3**	(3.9)	-4.9	(4.5)	1.2	(5.3)	9.4	(5.4)	0.46
Spain	**462.4**	(2.8)	**70.5**	(2.6)	**-28.4**	(7.5)	**-60.4**	(4.7)	**16.5**	(1.4)	**-21.0**	(1.6)	**-14.4**	(2.9)	-2.6	(3.4)	-3.0	(3.4)	11.3	(10.5)	0.37
Switzerland	**517.3**	(5.1)	**26.2**	(5.5)	**-50.0**	(6.3)	**-52.6**	(12.5)	**31.2**	(2.0)	**-36.7**	(2.2)	**-22.1**	(2.5)	**-14.6**	(6.4)	3.6	(7.3)	-4.0	(8.7)	0.29
Turkey	**471.5**	(6.9)	**50.3**	(6.8)	**-36.8**	(11.7)	c	c	**26.0**	(1.5)	**-32.5**	(3.1)	-14.6	(8.7)	-5.4	(6.7)	1.8	(7.2)	19.5	(13.7)	0.31
United States	**451.3**	(9.5)	**68.4**	(9.3)	-13.7	(11.4)	-14.5	(14.6)	**37.5**	(2.7)	**-22.5**	(2.7)	6.4	(3.4)	-2.5	(11.3)	-4.3	(11.5)	-1.9	(14.7)	0.26
Partners																					
Argentina	**395.8**	(5.7)	**69.8**	(6.2)	**-37.0**	(11.1)	**-41.1**	(8.8)	**32.1**	(1.9)	**-28.7**	(3.1)	**-20.7**	(9.8)	2.2	(5.9)	-7.9	(6.9)	-6.2	(11.1)	0.37
Brazil	**430.3**	(3.2)	**56.0**	(3.8)	**-33.1**	(5.2)	**-47.5**	(4.4)	**22.9**	(1.3)	**-25.1**	(1.8)	**-65.6**	(9.0)	-3.3	(3.6)	**-8.5**	(3.9)	2.4	(6.4)	0.31
Colombia	**434.8**	(5.5)	**32.7**	(5.4)	**-26.4**	(7.3)	**-20.9**	(8.4)	**25.1**	(1.7)	**-10.1**	(2.9)	**-66.0**	(14.4)	-5.3	(6.5)	-1.5	(6.8)	-0.4	(8.2)	0.25
Dubai (UAE)	**366.4**	(6.9)	**59.3**	(6.5)	14.9	(10.4)	**-44.6**	(11.5)	**37.0**	(1.5)	**-48.3**	(2.5)	**72.1**	(2.9)	**-19.5**	(9.3)	9.8	(10.0)	-2.5	(15.9)	0.32
Hong Kong-China	**506.7**	(8.4)	**66.6**	(8.3)	8.1	(9.9)	c	c	**14.8**	(2.0)	**-30.4**	(3.6)	2.8	(4.3)	5.0	(11.3)	-13.4	(11.3)	4.1	(13.2)	0.15
Indonesia	**396.2**	(10.9)	**52.6**	(10.0)	19.4	(10.1)	**31.1**	(8.5)	**17.0**	(2.0)	**-34.2**	(3.0)	6.6	(4.1)	2.9	(8.0)	-5.2	(8.2)	-15.2	(9.8)	0.21
Latvia	**460.1**	(8.3)	**58.3**	(8.3)	-14.4	(14.5)	c	c	**26.0**	(2.4)	**-42.6**	(2.6)	-7.6	(5.5)	2.1	(10.7)	-5.2	(11.4)	-6.7	(17.9)	0.26
Macao-China	**484.0**	(3.1)	**55.3**	(2.7)	-8.1	(4.7)	**-37.7**	(3.8)	**6.5**	(1.1)	**-25.4**	(1.7)	**-5.4**	(2.0)	-4.4	(3.0)	-0.8	(3.4)	-5.3	(5.1)	0.25
Panama	**376.0**	(9.0)	**62.9**	(8.1)	-19.1	(12.4)	-1.3	(9.3)	**26.8**	(2.8)	**-20.6**	(6.6)	**-39.7**	(14.4)	-6.4	(7.9)	0.6	(9.8)	3.6	(17.0)	0.34
Peru	**407.0**	(7.2)	**36.1**	(5.9)	**-20.1**	(7.5)	**-46.1**	(7.7)	**33.0**	(1.9)	**-16.4**	(2.9)	**-59.0**	(7.3)	-2.5	(7.6)	-1.3	(7.3)	-4.9	(10.1)	0.37
Qatar	**296.8**	(5.1)	**65.8**	(5.1)	**23.7**	(7.9)	**-31.7**	(6.9)	**19.5**	(1.1)	**-39.1**	(1.7)	**51.8**	(2.0)	-3.6	(7.5)	-2.5	(8.0)	-3.0	(12.7)	0.18
Trinidad and Tobago	**367.7**	(16.8)	**123.0**	(16.7)	**42.1**	(17.6)	-25.7	(16.8)	**33.0**	(1.8)	**-50.5**	(2.9)	**-25.8**	(9.4)	-15.1	(18.5)	11.6	(18.7)	19.8	(19.0)	0.29
Tunisia	**400.3**	(4.6)	**66.3**	(5.3)	**-37.8**	(5.7)	**-48.3**	(4.6)	**9.9**	(1.3)	**-20.2**	(2.1)	c	c	**-8.2**	(2.8)	2.0	(4.0)	9.7	(6.3)	0.38
Uruguay	**431.9**	(4.9)	**60.0**	(5.1)	**-34.7**	(6.1)	**-45.9**	(5.2)	**23.6**	(1.4)	**-33.8**	(2.8)	**-20.8**	(8.7)	-5.0	(4.8)	3.2	(5.7)	-0.3	(8.4)	0.41

Note: The outcome variable is student performance in reading (PV1READ to PV5READ). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.8. **Students' attitudes towards school and grade repetition, Model 8, Model 9**

