

OECD Economic Studies

ECONOMICS

OECD 

No. 33, 2001/2

© OECD, 2001.

© Software: 1987-1996, Acrobat is a trademark of ADOBE.

All rights reserved. OECD grants you the right to use one copy of this Program for your personal use only. Unauthorised reproduction, lending, hiring, transmission or distribution of any data or software is prohibited. You must treat the Program and associated materials and any elements thereof like any other copyrighted material.

All requests should be made to:

Head of Publications Service,
OECD Publications Service,
2, rue André-Pascal,
75775 Paris Cedex 16, France.

No. 33, 2001/II

OECD Economic Studies



ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Pursuant to Article 1 of the Convention signed in Paris on 14th December 1960, and which came into force on 30th September 1961, the Organisation for Economic Co-operation and Development (OECD) shall promote policies designed:

- to achieve the highest sustainable economic growth and employment and a rising standard of living in Member countries, while maintaining financial stability, and thus to contribute to the development of the world economy;
- to contribute to sound economic expansion in Member as well as non-member countries in the process of economic development; and
- to contribute to the expansion of world trade on a multilateral, non-discriminatory basis in accordance with international obligations.

The original Member countries of the OECD are Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The following countries became Members subsequently through accession at the dates indicated hereafter: Japan (28th April 1964), Finland (28th January 1969), Australia (7th June 1971), New Zealand (29th May 1973), Mexico (18th May 1994), the Czech Republic (21st December 1995), Hungary (7th May 1996), Poland (22nd November 1996), Korea (12th December 1996) and the Slovak Republic (14th December 2000). The Commission of the European Communities takes part in the work of the OECD (Article 13 of the OECD Convention).

Publié en français sous le titre :
Revue économique de l'OCDE
N° 33, 2001/2

© OECD 2001

Permission to reproduce a portion of this work for non-commercial purposes or classroom use should be obtained through the Centre français d'exploitation du droit de copie (CFC), 20, rue des Grands-Augustins, 75006 Paris, France, tel. (33-1) 44 07 47 70, fax (33-1) 46 34 67 19, for every country except the United States. In the United States permission should be obtained through the Copyright Clearance Center, Customer Service, (508)750-8400, 222 Rosewood Drive, Danvers, MA 01923 USA, or CCC Online: www.copyright.com. All other applications for permission to reproduce or translate all or part of this book should be made to OECD Publications, 2, rue André-Pascal, 75775 Paris Cedex 16, France.

FOREWORD

OECD *Economic Studies* is produced twice yearly. Many of the articles derive from papers produced for the Organisation's intergovernmental Committees, in particular those serviced by the Economics Department of the OECD Secretariat. Nonetheless, articles also drawn from work by other parts of the Secretariat, or by consultants, may be included. Most of the articles are in the area of applied economic and statistical analysis, generally with an international or cross-country dimension.

OECD *Economic Studies* is published on the responsibility of the Secretary-General of the OECD. Views expressed are those of the authors, and do not necessarily reflect those of the Organisation or its Member governments.

Editorial Board

Jørgen Elmeskov (Chairman)	Paul Atkinson
Sveinbjörn Blöndal	Mark Pearson
Martine Durand	Dirk Pilat
Deborah Roseveare	Hannes Suppanz
Joaquim Oliveira Martins	Paul Schreyer
Paul Swaim	

TABLE OF CONTENTS

THE DRIVING FORCES OF ECONOMIC GROWTH: PANEL DATA EVIDENCE FOR THE OECD COUNTRIES	9
--	----------

Andrea Bassanini and Stefano Scarpetta

This paper discusses links between policy settings, institutions and economic growth in OECD countries on the basis of pooled cross-country time-series regressions. The novel econometric approach used in the paper allows short-term adjustments and convergence speeds to vary across countries, in accordance with most theoretical models, while imposing restrictions only on the long-run coefficients. In addition to the "primary" influences of physical and human capital accumulation, the results confirm the importance for growth of R&D activity, the macroeconomic environment, trade openness and well developed financial markets. They also confirm that many of the policy influences operate not only via the overall efficiency of factor use but also indirectly via the mobilisation of resources for fixed investment.

GROWTH EFFECTS OF EDUCATION AND SOCIAL CAPITAL IN THE OECD COUNTRIES	57
---	-----------

Jonathan Temple

This paper surveys the empirical literature on the growth effects of education and social capital. The main focus is on the cross-country evidence for the OECD countries, but the paper also briefly reviews evidence from labour economics, to clarify where empirical work on education using macro data may be relatively useful. It is argued that on balance, the recent cross-country evidence points to productivity benefits of education that are at least as large as those identified by labour economists. The paper also discusses the implications of this finding. Finally, the paper reviews the emerging literature on the benefits of social capital. Since this literature is still in its early days, policy conclusions are accordingly harder to find.