	Model 8										Model 9									
	Intercept		drepN		drep1		drep12		R^2		Intercept		drepN		drep1		drep12		ESCS		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.			Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**-0.17**	(0.07)	**0.26**	(0.06)	**0.32**	(0.10)	0.01	(0.26)	0.01		**-0.17**	(0.07)	**0.26**	(0.06)	**0.32**	(0.10)	0.01	(0.27)	-0.01	(0.02)	0.01
Belgium	**-0.14**	(0.03)	0.02	(0.03)	**0.15**	(0.04)	-0.05	(0.08)	0.00		**-0.14**	(0.03)	0.01	(0.03)	**0.16**	(0.04)	-0.03	(0.08)	**0.03**	(0.01)	0.00
Chile	**0.15**	(0.04)	**0.19**	(0.05)	-0.07	(0.07)	-0.06	(0.18)	0.01		**0.13**	(0.05)	**0.20**	(0.05)	-0.07	(0.07)	-0.07	(0.18)	-0.03	(0.02)	0.01
France	**-0.13**	(0.04)	**0.27**	(0.05)	**0.23**	(0.07)	0.13	(0.16)	0.01		**-0.13**	(0.04)	**0.27**	(0.05)	**0.24**	(0.07)	0.13	(0.16)	0.02	(0.02)	0.01
Germany	**-0.25**	(0.05)	**0.16**	(0.05)	**0.21**	(0.07)	-0.22	(0.20)	0.00		**-0.24**	(0.05)	**0.17**	(0.05)	**0.17**	(0.07)	-0.26	(0.21)	**-0.09**	(0.02)	0.01
Hungary	**-0.19**	(0.09)	0.16	(0.10)	0.05	(0.10)	0.08	(0.16)	0.00		-0.18	(0.10)	0.15	(0.10)	0.06	(0.10)	0.08	(0.16)	0.01	(0.02)	0.00
Ireland	0.23	(0.17)	-0.08	(0.17)	-0.17	(0.18)	c	c	0.00		0.27	(0.18)	-0.14	(0.18)	-0.20	(0.19)	c	c	**0.14**	(0.02)	0.02
Italy	**-0.12**	(0.02)	**0.17**	(0.03)	0.20	(0.16)	-0.33	(0.26)	0.00		**-0.12**	(0.02)	**0.17**	(0.03)	0.19	(0.16)	-0.35	(0.26)	-0.01	(0.01)	0.00
Luxembourg	**-0.28**	(0.04)	**0.16**	(0.05)	**0.25**	(0.06)	0.16	(0.13)	0.00		**-0.27**	(0.04)	**0.18**	(0.05)	**0.21**	(0.06)	0.13	(0.13)	**-0.06**	(0.02)	0.01
Mexico	**0.25**	(0.05)	**0.28**	(0.05)	**-0.19**	(0.05)	**-0.50**	(0.10)	0.04		**0.30**	(0.05)	**0.29**	(0.05)	**-0.15**	(0.05)	**-0.48**	(0.10)	**0.05**	(0.01)	0.04
Netherlands	**-0.29**	(0.06)	0.06	(0.06)	0.10	(0.07)	c	c	0.00		**-0.30**	(0.06)	0.06	(0.06)	0.11	(0.07)	c	c	**0.03**	(0.01)	0.00
Portugal	**0.12**	(0.05)	**0.32**	(0.05)	**0.15**	(0.06)	0.08	(0.08)	0.02		**0.11**	(0.05)	**0.33**	(0.05)	**0.14**	(0.06)	0.07	(0.08)	-0.01	(0.01)	0.02
Spain	-0.01	(0.03)	**0.25**	(0.03)	**0.19**	(0.05)	0.02	(0.05)	0.01		0.00	(0.03)	**0.25**	(0.03)	**0.19**	(0.05)	0.02	(0.05)	0.01	(0.01)	0.01
Switzerland	**-0.11**	(0.05)	**0.11**	(0.04)	**0.27**	(0.05)	0.13	(0.18)	0.00		**-0.10**	(0.05)	**0.11**	(0.04)	**0.25**	(0.05)	0.12	(0.18)	-0.03	(0.02)	0.00
Turkey	0.08	(0.05)	**0.13**	(0.05)	0.21	(0.19)	c	c	0.00		-0.08	(0.06)	**0.16**	(0.05)	0.14	(0.18)	c	c	**-0.12**	(0.02)	0.02
United States	0.08	(0.08)	0.15	(0.09)	0.13	(0.11)	-0.15	(0.18)	0.00		0.13	(0.07)	0.08	(0.08)	0.11	(0.10)	-0.17	(0.18)	**0.12**	(0.02)	0.01
Partners																					
Argentina	**0.08**	(0.03)	0.05	(0.05)	0.10	(0.07)	-0.14	(0.11)	0.00		0.06	(0.03)	0.06	(0.04)	0.08	(0.07)	-0.15	(0.11)	-0.03	(0.02)	0.00
Brazil	**0.37**	(0.03)	**0.12**	(0.03)	-0.02	(0.04)	**-0.16**	(0.05)	0.01		**0.37**	(0.03)	**0.12**	(0.03)	-0.02	(0.04)	**-0.16**	(0.05)	0.00	(0.01)	0.01
Colombia	**0.38**	(0.05)	**0.15**	(0.05)	-0.11	(0.06)	0.02	(0.10)	0.01		**0.40**	(0.05)	**0.14**	(0.05)	-0.10	(0.06)	0.03	(0.10)	0.02	(0.02)	0.01
Dubai (UAE)	**-0.12**	(0.05)	**0.21**	(0.05)	0.02	(0.07)	-0.14	(0.14)	0.01		**-0.12**	(0.05)	**0.21**	(0.05)	0.02	(0.07)	-0.15	(0.14)	-0.01	(0.02)	0.01
Hong Kong-China	**-0.59**	(0.05)	**0.11**	(0.05)	**0.13**	(0.05)	c	c	0.00		**-0.55**	(0.05)	**0.10**	(0.05)	**0.15**	(0.05)	c	c	**0.04**	(0.01)	0.00
Indonesia	0.05	(0.09)	**0.48**	(0.09)	**0.27**	(0.11)	**0.42**	(0.12)	0.01		0.07	(0.09)	**0.48**	(0.09)	**0.28**	(0.11)	**0.43**	(0.12)	0.01	(0.01)	0.01
Latvia	0.10	(0.08)	0.11	(0.08)	-0.19	(0.12)	c	c	0.01		0.11	(0.08)	0.11	(0.07)	-0.19	(0.12)	c	c	0.01	(0.02)	0.01
Macao-China	**-0.49**	(0.02)	**0.06**	(0.03)	0.05	(0.03)	-0.01	(0.04)	0.00		**-0.45**	(0.02)	0.05	(0.03)	0.07	(0.04)	0.00	(0.04)	**0.06**	(0.01)	0.01
Panama	**0.23**	(0.08)	**0.26**	(0.08)	-0.05	(0.11)	-0.02	(0.12)	0.02		**0.27**	(0.08)	**0.24**	(0.07)	-0.02	(0.11)	-0.02	(0.12)	**0.05**	(0.02)	0.02
Peru	**0.16**	(0.05)	0.08	(0.06)	**-0.17**	(0.06)	**-0.33**	(0.10)	0.01		**0.24**	(0.06)	0.06	(0.06)	**-0.14**	(0.06)	**-0.32**	(0.10)	**0.05**	(0.01)	0.02
Qatar	**-0.24**	(0.04)	**0.13**	(0.04)	0.01	(0.06)	-0.14	(0.08)	0.00		**-0.25**	(0.04)	**0.12**	(0.04)	0.01	(0.06)	-0.14	(0.08)	**0.04**	(0.01)	0.00
Trinidad and Tobago	-0.17	(0.13)	**0.71**	(0.13)	**0.55**	(0.14)	-0.33	(0.18)	0.02		-0.17	(0.13)	**0.71**	(0.13)	**0.55**	(0.14)	-0.33	(0.18)	0.00	(0.02)	0.02
Tunisia	**0.43**	(0.04)	**0.18**	(0.05)	0.06	(0.05)	**-0.20**	(0.07)	0.01		**0.39**	(0.05)	**0.19**	(0.05)	0.05	(0.05)	**-0.21**	(0.07)	-0.02	(0.01)	0.02
Uruguay	**0.11**	(0.03)	0.08	(0.04)	0.09	(0.05)	-0.11	(0.06)	0.00		**0.09**	(0.04)	**0.09**	(0.04)	0.08	(0.05)	-0.12	(0.06)	-0.02	(0.02)	0.00

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.8. **Students' attitudes towards school and grade repetition, Model 10** *(cont.)*