R&D AND PRODUCTIVITY GROWTH: PANEL DATA ANALYSIS OF 16 OECD COUNTRIES 103

Dominique Guellec and Bruno van Pottelsberghe de la Potterie

This study investigates the long-term effects of various types of R&D on multi-factor productivity growth, which are the spillover effects of R&D activities. Econometric estimates are conducted on a panel of 16 OECD countries, over the period 1980-98. All results are averages over countries and time, and little can be said about country specificities. Major results are as follows: an increase of 1 per cent in business R&D generates 0.13 per cent in productivity growth. The effect is larger in countries that are intensive in business R&D, and in countries where the share of defence-related government funding is lower; a 1 per cent increase in foreign R&D generates 0.46 per cent in productivity growth, and the effect is larger in countries intensive in business R&D; 1 per cent more in public R&D generates 0.17 per cent in productivity growth. The effect is larger in countries where the share of universities (as opposed to government labs) is higher, in countries where the share of defence R&D is lower, and in countries which are intensive in business R&D.

MEASURING PRODUCTIVITY 127

Paul Schreyer and Dirk Pilat

This paper discusses the main measurement issues in calculating productivity indicators, and provides guidance to researchers and statisticians in addressing these difficulties. It draws on the OECD Productivity Manual and on recent OECD work on productivity levels. The paper examines a range of issues related to the measurement of productivity growth, including the choice of output measure (gross output versus value added), the measurement of output, labour and capital input, as well as index number issues. It also discusses OECD estimates of productivity levels and the key measurement issues in deriving these estimates, including the appropriate conversion from one currency unit to another. A final section discusses the interpretation of productivity measures, including their most common applications and the possible pitfalls. The paper concludes that substantial progress has been made in recent years to improve the comparability of productivity statistics. In many countries, however, basic source data are still the key limitation to the development of comparable indicators of productivity. In addition, statisticians, researchers and policy makers need to be more aware of the appropriate uses and interpretation of productivity statistics.

**ESTIMATING THE STRUCTURAL RATE OF UNEMPLOYMENT
FOR THE OECD COUNTRIES 171**

Dave Turner, Laurence Boone, Claude Giorno, Mara Meacci, Dave Rae and Pete Richardson

The paper first reviews the conceptual framework underlying different measures of structural unemployment as well as alternative empirical methods that have been used to provide estimates of them. Drawing on this review, it goes on to develop a method for estimating time-varying NAIRUs across a range of OECD countries using a Kalman filter. It then discusses the resulting econometric estimates, and the scope for their further refinement given the associated range of uncertainties. Recent trends in the NAIRU estimates are reviewed: they fell in many countries in the second half of the 1990s, although actual unemployment has remained well above the NAIRU for a majority of countries throughout much of the 1990s, particularly in Europe. Finally, the relevance of such measures to analysing inflation developments and monetary policy is discussed.

THE DRIVING FORCES OF ECONOMIC GROWTH: PANEL DATA EVIDENCE FOR THE OECD COUNTRIES

Andrea Bassanini and Stefano Scarpetta

TABLE OF CONTENTS

Introduction	10
The determinants of economic growth.....	11
Basic determinants of growth: the accumulation of physical and human capital.....	13
Research and development.....	15
Macroeconomic policy setting and growth	16
Financial development and growth	20
International trade and growth.....	20
Specification of the growth equation and estimation technique	21
The estimated growth equations	21
The econometric technique.....	23
Regression results and interpretation.....	24
The role of convergence and capital accumulation in the growth process.....	26
The role of macroeconomic policy and institutions on growth	27
Research and development.....	30
Policy and institutional influences on capital accumulation	32
Some quantitative implications of the regression results	34
The long-run effect of policy and institutional changes	34
Explaining cross-country differences in average growth rates	36
The role of policy and institutions in shaping the growth process over the past two decades	39
Concluding remarks	39
<i>Appendix 1. The Policy-and-Institutions Augmented Growth Model</i>	47
<i>Appendix 2. The data</i>	50
Bibliography	53

The authors wish to thank Phillip Hemmings for his help in comparing the present results with those in the empirical literature, and Catherine Chapuis-Grabiner for excellent statistical assistance. Sanghoon Ahn, Sven Blöndal, Gavin Cameron, Jørgen Elmeskov, Michael Feiner, Dirk Pilat, Ron Smith, Jonathan Temple, Nicholas Vanston and Ignazio Visco also provided many helpful comments. A previous version of this paper was presented at the "Séminaire d'Economie Monétaire Internationale", Paris, 24 November 2000 (CEPREMAP, DELTA) and at the 16th Annual Congress of the European Economic Association, 29 August to 1 September 2001, and we acknowledge useful comments from participants. The opinions expressed in the paper are those of the authors and do not necessarily reflect those of the OECD or its Member countries.

INTRODUCTION

The past decade has witnessed a renewed interest in the main factors driving economic growth in the OECD countries. A few countries – including the United States, the technology leader – have experienced an acceleration in growth of GDP per capita, but other major economies have lagged behind, raising questions as to the role of technological progress as well as policy and institutions. This paper aims at shedding some light on these issues by presenting evidence on the long-term links between policy settings, institutions and economic growth in OECD countries while controlling for underlying differences in technological progress. In particular, the focus is two-fold: first, on the possible influences of human capital, research and development activity, macroeconomic and structural policy settings, trade policy and financial market conditions on economic efficiency; second, on the effects of many of the same factors on the accumulation of physical capital.

While empirical studies support the relevance of these factors for economic growth, this literature often relies on a large set of countries, including many non-OECD economies and, once constrained to the OECD sample, results are often unsatisfactory (Temple, 1999). The cross-country variability in both growth patterns and potential explanatory variables is much smaller if one focuses on the OECD sub-sample. Hence, data quality and the estimation approach assume an even more crucial role in the empirical analysis. We tackle both issues by using harmonised OECD data and a novel econometric approach that reconciles growth model assumptions with available data.