	Model 10																
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		R^2
	Intercept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																	
Austria	**-0.21**	(0.07)	**0.28**	(0.06)	**0.32**	(0.10)	-0.01	(0.26)	0.01	(0.02)	-0.02	(0.04)	**0.20**	(0.06)	0.01	(0.04)	0.01
Belgium	**-0.14**	(0.04)	0.01	(0.03)	**0.15**	(0.04)	-0.04	(0.08)	**0.04**	(0.01)	**-0.05**	(0.02)	**0.07**	(0.03)	0.01	(0.02)	0.01
Chile	0.10	(0.05)	**0.21**	(0.05)	-0.07	(0.07)	-0.07	(0.18)	-0.03	(0.02)	**0.08**	(0.03)	**-0.35**	(0.15)	-0.04	(0.03)	0.01
France	-0.03	(0.05)	**0.25**	(0.05)	**0.24**	(0.07)	0.16	(0.15)	**0.05**	(0.02)	**-0.19**	(0.03)	**0.25**	(0.05)	-0.04	(0.04)	0.03
Germany	**-0.18**	(0.06)	**0.16**	(0.06)	**0.16**	(0.07)	-0.29	(0.21)	**-0.07**	(0.02)	**-0.14**	(0.04)	**0.11**	(0.05)	0.00	(0.03)	0.02
Hungary	-0.13	(0.10)	0.14	(0.10)	0.04	(0.10)	0.08	(0.16)	0.01	(0.02)	**-0.09**	(0.04)	0.04	(0.08)	0.01	(0.03)	0.00
Ireland	0.31	(0.18)	-0.15	(0.18)	-0.21	(0.19)	c	c	**0.14**	(0.02)	-0.03	(0.04)	-0.03	(0.06)	-0.02	(0.04)	0.02
Italy	-0.03	(0.03)	**0.16**	(0.03)	0.20	(0.15)	-0.31	(0.26)	-0.01	(0.01)	**-0.12**	(0.02)	-0.01	(0.03)	-0.02	(0.01)	0.01
Luxembourg	**-0.28**	(0.08)	**0.17**	(0.05)	**0.21**	(0.06)	0.15	(0.13)	**-0.06**	(0.02)	**-0.12**	(0.03)	0.09	(0.06)	-0.03	(0.03)	0.01
Mexico	**0.46**	(0.05)	**0.24**	(0.05)	**-0.16**	(0.05)	**-0.44**	(0.10)	**0.05**	(0.01)	**-0.17**	(0.02)	**-0.29**	(0.04)	**-0.07**	(0.02)	0.05
Netherlands	**-0.36**	(0.07)	0.06	(0.06)	0.09	(0.07)	c	c	**0.05**	(0.01)	0.02	(0.03)	**0.16**	(0.04)	**0.07**	(0.02)	0.01
Portugal	**0.23**	(0.05)	**0.30**	(0.05)	**0.14**	(0.06)	0.09	(0.08)	0.00	(0.01)	**-0.20**	(0.03)	0.05	(0.06)	-0.02	(0.03)	0.03
Spain	**0.10**	(0.04)	**0.23**	(0.03)	**0.19**	(0.05)	0.02	(0.05)	0.01	(0.01)	**-0.14**	(0.02)	**-0.09**	(0.03)	0.01	(0.02)	0.02
Switzerland	-0.10	(0.05)	**0.11**	(0.04)	**0.25**	(0.05)	0.08	(0.18)	0.00	(0.02)	**-0.17**	(0.03)	**0.18**	(0.04)	**0.06**	(0.03)	0.02
Turkey	-0.01	(0.07)	**0.14**	(0.05)	0.15	(0.18)	c	c	**-0.12**	(0.02)	**-0.13**	(0.03)	-0.10	(0.08)	0.03	(0.03)	0.02
United States	**0.16**	(0.08)	0.07	(0.08)	0.10	(0.10)	-0.18	(0.18)	**0.14**	(0.02)	-0.06	(0.03)	**0.12**	(0.04)	-0.05	(0.04)	0.02
Partners																	
Argentina	0.03	(0.04)	0.06	(0.04)	0.08	(0.07)	-0.15	(0.11)	-0.03	(0.02)	0.02	(0.03)	0.09	(0.11)	0.04	(0.03)	0.00
Brazil	**0.40**	(0.03)	**0.11**	(0.03)	-0.02	(0.04)	**-0.15**	(0.05)	0.00	(0.01)	**-0.06**	(0.02)	**-0.39**	(0.12)	0.01	(0.02)	0.01
Colombia	**0.42**	(0.06)	**0.14**	(0.05)	-0.10	(0.06)	0.03	(0.10)	0.02	(0.02)	-0.02	(0.04)	-0.16	(0.27)	0.00	(0.03)	0.01
Dubai (UAE)	-0.06	(0.06)	**0.19**	(0.05)	0.00	(0.07)	-0.13	(0.14)	-0.01	(0.02)	**-0.12**	(0.03)	0.03	(0.04)	-0.02	(0.03)	0.01
Hong Kong-China	**-0.59**	(0.05)	**0.10**	(0.05)	**0.14**	(0.05)	c	c	**0.04**	(0.01)	0.01	(0.02)	0.04	(0.02)	0.03	(0.02)	0.00
Indonesia	0.09	(0.10)	**0.45**	(0.09)	**0.26**	(0.11)	**0.40**	(0.12)	**0.03**	(0.01)	**-0.09**	(0.03)	**0.10**	(0.04)	0.04	(0.03)	0.02
Latvia	**0.22**	(0.08)	0.06	(0.07)	-0.20	(0.12)	c	c	0.02	(0.02)	**-0.16**	(0.04)	0.00	(0.06)	0.01	(0.03)	0.01
Macao-China	**-0.47**	(0.03)	0.05	(0.03)	0.07	(0.04)	0.00	(0.04)	**0.06**	(0.01)	-0.01	(0.02)	0.03	(0.02)	0.00	(0.02)	0.01
Panama	**0.28**	(0.09)	**0.25**	(0.07)	0.01	(0.12)	-0.03	(0.11)	**0.05**	(0.02)	0.07	(0.06)	**-0.40**	(0.14)	-0.03	(0.05)	0.03
Peru	**0.28**	(0.06)	0.04	(0.06)	**-0.14**	(0.06)	**-0.31**	(0.10)	**0.04**	(0.01)	**-0.07**	(0.03)	**-0.35**	(0.08)	0.02	(0.03)	0.03
Qatar	**-0.38**	(0.04)	0.07	(0.04)	-0.01	(0.06)	-0.12	(0.08)	**0.03**	(0.01)	-0.04	(0.02)	**0.27**	(0.02)	0.04	(0.02)	0.02
Trinidad and Tobago	-0.01	(0.14)	**0.66**	(0.14)	**0.52**	(0.14)	-0.34	(0.18)	0.00	(0.02)	**-0.19**	(0.04)	-0.11	(0.10)	-0.04	(0.03)	0.03
Tunisia	**0.52**	(0.05)	**0.15**	(0.05)	0.03	(0.05)	**-0.22**	(0.07)	-0.02	(0.01)	**-0.19**	(0.03)	c	c	-0.02	(0.03)	0.03
Uruguay	**0.12**	(0.05)	**0.08**	(0.04)	0.08	(0.05)	-0.11	(0.06)	-0.02	(0.02)	-0.04	(0.03)	-0.14	(0.08)	0.00	(0.03)	0.00

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.9. **Students' attitudes towards school, grade repetition and students' background characteristics, Model 11**