As in a number of recent studies, we use pooled cross-country time-series data in order to explain both cross-country differences in growth performance as well as the evolution of performance over time in each country. In addition, our econometric technique allows short-term adjustments and convergence speeds to vary across countries while imposing (and testing) restrictions only on long-run coefficients (*i.e.* those related to the production function). To anticipate the main results of this study, we find the accumulation of physical as well as human capital to be the main drivers of economic growth. In addition, R&D activity, a sound macroeconomic environment, trade openness and well-developed financial markets contribute to raise living standard in OECD countries. Some of the same factors

that operate “directly” on growth also influence it indirectly *via* the mobilisation of resources for fixed investment.

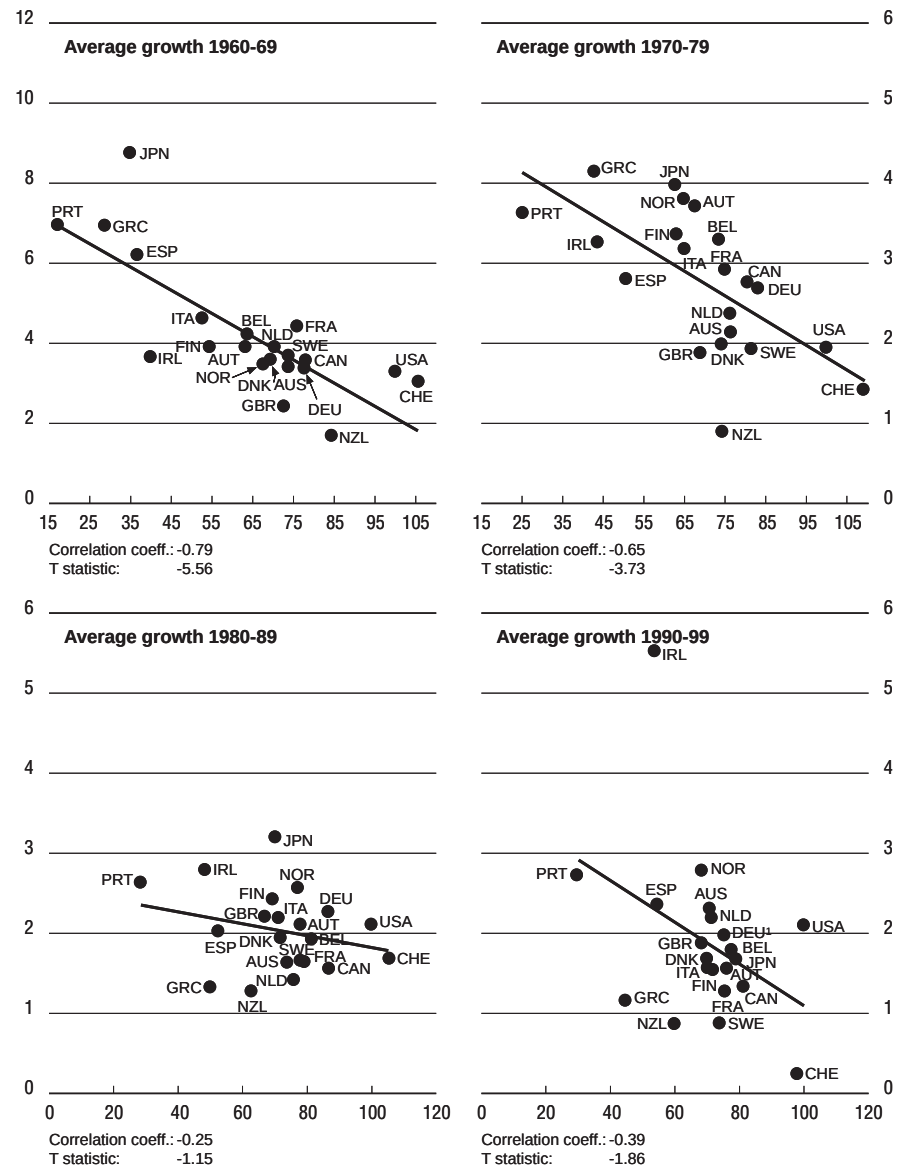
The paper is organised as follows. In the first section, we briefly introduce the policies and institutional dimensions that are considered in the empirical investigation of the sources of economic growth. The section focuses on the transmission mechanisms linking policy to growth as well as on cross-country differences in policy settings and their evolution over time. The institutional and policy variables considered have three basic characteristics: *i*) they are largely macroeconomic in nature; *ii*) they yield testable implications for economic growth; and *iii*) they can be evaluated using available data across countries and over time. In the second section we introduce the estimated growth equation and discuss the econometric approach used in the regressions. The third section presents the econometric results. The estimated coefficients from the growth regressions are also used in the fourth section to assess the role of different policy settings in the evolution of growth both over time and across the OECD countries. The final section concludes.

THE DETERMINANTS OF ECONOMIC GROWTH

The literature on economic growth is vast and policy-oriented studies, in particular, have flourished in the past decade (see Temple, 1999 and Ahn and Hemmings, 2000 for surveys). Yet, there is little agreement on the exact mechanisms linking policy settings to growth. For example, if one assumes, consistent with the traditional neo-classical growth model, diminishing returns to reproducible factors and exogenous saving rates, population growth and technological progress, then policies have no role in shaping long-term economic growth. Under these circumstances, richer countries grow at a slower rate than poorer countries adjusted for demographic differences. However, evidence of this process of *unconditional* convergence has weakened, at least amongst the OECD countries, in the most recent decades (Figure 1).¹ Thus, the concept of convergence can only be reconciled with the data if one moves to *conditional* convergence; that is to say, a relation between growth rate and initial conditions after controlling for other variables.

Relaxing the hypothesis of exogenous saving and capital formation gives room for policy to affect growth in the short and medium-term via an impact on saving and the level and composition of investment. Indeed, a number of studies suggest that policy and institutions affect the *level* of economic efficiency with which resources are allocated in the economy. Nevertheless, whether through its effect on investment or on the level of economic efficiency, a one-time change in policy leads only to a transitory change in output growth in such models. When the capital stock and output have risen to levels at which the new rate of gross

Figure 1. Comparison of GDP per capita growth rates and initial conditions over four decades



Note: GDP per capita is based on the OECD 1993 purchasing power parities estimates, average growth rates are based on H-P cyclically-adjusted real GDP volumes. Horizontal axis show the PPP level of GDP per capita (USA = 100).

1. 1991-1999 period.

Source: OECD.

investment is only sufficient to maintain a constant capital/labour ratio plus an amount to cover physical depreciation, growth reverts back to the steady state rate. In other words, any policy change will affect output growth only in the short to medium-term by shifting the growth path, although the *underlying rate of growth* remains determined by exogenous population growth and technical progress (that are potentially different across countries).

Another class of growth models relaxes the assumption of diminishing returns to reproducible factors. Some authors add human to physical capital to derive a concept of “broad” capital characterised by constant or even increasing returns to scale (*e.g.* Lucas, 1988; Rebelo, 1991). Others introduce externalities to the accumulation of physical capital whereby private returns to scale may be diminishing, but social returns can be constant or increasing – due to either learning by doing (*e.g.* Romer, 1986; Young, 1991) or R&D (*e.g.* Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992). With constant (or increasing) returns to (“broad”) capital, the long-term rate of growth becomes endogenous, in the sense that it depends on investment decisions which, in turn, could be influenced by policy and institutions. Some of these endogenous growth models imply “conditional” convergence, while others do not, depending on assumptions about the specification of the production function and the evolution of broad capital accumulation (see Barro and Sala-i-Martin, 1995; Durlauf and Quah, 1999 for reviews).

Only empirical observation can provide evidence on which view of the link between policy and growth is most relevant, but results of such studies are often ambiguous. Aggregate analyses, such as the one we present in this paper, can only shed some light, while microeconomic evidence is needed to better assess the links between capital accumulation and technological progress. Bearing this in mind, we look at the influence of policy and institutions on GDP per capita by estimating a growth equation and an investment equation. The GDP per capita growth equation aims at identifying the effect on output of a policy variable over and above its potential impact on investment, while the investment equation is intended to identify the possible impact of the policy variable on the level of investment. The remainder of this section presents the different variables considered in our analysis.

Basic determinants of growth: the accumulation of physical and human capital

The accumulation of physical capital

The rate of accumulation of physical capital is one of the main factors determining the level of real output per capita although, as stressed above, its effects could be more or less permanent depending on the extent to which technological

innovation is embodied in new capital. Whatever the transition mechanism from capital accumulation to growth, the significant differences in the investment rate across countries and over time point to it as a possible source of cross-country differences in output per capita. In particular, long-run averages of business-sector investment rates range from around 10 per cent to over 20 per cent of GDP. Furthermore, major shifts in investment rates within countries are not uncommon, a notable example being the rapid rise in the US investment rate in recent years.

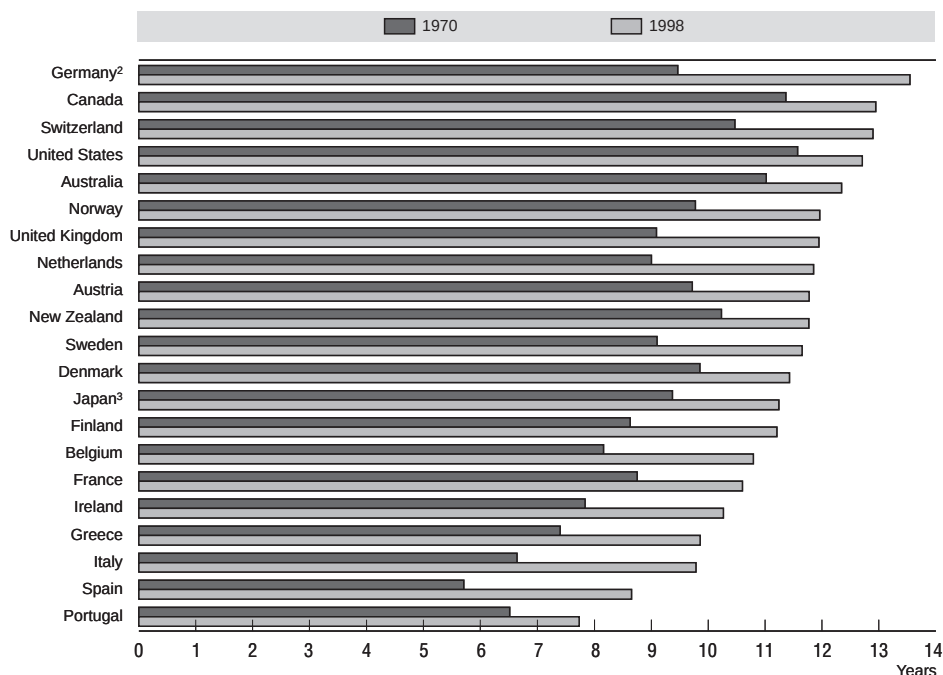
In the empirical analysis, we consider the accumulation of physical capital by private agents (proxied by the share of business investment in GDP). Public-sector investment is also considered in extended growth equations to assess its independent impact on output, as suggested by Aschauer (1989), as well as its potential effect on the estimated coefficient of the business-sector investment rate.