	Model 11																				
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*ESCS		drep1*ESCS		R^2
	Inter-cept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**-0.22**	(0.07)	**0.28**	(0.06)	**0.30**	(0.11)	0.02	(0.26)	0.09	(0.08)	-0.02	(0.04)	**0.20**	(0.06)	0.01	(0.04)	-0.08	(0.08)	-0.13	(0.15)	0.01
Belgium	**-0.14**	(0.04)	0.00	(0.04)	**0.14**	(0.04)	-0.04	(0.08)	0.03	(0.03)	**-0.05**	(0.02)	**0.07**	(0.03)	0.01	(0.02)	0.02	(0.03)	-0.04	(0.06)	0.01
Chile	0.08	(0.06)	**0.24**	(0.06)	-0.04	(0.09)	-0.07	(0.18)	-0.06	(0.05)	**0.08**	(0.03)	**-0.35**	(0.15)	-0.04	(0.03)	0.03	(0.05)	0.04	(0.07)	0.01
France	-0.02	(0.06)	**0.24**	(0.05)	**0.19**	(0.08)	0.17	(0.15)	0.08	(0.06)	**-0.20**	(0.03)	**0.24**	(0.05)	-0.04	(0.04)	-0.03	(0.06)	-0.09	(0.08)	0.03
Germany	**-0.18**	(0.06)	**0.16**	(0.06)	**0.17**	(0.08)	-0.29	(0.21)	-0.08	(0.05)	**-0.14**	(0.04)	**0.11**	(0.05)	0.00	(0.03)	0.01	(0.06)	0.07	(0.08)	0.02
Hungary	-0.15	(0.11)	0.16	(0.11)	0.06	(0.15)	0.08	(0.16)	-0.02	(0.09)	**-0.09**	(0.04)	0.03	(0.08)	0.01	(0.03)	0.03	(0.09)	0.02	(0.14)	0.00
Ireland	0.38	(0.21)	-0.21	(0.20)	-0.27	(0.22)	c	c	**0.37**	(0.18)	-0.03	(0.04)	-0.03	(0.06)	-0.02	(0.04)	-0.23	(0.18)	-0.23	(0.20)	0.02
Italy	-0.03	(0.03)	**0.16**	(0.03)	-0.11	(0.17)	-0.31	(0.25)	-0.01	(0.03)	**-0.12**	(0.02)	-0.01	(0.03)	-0.02	(0.01)	0.00	(0.03)	**-0.34**	(0.13)	0.01
Luxembourg	**-0.28**	(0.08)	**0.19**	(0.05)	**0.27**	(0.07)	0.16	(0.13)	-0.04	(0.04)	**-0.12**	(0.03)	0.09	(0.06)	-0.03	(0.03)	-0.05	(0.04)	0.11	(0.06)	0.02
Mexico	**0.53**	(0.04)	**0.18**	(0.04)	**-0.32**	(0.06)	**-0.40**	(0.10)	**0.12**	(0.03)	**-0.17**	(0.02)	**-0.29**	(0.04)	**-0.07**	(0.02)	**-0.07**	(0.03)	**-0.12**	(0.03)	0.05
Netherlands	**-0.36**	(0.07)	0.06	(0.07)	0.09	(0.07)	c	c	0.05	(0.08)	0.02	(0.03)	**0.16**	(0.04)	**0.07**	(0.02)	0.00	(0.09)	0.01	(0.08)	0.01
Portugal	**0.23**	(0.06)	**0.30**	(0.06)	0.16	(0.09)	0.09	(0.08)	0.00	(0.04)	**-0.20**	(0.03)	0.05	(0.06)	-0.02	(0.03)	0.00	(0.04)	0.02	(0.07)	0.03
Spain	**0.11**	(0.04)	**0.22**	(0.04)	**0.20**	(0.07)	0.02	(0.05)	0.02	(0.02)	**-0.14**	(0.02)	**-0.09**	(0.03)	0.01	(0.02)	-0.02	(0.03)	0.01	(0.06)	0.02
Switzerland	-0.10	(0.05)	**0.11**	(0.04)	**0.24**	(0.06)	0.09	(0.18)	0.00	(0.06)	**-0.17**	(0.03)	**0.18**	(0.04)	**0.06**	(0.03)	0.00	(0.06)	-0.05	(0.08)	0.02
Turkey	0.18	(0.10)	-0.06	(0.10)	-0.44	(0.49)	c	c	0.01	(0.06)	**-0.13**	(0.03)	-0.11	(0.08)	0.03	(0.03)	**-0.14**	(0.06)	-0.35	(0.22)	0.03
United States	**0.17**	(0.08)	0.06	(0.09)	0.10	(0.12)	-0.18	(0.17)	**0.18**	(0.09)	-0.06	(0.03)	**0.12**	(0.04)	-0.05	(0.04)	-0.04	(0.09)	0.01	(0.13)	0.02
Partners																					
Argentina	0.04	(0.04)	0.04	(0.05)	0.04	(0.11)	-0.15	(0.11)	0.00	(0.04)	0.02	(0.03)	0.09	(0.11)	0.04	(0.03)	-0.03	(0.04)	-0.04	(0.07)	0.00
Brazil	**0.43**	(0.03)	0.08	(0.04)	-0.11	(0.06)	**-0.14**	(0.05)	0.03	(0.02)	**-0.06**	(0.02)	**-0.40**	(0.12)	0.01	(0.02)	-0.03	(0.02)	**-0.06**	(0.03)	0.01
Colombia	**0.45**	(0.06)	0.11	(0.06)	**-0.17**	(0.08)	0.04	(0.10)	0.05	(0.03)	-0.02	(0.04)	-0.17	(0.27)	0.00	(0.03)	-0.03	(0.04)	-0.05	(0.04)	0.01
Dubai (UAE)	-0.06	(0.06)	**0.20**	(0.06)	0.00	(0.07)	-0.13	(0.14)	0.01	(0.05)	**-0.12**	(0.03)	0.03	(0.04)	-0.02	(0.03)	-0.02	(0.06)	-0.04	(0.08)	0.01
Hong Kong-China	**-0.59**	(0.06)	0.10	(0.06)	0.13	(0.08)	c	c	0.04	(0.05)	0.01	(0.02)	0.04	(0.02)	0.03	(0.02)	0.00	(0.05)	-0.01	(0.06)	0.00
Indonesia	0.11	(0.12)	**0.43**	(0.13)	**0.27**	(0.13)	**0.40**	(0.13)	0.05	(0.05)	**-0.09**	(0.03)	**0.10**	(0.04)	0.04	(0.03)	-0.02	(0.06)	0.00	(0.06)	0.02
Latvia	**0.22**	(0.09)	0.06	(0.08)	-0.20	(0.15)	c	c	0.03	(0.11)	**-0.16**	(0.04)	0.00	(0.06)	0.01	(0.03)	-0.01	(0.11)	0.00	(0.14)	0.01
Macao-China	**-0.50**	(0.03)	**0.09**	(0.04)	0.10	(0.05)	0.00	(0.04)	0.02	(0.02)	-0.01	(0.02)	0.03	(0.02)	0.00	(0.02)	**0.07**	(0.03)	0.04	(0.04)	0.01
Panama	**0.26**	(0.09)	**0.26**	(0.08)	0.27	(0.18)	-0.03	(0.11)	0.02	(0.04)	0.06	(0.06)	**-0.42**	(0.14)	-0.03	(0.05)	0.01	(0.05)	**0.18**	(0.07)	0.03
Peru	0.16	(0.09)	0.17	(0.09)	0.02	(0.11)	**-0.34**	(0.10)	-0.04	(0.04)	**-0.06**	(0.03)	**-0.35**	(0.08)	0.02	(0.03)	**0.09**	(0.04)	**0.11**	(0.05)	0.03
Qatar	**-0.38**	(0.04)	0.07	(0.04)	0.00	(0.07)	-0.12	(0.08)	0.05	(0.05)	-0.04	(0.02)	**0.27**	(0.02)	0.04	(0.02)	-0.02	(0.05)	-0.04	(0.06)	0.02
Trinidad and Tobago	0.08	(0.20)	**0.56**	(0.20)	**0.49**	(0.21)	-0.36	(0.19)	0.11	(0.11)	**-0.19**	(0.04)	-0.12	(0.10)	-0.04	(0.03)	-0.14	(0.11)	-0.02	(0.12)	0.03
Tunisia	**0.56**	(0.06)	0.10	(0.07)	0.01	(0.09)	**-0.20**	(0.07)	0.01	(0.02)	**-0.19**	(0.03)	c	c	-0.02	(0.03)	-0.05	(0.03)	-0.01	(0.05)	0.03
Uruguay	**0.13**	(0.06)	0.07	(0.05)	0.12	(0.08)	-0.11	(0.06)	-0.01	(0.03)	-0.05	(0.03)	-0.14	(0.08)	0.00	(0.03)	-0.01	(0.03)	0.02	(0.05)	0.00

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.9. **Students' attitudes towards school, grade repetition and students' background characteristics, Model 12** *(cont.)*

	Model 12																								
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*MALE		drep1*MALE		R^2
	Inter-cept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**-0.22**	(0.09)	**0.28**	(0.09)	**0.52**	(0.17)	-0.01	(0.26)	0.01	(0.02)	-0.01	(0.11)	**0.19**	(0.06)	0.01	(0.04)	0.00	(0.11)	-0.37	(0.27)	0.01
Belgium	**-0.13**	(0.05)	0.00	(0.05)	**0.12**	(0.05)	-0.04	(0.08)	**0.04**	(0.01)	-0.07	(0.05)	**0.07**	(0.03)	0.01	(0.02)	0.01	(0.06)	0.07	(0.07)	0.01
Chile	0.12	(0.06)	**0.19**	(0.07)	0.07	(0.11)	-0.07	(0.18)	-0.03	(0.02)	0.06	(0.08)	**-0.34**	(0.15)	-0.04	(0.03)	0.05	(0.09)	-0.23	(0.13)	0.01
France	-0.03	(0.06)	**0.25**	(0.05)	**0.24**	(0.07)	0.16	(0.15)	**0.05**	(0.02)	**-0.20**	(0.07)	**0.25**	(0.05)	-0.04	(0.04)	0.01	(0.08)	-0.01	(0.13)	0.03
Germany	**-0.24**	(0.07)	**0.22**	(0.07)	**0.31**	(0.10)	-0.27	(0.21)	**-0.07**	(0.02)	-0.04	(0.09)	**0.11**	(0.05)	0.00	(0.03)	-0.11	(0.10)	-0.28	(0.15)	0.02
Hungary	-0.23	(0.16)	0.25	(0.16)	0.16	(0.19)	0.08	(0.16)	0.01	(0.02)	0.07	(0.16)	0.04	(0.08)	0.01	(0.03)	-0.17	(0.16)	-0.20	(0.27)	0.01
Ireland	0.27	(0.23)	-0.10	(0.23)	-0.17	(0.23)	c	c	**0.14**	(0.02)	0.06	(0.27)	-0.03	(0.06)	-0.02	(0.04)	-0.09	(0.27)	-0.08	(0.29)	0.02
Italy	0.01	(0.03)	**0.11**	(0.03)	0.03	(0.41)	-0.29	(0.26)	-0.01	(0.01)	**-0.19**	(0.04)	-0.01	(0.03)	-0.02	(0.01)	0.08	(0.04)	0.27	(0.47)	0.01
Luxembourg	**-0.32**	(0.09)	**0.20**	(0.07)	**0.35**	(0.08)	0.14	(0.13)	**-0.06**	(0.02)	-0.04	(0.08)	0.08	(0.06)	-0.02	(0.03)	-0.05	(0.09)	**-0.26**	(0.12)	0.01
Mexico	**0.44**	(0.05)	**0.27**	(0.05)	**-0.16**	(0.07)	**-0.44**	(0.10)	**0.05**	(0.01)	**-0.14**	(0.06)	**-0.29**	(0.04)	**-0.07**	(0.02)	-0.04	(0.06)	0.01	(0.07)	0.05
Netherlands	**-0.38**	(0.11)	0.09	(0.10)	0.07	(0.12)	c	c	**0.05**	(0.01)	0.06	(0.13)	**0.16**	(0.04)	**0.07**	(0.02)	-0.06	(0.11)	0.04	(0.15)	0.01
Portugal	**0.23**	(0.07)	**0.31**	(0.07)	0.13	(0.10)	0.09	(0.08)	0.00	(0.01)	**-0.19**	(0.07)	0.05	(0.06)	-0.02	(0.03)	-0.01	(0.08)	0.02	(0.11)	0.03
Spain	0.08	(0.04)	**0.25**	(0.04)	**0.22**	(0.08)	0.02	(0.05)	0.01	(0.01)	**-0.11**	(0.05)	**-0.08**	(0.03)	0.01	(0.02)	-0.04	(0.06)	-0.06	(0.11)	0.02
Switzerland	-0.07	(0.06)	0.07	(0.06)	**0.27**	(0.08)	0.08	(0.18)	0.00	(0.02)	**-0.23**	(0.07)	**0.18**	(0.04)	**0.06**	(0.03)	0.08	(0.07)	-0.04	(0.12)	0.02
Turkey	-0.04	(0.10)	0.17	(0.09)	0.23	(0.28)	c	c	**-0.12**	(0.02)	-0.08	(0.10)	-0.11	(0.08)	0.03	(0.03)	-0.05	(0.10)	-0.11	(0.25)	0.02
United States	**0.21**	(0.10)	0.02	(0.11)	0.09	(0.15)	-0.17	(0.18)	**0.14**	(0.02)	-0.14	(0.13)	**0.12**	(0.04)	-0.05	(0.04)	0.09	(0.13)	0.02	(0.19)	0.02
Partners																					
Argentina	0.05	(0.05)	0.03	(0.06)	0.04	(0.10)	-0.15	(0.11)	-0.03	(0.02)	-0.03	(0.07)	0.09	(0.11)	0.04	(0.03)	0.06	(0.08)	0.07	(0.13)	0.00
Brazil	**0.37**	(0.03)	**0.14**	(0.04)	0.04	(0.05)	**-0.15**	(0.05)	0.00	(0.01)	-0.01	(0.05)	**-0.40**	(0.12)	0.01	(0.02)	-0.05	(0.06)	-0.11	(0.08)	0.01
Colombia	**0.44**	(0.05)	**0.13**	(0.06)	-0.17	(0.09)	0.03	(0.10)	0.02	(0.02)	-0.06	(0.08)	-0.15	(0.27)	0.00	(0.03)	0.02	(0.09)	0.13	(0.14)	0.01
Dubai (UAE)	0.02	(0.10)	0.11	(0.09)	-0.04	(0.12)	-0.12	(0.14)	-0.01	(0.02)	**-0.24**	(0.11)	0.03	(0.04)	-0.02	(0.03)	0.13	(0.12)	0.05	(0.16)	0.01
Hong Kong-China	**-0.57**	(0.06)	0.07	(0.06)	**0.14**	(0.07)	c	c	**0.04**	(0.01)	-0.02	(0.09)	0.04	(0.02)	0.03	(0.02)	0.04	(0.10)	-0.01	(0.10)	0.00
Indonesia	0.07	(0.14)	**0.47**	(0.14)	**0.31**	(0.16)	**0.40**	(0.12)	**0.03**	(0.01)	-0.06	(0.14)	**0.10**	(0.04)	0.04	(0.03)	-0.03	(0.14)	-0.07	(0.15)	0.02
Latvia	**0.40**	(0.11)	-0.12	(0.12)	**-0.49**	(0.17)	c	c	0.02	(0.02)	**-0.40**	(0.17)	0.01	(0.06)	0.01	(0.03)	0.25	(0.18)	0.41	(0.24)	0.02
Macao-China	**-0.47**	(0.03)	0.05	(0.03)	0.07	(0.04)	0.00	(0.04)	**0.06**	(0.01)	-0.01	(0.03)	0.03	(0.02)	0.00	(0.02)	0.00	(0.04)	0.00	(0.07)	0.01
Panama	**0.29**	(0.10)	**0.24**	(0.09)	-0.02	(0.14)	-0.03	(0.11)	**0.05**	(0.02)	0.05	(0.09)	**-0.40**	(0.14)	-0.03	(0.05)	0.01	(0.12)	0.06	(0.13)	0.03
Peru	**0.29**	(0.08)	0.03	(0.08)	**-0.20**	(0.09)	**-0.31**	(0.10)	**0.04**	(0.01)	-0.08	(0.08)	**-0.34**	(0.08)	0.02	(0.03)	0.00	(0.08)	0.11	(0.09)	0.03
Qatar	**-0.42**	(0.06)	0.11	(0.06)	0.00	(0.09)	-0.13	(0.08)	**0.03**	(0.01)	0.03	(0.07)	**0.27**	(0.02)	0.04	(0.02)	-0.08	(0.08)	-0.03	(0.13)	0.02
Trinidad and Tobago	-0.24	(0.17)	**0.89**	(0.17)	**0.79**	(0.17)	-0.33	(0.18)	0.00	(0.02)	0.16	(0.18)	-0.12	(0.10)	-0.04	(0.03)	-0.34	(0.19)	**-0.42**	(0.19)	0.03
Tunisia	**0.47**	(0.06)	**0.20**	(0.06)	**0.14**	(0.07)	**-0.21**	(0.07)	-0.02	(0.01)	**-0.12**	(0.06)	c	c	-0.02	(0.03)	-0.09	(0.07)	**-0.24**	(0.10)	0.03
Uruguay	**0.15**	(0.06)	0.05	(0.06)	0.04	(0.08)	-0.12	(0.06)	-0.02	(0.02)	-0.10	(0.08)	-0.15	(0.08)	0.00	(0.03)	0.07	(0.08)	0.09	(0.11)	0.00

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.9. **Students' attitudes towards school, grade repetition and students' background characteristics, Model 13** *(cont.)*