Human capital

Recent studies on growth also assume that formal skills and experience embodied in the labour force represent a form of (human) capital. On the one hand, it could be argued that human capital is subject to some kind of diminishing returns so that a more highly trained and skilled workforce would enjoy higher levels of income in the long term, but not necessarily permanently higher growth rates of income. On the other hand, investment in human capital (*e.g.* expenditures on education and training) could have a more permanent impact on the growth process if high skills and training go hand-in-hand with more intensive research and development and a faster rate of technological progress, or if the adoption of new technologies is facilitated by a highly skilled workforce.² As in the case of physical capital, only empirical evidence can shed some light on social returns to investment in human capital and, thus, help in discriminating amongst competing theories.

In this study, human capital is measured by estimates of the average number of years of formal education among the working-age population, based on figures on educational attainment and assumptions about how many years of education a particular level of attainment represents. These indicators are admittedly only crude and somewhat narrow proxies, taking little account of quality aspects of formal education or other important dimensions of human capital. Nonetheless, estimates of the average years of schooling amongst the working-age population suggest that, despite some convergence over time, there remain significant differences between countries as to the level of educational attainment. In 1970, the average length of formal education of the working age population ranged from 5.7 years (Spain) to 11.6 years (United States), whilst the most recent observations still indicate a range from 7.7 to 13.6 years (Portugal and Germany, respectively) (Figure 2). The figure also indicates that increases in average length of education

Figure 2. **Average years of education¹ of the working-age population in OECD countries**
1970 and 1998



1. Average number of years of education in the population aged 15-64 years, based on data on level of education attained and assumptions about the number of years of education implied by different levels of education attainment.
2. Western Germany in 1970.
3. 1990 instead of 1998 for Japan.

Source: OECD.

range from less than half a year on average per decade (*e.g.* the United States) to more than one year on average per decade (*e.g.* Germany and Italy, the latter from a relatively low level).

Research and development

Expenditure on research and development (R&D) can be considered as an investment in knowledge that translates into new technologies as well as more efficient ways of using existing resources of physical and human capital. Indeed, in the case of R&D, there seems to be stronger consensus that it may have a persis-

tent effect on growth, that is, higher R&D expenditure would, *ceteris paribus*, be associated with permanently higher growth rates.

The amount of resources that are devoted to R&D can be influenced by government intervention. In particular, the potential benefits from new ideas may not be fully appropriated by the innovators themselves due to spillover effects, which imply that without policy intervention the private sector would likely engage in less R&D than what could be socially optimal. This can justify some government involvement in R&D, both through direct provision and funding, but also through indirect measures such as tax incentives and protection of intellectual property rights to encourage private-sector R&D (see Nadiri, 1993 and Cameron, 1998, for reviews).

Overall expenditure on R&D as a share of GDP has risen since the 1980s in most countries (Figure 3), mainly reflecting increases in R&D activity in the business sector that accounts for the majority of expenditure in this area in most OECD countries. Indeed, the share of publicly financed R&D has declined over the past decade in most countries as a result of reductions in military R&D budgets.

Several issues have to be considered in assessing the role of R&D on growth. First, the relationship between public and private R&D could be one of complementarity or substitution. Second, public-sector R&D is often directed at making improvements in areas not directly related to growth, such as defence and medical research, and any impact on output growth could be diffused and slow to come about (see OECD, 1998). These considerations suggest that any quantitative analysis of growth must take R&D activity into account as an additional form of investment and differentiate between various types of R&D expenditures. Given data availability, we consider total R&D expenditure (as a share of GDP) and its components, public and business sector R&D expenditure.

Macroeconomic policy setting and growth

In the context of growth studies, three issues have generally been considered with respect to macroeconomic policy settings: the benefits of establishing and maintaining low inflation, the impact of government deficits on private investment, and the possibility of negative impacts on growth stemming from a too-large government sector (with associated high tax pressure to finance high government expenditure).

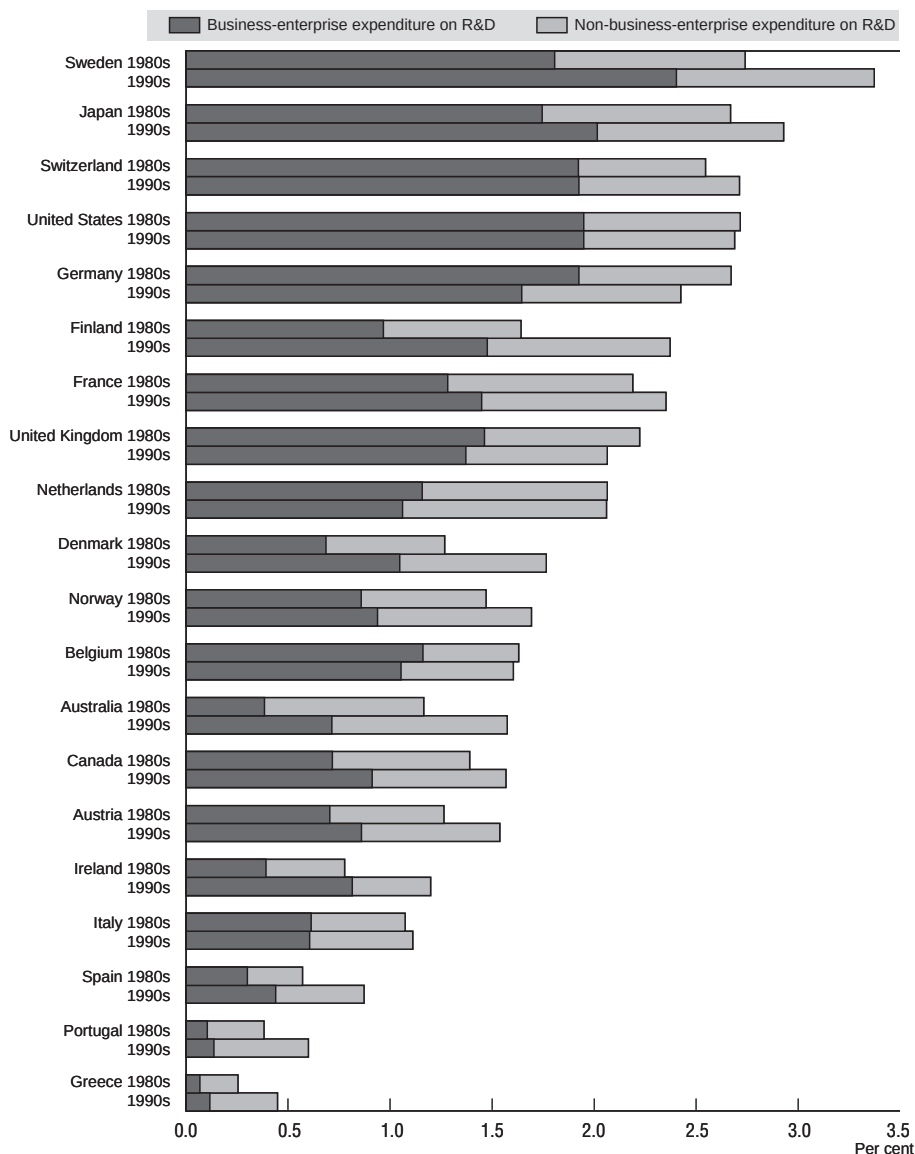
Inflation and growth

The usual arguments for lower and more stable inflation rates include reduced uncertainty in the economy and enhanced efficiency of the price mechanism. A reduction in the level of inflation could have an overall effect on the level of capital accumulation in cases of tax distortions (*e.g.* nominally-denominated

Figure 3. Expenditure on R&D in the OECD countries

1980s and 1990s

Total expenditure on R&D as a percentage of GDP



Source: OECD.

allowances) or when investment decisions are made with a long-run perspective (*e.g.* shift in technologies). Moreover, uncertainty related to higher volatility in inflation could discourage firms from investing in projects that have high returns, but also a higher inherent degree of risk.³

Evidence on the relationship between inflation and growth is somewhat mixed: while there is evidence that investment suffers in cases of high inflation, the relation is less clear in cases of moderate or low inflation (see *e.g.* Edey, 1994; Bruno and Easterly, 1998). Moreover, to the extent uncertainty is the link to investment and growth, it would suggest a focus on variation in inflation. However, given the correlation between level and variability of inflation, the two effects could be difficult to distinguish.

Based on the above discussion, two indicators of inflation are considered in the empirical analysis: the level of inflation and its variability. These indicators are included in the growth equation, which incorporates the investment share, whereby the estimated impact on growth occurs *via* the effects of these variables on overall efficiency and the choice of investment projects. They are also included in the investment equation, which permits testing for an effect of both variables on the level of investment.

Fiscal policy and growth

Fiscal policy settings can affect output and growth in the medium term as well as over the business cycle. In particular, where government deficits finance consumption or transfers, a traditional argument for prudent policy is to reduce the crowding out effects on private sector investment. Also, if fiscal policy is seen as being at odds with monetary policy, the credibility of the latter could be undermined leading to risk premia in interest rates and pressures on exchange rates, with repercussions on capital accumulation.

It has been argued that taxes necessary to finance government spending could also distort incentives, with negative implications for the efficient allocation of resources and hence the level or the growth of output. The main conclusion from the literature is that there may be both a “size” effect of government intervention as well as specific effects stemming from the financing and composition of public expenditure. At a low level, the productive effects of public spending are likely to exceed the social costs of raising funds. However, government expenditure and the required taxes may reach levels where the negative effects on efficiency and hence growth starts dominating. These negative effects may be more evident where the financing relies heavily on more “distortionary” taxes (*e.g.* direct taxes) and where public expenditure focuses on “unproductive” activities.