	Model 13																	
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN* MIGLAN		drep1* MIGLAN		R^2
	Inter-cept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																					
Austria	**-0.28**	(0.07)	**0.35**	(0.07)	**0.38**	(0.13)	-0.03	(0.25)	0.01	(0.02)	-0.02	(0.04)	**0.45**	(0.13)	0.01	(0.04)	**-0.29**	(0.13)	-0.23	(0.30)	0.01
Belgium	**-0.15**	(0.04)	0.03	(0.04)	**0.15**	(0.05)	-0.04	(0.08)	**0.04**	(0.01)	**-0.05**	(0.02)	**0.10**	(0.05)	0.01	(0.02)	-0.06	(0.05)	0.01	(0.10)	0.01
Chile	0.10	(0.05)	**0.21**	(0.05)	-0.07	(0.07)	-0.07	(0.18)	-0.03	(0.02)	**0.08**	(0.03)	-0.21	(0.41)	-0.04	(0.03)	-0.18	(0.40)	0.15	(0.52)	0.01
France	-0.03	(0.05)	**0.24**	(0.05)	**0.21**	(0.08)	0.16	(0.15)	**0.05**	(0.02)	**-0.19**	(0.03)	**0.20**	(0.10)	-0.05	(0.04)	0.03	(0.12)	0.13	(0.18)	0.03
Germany	**-0.17**	(0.06)	**0.15**	(0.06)	0.14	(0.09)	-0.28	(0.21)	**-0.07**	(0.02)	**-0.14**	(0.03)	0.06	(0.12)	0.00	(0.03)	0.06	(0.13)	0.07	(0.16)	0.02
Hungary	-0.13	(0.10)	0.14	(0.10)	0.08	(0.11)	0.06	(0.16)	0.01	(0.02)	**-0.09**	(0.04)	0.26	(0.60)	0.01	(0.03)	-0.19	(0.64)	-0.73	(0.72)	0.01
Ireland	0.27	(0.22)	-0.10	(0.22)	-0.17	(0.23)	c	c	**0.14**	(0.02)	-0.03	(0.04)	0.08	(0.27)	-0.02	(0.04)	-0.11	(0.29)	-0.11	(0.33)	0.02
Italy	-0.04	(0.03)	**0.17**	(0.03)	0.17	(0.16)	-0.32	(0.26)	-0.01	(0.01)	**-0.12**	(0.02)	0.03	(0.07)	-0.02	(0.01)	-0.06	(0.07)	0.03	(0.31)	0.01
Luxembourg	-0.29	(0.16)	0.18	(0.18)	0.22	(0.26)	0.15	(0.13)	**-0.06**	(0.02)	**-0.12**	(0.03)	0.10	(0.16)	-0.03	(0.03)	-0.01	(0.18)	-0.01	(0.25)	0.01
Mexico	**0.48**	(0.04)	**0.23**	(0.04)	**-0.18**	(0.05)	**-0.39**	(0.10)	**0.05**	(0.01)	**-0.17**	(0.02)	**-0.60**	(0.19)	**-0.07**	(0.02)	0.30	(0.20)	0.43	(0.23)	0.05
Netherlands	**-0.33**	(0.06)	0.03	(0.06)	0.04	(0.07)	c	c	**0.05**	(0.01)	0.02	(0.03)	-0.05	(0.15)	**0.07**	(0.02)	0.18	(0.17)	0.31	(0.17)	0.01
Portugal	**0.22**	(0.05)	**0.32**	(0.05)	**0.15**	(0.06)	0.09	(0.08)	0.00	(0.01)	**-0.20**	(0.03)	0.19	(0.14)	-0.02	(0.03)	-0.21	(0.14)	-0.11	(0.18)	0.03
Spain	0.08	(0.04)	**0.25**	(0.04)	**0.14**	(0.06)	0.02	(0.05)	0.01	(0.01)	**-0.14**	(0.02)	-0.03	(0.05)	0.01	(0.02)	-0.10	(0.05)	0.12	(0.10)	0.02
Switzerland	**-0.12**	(0.06)	**0.13**	(0.05)	**0.27**	(0.08)	0.07	(0.18)	0.00	(0.02)	**-0.17**	(0.03)	**0.24**	(0.11)	**0.06**	(0.03)	-0.07	(0.11)	-0.07	(0.14)	0.02
Turkey	-0.01	(0.07)	**0.13**	(0.06)	0.18	(0.15)	c	c	**-0.12**	(0.02)	**-0.13**	(0.03)	-0.11	(0.25)	0.03	(0.03)	0.03	(0.27)	-0.15	(0.49)	0.02
United States	**0.19**	(0.08)	0.03	(0.08)	0.10	(0.12)	-0.17	(0.18)	**0.14**	(0.02)	-0.06	(0.03)	-0.02	(0.19)	-0.05	(0.04)	0.16	(0.21)	0.04	(0.22)	0.02
Partners																					
Argentina	0.02	(0.04)	0.06	(0.04)	0.07	(0.08)	-0.15	(0.11)	-0.03	(0.02)	0.02	(0.03)	0.11	(0.13)	0.04	(0.03)	-0.04	(0.22)	0.04	(0.27)	0.00
Brazil	**0.40**	(0.03)	**0.11**	(0.03)	-0.01	(0.04)	**-0.15**	(0.05)	0.00	(0.01)	**-0.06**	(0.02)	-0.20	(0.24)	0.01	(0.02)	-0.27	(0.27)	-0.39	(0.29)	0.01
Colombia	**0.42**	(0.05)	**0.14**	(0.05)	-0.09	(0.06)	0.03	(0.10)	0.02	(0.02)	-0.03	(0.04)	-0.02	(0.37)	0.00	(0.03)	0.19	(0.47)	-0.95	(0.51)	0.01
Dubai (UAE)	-0.13	(0.10)	**0.27**	(0.10)	0.13	(0.12)	-0.11	(0.14)	-0.01	(0.02)	**-0.12**	(0.03)	0.14	(0.11)	-0.02	(0.03)	-0.11	(0.12)	-0.20	(0.15)	0.01
Hong Kong-China	**-0.59**	(0.07)	0.11	(0.07)	0.11	(0.09)	c	c	**0.04**	(0.01)	0.01	(0.02)	0.05	(0.10)	0.03	(0.02)	-0.02	(0.10)	0.03	(0.12)	0.00
Indonesia	-0.04	(0.10)	**0.59**	(0.11)	**0.41**	(0.12)	**0.33**	(0.12)	**0.03**	(0.01)	**-0.09**	(0.03)	**0.36**	(0.10)	0.04	(0.03)	**-0.27**	(0.11)	**-0.28**	(0.13)	0.02
Latvia	**0.22**	(0.08)	0.06	(0.08)	-0.18	(0.13)	c	c	0.02	(0.02)	**-0.16**	(0.04)	0.03	(0.15)	0.01	(0.03)	-0.02	(0.18)	-0.13	(0.29)	0.01
Macao-China	**-0.46**	(0.04)	0.05	(0.05)	0.05	(0.06)	0.00	(0.04)	**0.06**	(0.01)	-0.01	(0.02)	0.03	(0.04)	0.00	(0.02)	0.00	(0.06)	0.02	(0.07)	0.01
Panama	**0.32**	(0.08)	**0.21**	(0.07)	-0.11	(0.11)	-0.05	(0.11)	**0.05**	(0.02)	0.06	(0.05)	**-0.82**	(0.18)	-0.03	(0.06)	**0.40**	(0.20)	**0.96**	(0.22)	0.04
Peru	**0.27**	(0.06)	0.05	(0.06)	**-0.14**	(0.07)	**-0.31**	(0.10)	**0.04**	(0.01)	**-0.07**	(0.03)	-0.25	(0.20)	0.02	(0.03)	-0.13	(0.19)	-0.10	(0.20)	0.03
Qatar	**-0.31**	(0.06)	-0.02	(0.06)	-0.02	(0.09)	-0.13	(0.08)	**0.03**	(0.01)	-0.04	(0.02)	0.13	(0.08)	0.04	(0.02)	**0.16**	(0.08)	0.03	(0.12)	0.02
Trinidad and Tobago	0.01	(0.15)	**0.64**	(0.15)	**0.49**	(0.15)	-0.36	(0.19)	0.00	(0.02)	**-0.19**	(0.04)	-0.29	(0.27)	-0.04	(0.03)	0.11	(0.29)	0.34	(0.34)	0.03
Tunisia	**0.52**	(0.05)	**0.16**	(0.05)	0.03	(0.05)	**-0.22**	(0.07)	-0.02	(0.01)	**-0.19**	(0.03)	c	c	-0.02	(0.03)	c	c	c	c	0.03
Uruguay	**0.12**	(0.05)	**0.09**	(0.04)	0.09	(0.05)	-0.12	(0.06)	-0.02	(0.02)	-0.04	(0.03)	-0.01	(0.25)	0.00	(0.03)	-0.17	(0.26)	-0.17	(0.32)	0.00

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.9. **Students' attitudes towards school, grade repetition and students' background characteristics, Model 14** *(cont.)*