Between the 1980s and 1990s the “size” of the public sector tended to increase in most OECD countries as did government gross liabilities. More

recently, public-sector balances have improved significantly. Notwithstanding these important developments, the share of total government expenditure in GDP was still in the range of 40-50 per cent in a number of OECD countries in 1999 (Table 1). Moreover, less than 20 per cent of public expenditure in OECD countries consists of expenditure that could be classified as directly "productive" (*e.g.* schooling, infrastructure and R&D). And in a number of countries, the share of "productive" expenditure declined over the past decade.

In our empirical analysis of both growth and investment, supply-side hypotheses relating government size to growth are tested by looking at both taxation and

Table 1. **Total government outlays and "productive" government spending as a share of total spending in OECD countries, 1985, 1995 and 1999**

Percentage

	A Education		B Transport and communication		C R&D		A + B + C		Share of total government outlays in GDP		
	1985	1995	1985	1995	1985	1995 ⁴	1985	1995	1985	1995	1999
Australia	14.6	13.2	10.1	8.3	2.1 ⁵	2.2 ⁴	26.8	23.6	37.3	35.5	32.3
Austria	9.6	9.5	3.3	2.1	1.2	1.4	14.1	13.0	50.1	52.6	50.7
Belgium	12.7	..	8.7	..	0.9	..	22.3	..	57.3	50.1	47.9
Canada	13.0	..	5.4	..	1.5	..	19.8	..	46.0	46.3	40.2
Denmark	11.3	11.7	4.0	3.0	1.2	1.2	16.4	15.9	56.6 ³	59.0	54.3
France ¹	10.5	10.7	2.9	1.9	2.3	1.8	15.7	14.4	51.8	53.5	52.2
Germany	9.5	7.6	4.3	3.4	2.2	1.8	16.0	12.9	45.6	46.3	45.6
Iceland	13.0	12.3	9.0	7.6	1.6	2.5	23.6	22.4	32.5 ⁷	36.0	33.4
Ireland ¹	10.6	12.2	4.5	5.0	0.8	0.8	15.9	18.0	46.6	36.4	31.5
Italy	10.0	8.9	7.7	4.6	1.2	1.0	18.8	14.5	50.6	52.3	48.3
Japan	12.8	10.8 ⁴	..	..	1.8	1.9	..	..	31.6	35.6	38.1
Korea	17.8	18.1	7.1	9.6	..	2.7	..	30.4	17.6	19.1	25.5
Netherlands	9.9	..	..	..	1.8	..	..	..	51.9	47.7	43.2
New Zealand	..	13.3 ⁴	..	..	..	1.3 ¹	..	..	52.6 ⁶	38.8	40.8
Norway	12.0 ³	13.7	6.6 ³	5.9	1.6	1.6	20.2	21.3	41.5	47.6	46.1
Portugal ²	8.7	13.3	3.6	4.8	0.5 ⁵	0.9	12.9	19.0	40.2	44.5	44.7
Spain	8.8	10.3	6.3	6.0	0.7	0.9	15.8	17.1	37.7	42.5	38.6
Sweden	..	..	..	..	1.7	1.7	..	..	60.3	62.4	55.9
Switzerland	19.7	..	11.4	..	..	..	..	..	..	..	..
United Kingdom	10.2	12.1	3.2	3.6	2.0	1.5	15.5	17.2	43.0 ⁷	44.4	39.9
United States	..	..	..	..	4.1	2.8	..	..	33.8	32.9	30.1

1. 1993 instead of 1995.

2. 1992.

3. 1988.

4. 1994.

5. 1984.

6. 1986.

7. 1987.

The concept of "productive" government spending is based on a taxonomy used by Barro (1991).

Source: OECD.

government spending. The potential role of the structure of financing and expenditure is considered by looking separately at direct and indirect taxes and different components of government expenditure. In this context, the human capital variable discussed above may be taken to represent past and present governments' efforts in financing education.⁴ Likewise, public spending on research and development are clearly identified in the extended models that include R&D (see below). Finally, public investment has been included as distinct from consumption to test if that distinction is pertinent.

Financial development and growth

Financial systems also contribute to economic growth by providing funding for capital accumulation and by helping the diffusion of new technologies. A well-developed financial system is likely to mobilise savings by channelling small savings of individuals into profitable large-scale investments, while offering savers a high degree of liquidity. It also provides insurance to individual savers against idiosyncratic risk through diversification, and it reduces the costs of acquiring and evaluating information on prospective projects, for example through specialised financial advisory services. All these services are likely to contribute to economic growth, but there could, in theory, also be opposite effects. For example, lower risk and higher returns resulting from diversification may prompt households to save less, if income effects dominate substitution effects. Furthermore, to some extent growth may prompt the development of financial systems, thus there may be an element of reverse causality.

Ideally, we would like to use qualitative indicators of the possibilities offered to firms to access external funds and the ease with which investors can get adequate returns. However, available information is only limited to quantity indicators (Leahy *et al.*, 2001). In particular, we consider the total claims of deposit money banks on the private sector, which measures the degree of financial intermediation via the banking system. Moreover, we look at stock market capitalisation (the value of listed shares), which is an imperfect indicator of the ease with which funds can be raised in the equity market.