	Model 14																								
	Intercept		drepN		drep1		drep12		ESCS		MALE		MIGLAN		rAGE		drepN*rAGE		drep1*rAGE		R^2
	Inter-cept	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	Reg. coef.	S.E.	
OECD																									
Austria	-0.13	(0.08)	**0.19**	(0.08)	**0.27**	(0.12)	0.00	(0.26)	0.01	(0.02)	-0.02	(0.04)	**0.19**	(0.06)	-0.18	(0.15)	0.21	(0.16)	0.13	(0.19)	0.01
Belgium	**-0.12**	(0.05)	-0.02	(0.05)	**0.15**	(0.07)	-0.03	(0.08)	**0.04**	(0.01)	**-0.05**	(0.02)	**0.07**	(0.03)	-0.02	(0.05)	0.05	(0.05)	0.00	(0.08)	0.01
Chile	**0.14**	(0.06)	**0.17**	(0.06)	-0.08	(0.08)	-0.07	(0.18)	-0.03	(0.02)	**0.08**	(0.03)	**-0.35**	(0.15)	-0.10	(0.07)	0.08	(0.08)	0.02	(0.14)	0.01
France	-0.03	(0.06)	**0.23**	(0.07)	**0.26**	(0.08)	0.16	(0.15)	**0.05**	(0.02)	**-0.19**	(0.03)	**0.25**	(0.05)	-0.06	(0.08)	0.03	(0.09)	-0.03	(0.13)	0.03
Germany	**-0.17**	(0.07)	**0.14**	(0.07)	0.14	(0.08)	-0.28	(0.20)	**-0.07**	(0.02)	**-0.14**	(0.04)	**0.11**	(0.05)	-0.03	(0.07)	0.04	(0.08)	0.04	(0.12)	0.02
Hungary	-0.12	(0.11)	0.13	(0.11)	0.09	(0.16)	0.08	(0.16)	0.01	(0.02)	**-0.09**	(0.04)	0.04	(0.08)	0.00	(0.16)	0.02	(0.17)	-0.07	(0.26)	0.00
Ireland	**0.73**	(0.24)	**-0.57**	(0.24)	**-0.61**	(0.26)	c	c	**0.14**	(0.02)	-0.03	(0.04)	-0.03	(0.06)	**-0.75**	(0.26)	**0.74**	(0.26)	**0.68**	(0.30)	0.02
Italy	-0.02	(0.03)	**0.14**	(0.03)	**0.36**	(0.17)	-0.31	(0.26)	-0.01	(0.01)	**-0.12**	(0.02)	-0.01	(0.03)	-0.05	(0.04)	0.04	(0.05)	-0.31	(0.23)	0.01
Luxembourg	**-0.30**	(0.08)	**0.17**	(0.05)	**0.27**	(0.07)	0.14	(0.13)	**-0.06**	(0.02)	**-0.11**	(0.03)	0.09	(0.06)	0.01	(0.08)	-0.01	(0.08)	-0.15	(0.09)	0.01
Mexico	**0.45**	(0.04)	**0.26**	(0.05)	**-0.18**	(0.06)	**-0.44**	(0.10)	**0.05**	(0.01)	**-0.17**	(0.02)	**-0.29**	(0.04)	-0.04	(0.07)	-0.04	(0.08)	0.04	(0.09)	0.05
Netherlands	**-0.41**	(0.06)	0.12	(0.06)	0.12	(0.07)	c	c	**0.05**	(0.01)	0.02	(0.03)	**0.16**	(0.04)	0.22	(0.13)	-0.17	(0.13)	-0.10	(0.14)	0.01
Portugal	**0.25**	(0.05)	**0.27**	(0.06)	**0.18**	(0.07)	0.09	(0.09)	0.00	(0.01)	**-0.20**	(0.03)	0.05	(0.06)	-0.05	(0.08)	0.06	(0.08)	-0.06	(0.10)	0.03
Spain	**0.08**	(0.04)	**0.25**	(0.04)	**0.20**	(0.10)	0.02	(0.05)	0.01	(0.01)	**-0.14**	(0.02)	**-0.08**	(0.03)	0.03	(0.03)	-0.03	(0.04)	-0.03	(0.12)	0.02
Switzerland	-0.09	(0.07)	0.10	(0.07)	**0.20**	(0.07)	0.08	(0.18)	0.00	(0.02)	**-0.17**	(0.03)	**0.18**	(0.04)	0.03	(0.09)	0.02	(0.09)	0.09	(0.10)	0.02
Turkey	-0.01	(0.08)	0.14	(0.07)	0.02	(0.19)	c	c	**-0.12**	(0.02)	**-0.13**	(0.03)	-0.10	(0.08)	0.03	(0.12)	-0.01	(0.13)	0.29	(0.23)	0.03
United States	**0.18**	(0.09)	0.04	(0.09)	0.13	(0.12)	-0.18	(0.18)	**0.14**	(0.02)	-0.06	(0.03)	**0.12**	(0.04)	-0.10	(0.13)	0.06	(0.13)	-0.05	(0.16)	0.02
Partners																									
Argentina	0.09	(0.05)	-0.02	(0.06)	-0.06	(0.10)	-0.15	(0.11)	-0.03	(0.02)	0.02	(0.03)	0.09	(0.11)	-0.09	(0.06)	**0.15**	(0.07)	**0.26**	(0.12)	0.01
Brazil	**0.38**	(0.03)	**0.13**	(0.05)	0.05	(0.05)	**-0.15**	(0.05)	0.00	(0.01)	**-0.06**	(0.02)	**-0.39**	(0.12)	0.06	(0.04)	-0.05	(0.05)	-0.12	(0.06)	0.01
Colombia	**0.39**	(0.07)	**0.18**	(0.08)	-0.08	(0.08)	0.03	(0.10)	0.02	(0.02)	-0.02	(0.04)	-0.16	(0.27)	0.06	(0.09)	-0.08	(0.11)	-0.04	(0.13)	0.01
Dubai (UAE)	0.02	(0.07)	0.11	(0.07)	-0.02	(0.10)	-0.15	(0.14)	-0.01	(0.02)	**-0.12**	(0.03)	0.03	(0.04)	**-0.20**	(0.09)	**0.21**	(0.10)	0.08	(0.15)	0.01
Hong Kong-China	**-0.65**	(0.07)	**0.16**	(0.07)	**0.24**	(0.08)	c	c	**0.04**	(0.01)	0.01	(0.02)	0.04	(0.02)	0.17	(0.11)	-0.15	(0.11)	-0.22	(0.12)	0.01
Indonesia	0.05	(0.11)	**0.48**	(0.12)	**0.39**	(0.12)	**0.39**	(0.12)	**0.03**	(0.01)	**-0.09**	(0.03)	**0.10**	(0.04)	0.14	(0.11)	-0.08	(0.12)	**-0.24**	(0.12)	0.02
Latvia	0.20	(0.11)	0.08	(0.11)	-0.32	(0.17)	c	c	0.02	(0.02)	**-0.16**	(0.04)	0.00	(0.06)	0.05	(0.17)	-0.05	(0.17)	0.16	(0.25)	0.01
Macao-China	**-0.42**	(0.03)	-0.01	(0.04)	0.01	(0.05)	0.00	(0.04)	**0.06**	(0.01)	-0.01	(0.02)	0.03	(0.02)	**-0.08**	(0.03)	**0.12**	(0.04)	0.11	(0.07)	0.01
Panama	**0.30**	(0.12)	0.22	(0.12)	0.02	(0.17)	-0.03	(0.11)	**0.05**	(0.02)	0.06	(0.06)	**-0.40**	(0.14)	-0.06	(0.12)	0.05	(0.13)	-0.01	(0.16)	0.03
Peru	**0.30**	(0.07)	0.00	(0.07)	-0.13	(0.08)	**-0.31**	(0.10)	**0.04**	(0.01)	**-0.07**	(0.03)	**-0.35**	(0.08)	-0.03	(0.08)	0.07	(0.09)	-0.02	(0.13)	0.03
Qatar	**-0.35**	(0.06)	0.04	(0.06)	-0.15	(0.08)	-0.12	(0.08)	**0.03**	(0.01)	-0.04	(0.02)	**0.27**	(0.02)	-0.03	(0.07)	0.06	(0.08)	**0.29**	(0.12)	0.02
Trinidad and Tobago	-0.03	(0.14)	**0.69**	(0.15)	**0.52**	(0.14)	-0.35	(0.19)	0.00	(0.02)	**-0.19**	(0.04)	-0.11	(0.10)	0.00	(0.20)	-0.05	(0.21)	-0.01	(0.21)	0.03
Tunisia	**0.51**	(0.06)	**0.18**	(0.06)	0.00	(0.07)	**-0.22**	(0.07)	-0.02	(0.01)	**-0.19**	(0.03)	c	c	0.00	(0.05)	-0.05	(0.06)	0.04	(0.09)	0.03
Uruguay	**0.16**	(0.05)	0.04	(0.05)	0.00	(0.07)	-0.11	(0.06)	-0.02	(0.02)	-0.04	(0.03)	-0.14	(0.08)	-0.09	(0.06)	0.10	(0.07)	0.16	(0.09)	0.01

Note: The outcome variable is the index of student attitude towards school (ATSCHL). Values that are statistically significant are indicated in bold. When there are too few observations (fewer than 30 students or less than five schools with valid data) it is indicated as "c".

Table A.10. **Total number of observations and percentage of missing values**

	Number of total cases	DREPN/ DREP1/ DREP12	ESCS	MALE	MIGLAN	rAGE	ATSCHL	Missing in either: ESCS, MALE, MIGLAN or rAGE	Missing in either: ESCS, MALE, MIGLAN, rAGE or ATSCHL
		%	%	%	%	%	%	%	%
Austria	6 590	4.2	2.4	0.0	1.2	0.0	8.9	6.0	10.9
Belgium	8 501	4.1	1.5	0.0	0.8	0.0	8.3	5.2	9.8
Chile	5 669	2.0	1.6	0.0	1.6	0.0	5.5	2.2	6.2
France	4 298	1.5	1.7	0.0	0.6	0.0	4.4	2.9	6.6
Germany	4 979	10.4	8.5	0.0	7.6	0.0	14.5	12.2	16.2
Hungary	4 605	0.5	0.3	0.0	0.2	0.0	5.0	0.9	5.8
Ireland	3 937	2.5	2.4	0.0	2.2	0.0	6.2	3.0	6.9
Italy	30 905	0.8	0.2	0.0	0.4	0.0	3.4	1.3	4.6
Luxembourg	4 622	0.7	1.7	0.0	0.8	0.0	6.7	2.5	8.4
Mexico	38 250	1.7	0.5	0.0	0.6	0.0	4.9	2.1	6.6
Netherlands	4 760	3.3	1.1	0.0	0.6	0.0	7.6	4.0	8.5
Portugal	6 298	0.8	0.4	0.0	0.3	0.0	3.6	1.0	4.3
Spain	25 887	0.7	0.7	0.0	0.5	0.0	4.9	1.1	5.4
Switzerland	11 812	1.1	0.5	0.0	0.7	0.0	7.2	2.0	8.8
Turkey	4 996	0.7	0.6	0.0	0.2	0.0	5.7	1.2	6.5
United States	5 233	3.0	0.8	0.0	0.8	0.0	4.7	3.8	7.8
Argentina	4 596	2.1	0.8	0.0	0.7	0.0	4.4	3.2	6.7
Brazil	20 127	1.9	0.9	0.0	1.2	0.0	6.1	3.1	8.0
Colombia	7 921	1.0	0.4	0.0	0.4	0.0	6.3	1.2	7.0
Dubai (UAE)	5 620	1.4	0.6	0.0	1.2	0.0	6.8	2.4	8.4
Hong Kong-China	4 837	0.2	0.5	0.0	0.1	0.0	4.9	0.5	5.2
Indonesia	5 136	1.4	0.3	0.0	0.3	0.0	3.7	1.7	5.1
Latvia	4 502	0.4	0.6	0.0	0.2	0.0	3.9	1.1	4.8
Macao-China	5 952	0.2	0.2	0.0	0.3	0.0	4.5	0.5	4.9
Panama	3 969	8.7	7.9	0.0	7.1	0.0	14.9	10.7	18.5
Peru	5 985	1.3	0.7	0.0	1.0	0.0	9.4	1.9	10.4
Qatar	9 078	1.6	1.4	0.0	1.4	0.0	14.6	2.5	15.6
Trinidad and Tobago	4 778	1.8	2.9	0.0	1.7	0.0	6.9	4.3	9.6
Tunisia	4 955	1.2	0.3	0.0	0.1	0.0	3.7	1.4	4.8
Uruguay	5 957	0.7	1.1	0.0	0.7	0.0	7.2	1.4	8.0

APPENDIX B

Plausible values

Plausible Values are "random numbers drawn from the distribution of scores that could be reasonably assigned to each individual – that is, the marginal posterior distribution" (OECD, 2012). The PISA Data Analysis Manual compared the estimates based on PVs and those based on classical IRT estimates such as the Maximum Likelihood Estimates or Weighted Maximum Likelihood Estimates, and presented several methodological advantages of PVs (OECD, 2009). The manual also concluded that PVs return unbiased estimates of population performance parameters, percentage of students per sub-group, and bivariate or multivariate relationships between performance and background variables.*

For each assessment domain, five Plausible Values (PVs) are generated. For example, for reading, each student has five variables: ***PV1READ, PV2READ, PV3READ, PV4READ*** and ***PV5READ***.

The required statistic and its respective standard error have to be computed for each plausible value and combined at the very end to obtain the final estimate and its correct standard error. There are six major steps (OECD, 2009):

1. Any analysis that involves five PVs requires 405 estimates: 81 estimates (one with the student final weights and 80 with replicate weights) by five PVs.
2. The final estimate is the average of five estimates derived with each of five PVs with the student final weights.
3. The final sampling variance is the average of five sampling variances (i.e. a sampling variance should be computed for each of five PVs, first). See Section 3.2 on "Replicate Weights" for the equation on computing a sampling variance.
4. The measurement-error variance is derived by aggregating the differences between each of the five estimates with five PVs with the student final weights and the final estimate (see Step 2). For example, when population mean ($\hat{\mu}$) is estimated:

$$MeasurementErrorVariance = \frac{1}{4}\sum_{i=1}^{5}(\hat{\mu}_i - \hat{\mu})^2$$

* Regression coefficients are unbiased even with Weighted Maximum Likelihood Estimates (WLEs) or Expected *A Posteriori* Estimator (EAP), but PVs are necessary for obtaining correct standard errors for regression coefficients.

5. The final error variance is computed by combining the sampling variance and the measurement-error variance:

   (error variance) = (sampling variance) + 1.2 (measurement-error variance)

6. The final standard error is the square root of the error variance.

References

Allen, C.S., Q. Chen, et al. (2009), "Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-analytic, Multilevel Analysis", *Educational Evaluation and Policy Analysis* 31(4), pp. 480-499.

Bedard, K. and Dhuey, E. (2006), "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects", *Quarterly Journal of Economics*, 2006, 121(4), pp. 1437-1472.

Eide, E. and M.H. Showalter (2001), "The Effect of Grade Retention on Educational and Labor Market Outcomes", *Economics of Education Review*, Vol. 20, Issue 6.

García Jaramillo, S., C. Fernández Monsalve and F. Sánchez Torres (2011), "Deserción y Repetición en los primeros grados de la básica primaria: factores de riesgo y alternativas de política pública".

Goos, M., B.M. Schreier, et al. (2013), "How Can Cross-Country Differences in the Practice of Grade Retention Be Explained? A Closer Look at National Educational Policy Factors", *Comparative Education Review* 57(1), pp. 54-84.

Jacob, B.A. and L. Lefgren (2004), "Remedial Education and Student Achievement: A Regression-discontinuity Analysis", *Review of Economics and Statistics* 86(1), pp. 226-244.

Jacob, B.A. and L. Lefgren (2009), "The Effect of Grade Retention on High School Completion", *American Economic Journal: Applied Economics*, Vol. 1, No. 3, pp. 33-58.

Lee, J.W. and R.J. Barro (2001), "Schooling Quality in a Cross–section of Countries", *Economica* 68(272), pp. 465-488.

Manacorda, M. (2012), "The Cost of Grade Retention", *Review of Economics and Statistics* 94(2), pp. 596-606.

Murnane, R.J. and J.B. Willett (2011), *Methods Matter: Improving Causal Inference in Educational and Social Science Research*, Oxford Univ. Press.

OECD (2009), *PISA Data Analysis Manual*, OECD Publishing.

OECD (2010), *PISA 2009 Results: What Makes a School Successful?*, Vol. IV, OECD Publishing.

OECD (2011), "When Students Repeat Grades or are Transferred out of School: What Does it Mean for Education Systems?", PISA in Focus, educational policy brief.

OECD (2012). *PISA 2009 Technical Report*, OECD Publishing.

Rumberger, R.W. and Lim, S.A. (2008), "Why Students Drop Out of School: A Review of 25 Years of Research", Santa Barbara, CA, California Dropout Research Project, retrieved 14 October 2009, from *http://cdrp.ucsb.edu/dropouts/pubs_reports.htm#15*.

Schwerdt, G. and M. West (2012), "Effects of Early Grade Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida", mimeo.

Thompson, C. and E. Cunningham (2000), "Retention and Social Promotion: Research and Implications for Policy", *ERIC Digest* No. 161, New York, ERIC Clearinghouse on Urban Education, Teachers College, Columbia University.

West, M.R. (2012), "Is Retaining Students in the Early Grades Self-Defeating?", CCF Brief# 49, *Center on Children and Families at Brookings*.

Xia, N. and S.N. Kirby (2009), "Retaining Students in Grade: A Literature Review of the Effects of Retention on Students' Academic and Nonacademic Outcomes", RAND Technical report. *www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR678.pdf*.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments 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, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union 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.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(13 2013 01 1 P) ISSN 1995-2848 – 2014

OECD Journal

Economic Studies

Reconciling fiscal consolidation with growth and equity
Boris Cournède, Antoine Goujard and *Álvaro Pina*

Choosing the pace of fiscal consolidation
Łukasz Rawdanowicz

Policy incentives for private innovation and maximising the returns
Ben Westmore

Technological effects of intra-OECD trade in manufacturing: A panel data analysis over the period 1988-2008
Javad Abedini

Demographic or labour market trends: What determines the distribution of household earnings in OECD countries?
Wen-Hao Chen, Michael Förster and *Ana Llena-Nozal*

Towards global carbon pricing: Direct and indirect linking of carbon markets
Rob B. Dellink, Stéphanie Jamet, Jean Château and *Romain Duval*

Using a quasi-natural experiment to identify the effects of birth-related leave policies on subjective well-being in Europe
Anna Cristina d'Addio, Simon Chapple, Andreas Hoherz and *Bert Van Landeghem*

Grade repetition: A comparative study of academic and non-academic consequences
Miyako Ikeda and *Emma García*

Consult this publication on line at ***http://dx.doi.org/10.1787/eco_studies-v2013-1-en***

This work is published on the OECD iLibrary, which gathers all OECD books, periodicals and statistical databases.
Visit ***www.oecd-ilibrary.org*** for more information.

Volume 2013

OECD publishing
www.oecd.org/publishing

ISSN 1995-2848
VOLUME 2013 (1 ISSUE)
13 2013 01 1 P

9 771995 284003