International trade and growth

Aside from the benefits of exploiting comparative advantages, theories have suggested additional gains from trade arising through economies of scale, exposure to competition and the diffusion of knowledge. These could result both in higher overall efficiency and possibly a higher level of investment (*e.g.* if the adoption of foreign technologies requires investment in new types of capital). The progress OECD countries have made in reducing tariff barriers and dismantling non-tariff barriers would therefore suggest positive gains from trade.⁵ However,

trade may also be endogenous to the process of growth. The relatively open stance towards trade in OECD countries would suggest that the amount of trade conducted reflects patterns of growth (and to some extent geography, size and transport costs) as much as it reflects constraints in the form of tariff and non-tariff barriers.⁶

The possible reverse causality problem in the relationship between trade and economic growth suggests some caution in interpreting empirical results. In particular, we treat the intensity of trade in the growth equation as an indicator of trade exposure – capturing features such as competitive pressures – rather than one with direct policy implications. Apart from bearing this caveat in mind, the empirical analysis also has to take into account that small countries are naturally more exposed to foreign trade, regardless of their trade policy or competitiveness, while competitive pressure in large countries to a large extent stems from domestic competitors. To better reflect overall competitive pressures, the indicator of trade exposure was adjusted for country size by regressing the crude trade exposure variable on population size and taking the estimated residuals from this exercise as the (adjusted) trade variable in the analysis.

SPECIFICATION OF THE GROWTH EQUATION AND ESTIMATION TECHNIQUE

The estimated growth equations

Formally, the policy-augmented growth equation can be derived from a growth model built around a constant-returns-to-scale technology (see Appendix I for the derivation of the model). Output is a function of capital, employment, the efficiency with which they act together, and the level of technology. Given straightforward assumptions on how the factors of production evolve over time, the steady-state level of output per capita can be expressed as a function of the propensity to accumulate physical capital, the population growth rate, the level and growth rates of technological and economic efficiency, and the rate of depreciation of capital. Moreover, if the concept of capital is widened to include human capital, then the propensity to accumulate the latter is also a factor shaping the steady-state path of output per capita.

If countries were at their steady state – or if deviations from the steady state were random – growth equations could be simply based on the relationship linking steady-state output to its determinants. However, actual data may well include out-of-steady-state dynamics due, among other things, to a slow convergence to the steady state (see, amongst others, Mankiw *et al.*, 1992, for a discussion). Hence, the observed growth in output in any given period, abstracting from cyclical fluctuations, can be seen as the combination of three different forces: i) underlying tech-

nological progress – which is assumed to be exogenous; ii) a convergence process towards the country-specific steady-state path of output per capita; and iii) shifts in the steady state (growth or level of GDP per capita, see below) that can arise from changes in policy and institutions as well as investment rates and changes in population growth rates.

The empirical approach adopted in this paper starts with a parsimonious specification of the growth equation and then analyses extended models. The initial specification is consistent with the standard neo-classical growth model and includes only a convergence factor and the basic determinants of the steady state, namely the accumulation of physical capital and population growth. The first extension involves the introduction of human capital while further extensions consider R&D and a set of policy and institutional factors potentially affecting economic efficiency.

The OECD sample permits the use of annual data instead of averages over time, as often done in the cross-country empirical literature.⁷ However, year-to-year variations in output include cyclical components. These have been controlled for by including first differences of the steady-state determinants as short-run regressors in the estimated equations. Considering pooled cross-country time series (i denotes countries, t time) the growth equation, in its more general form, can be written as follows:

$$\begin{aligned} \Delta \ln y_{i,t} = & a_{0,i} - \phi_i \ln y_{i,t-1} + a_{1,i} \ln s k_{i,t} + a_{2,i} \ln h_{i,t} - a_{3,i} n_{i,t} + \sum_{j=4}^m a_{j,i} \ln V_{i,t}^j + a_{m+1,i} t \\ & + b_{1,i} \Delta \ln s k_{i,t} + b_{2,i} \Delta \ln h_{i,t} + b_{3,i} \Delta n_{i,t} + \sum_{j=4}^m b_{j,i} \Delta \ln V_{i,t}^j + \varepsilon_{i,t} \end{aligned} \quad (1)$$

where y is GDP per capita, sk is the propensity to accumulate physical capital; h is human capital; n is population growth; the V^j is a vector of variables affecting economic efficiency; t is a time trend; the b -regressors capture short-term dynamics and ε is the usual error term.

It should be stressed that equation [1] is a fairly general specification, and different growth models are nested in it. This is important for the interpretation of the policy variables which could be taken to represent either growth effects or level effects (see above). The estimated parameters of equation [1] allow to distinguish amongst some of these models. In particular, a significant coefficient on the lagged level of GDP per capita, *i.e.* the existence of convergence towards a country-specific steady state, would exclude one class of endogenous growth models (*i.e.* those *à la* Romer, 1986).⁸ However, this would not be sufficient evidence to rule out other endogenous models (*e.g.* *à la* Lucas, 1988).⁹ Indeed, even in the presence of convergence, a number of empirical papers have interpreted the estimated policy coefficients as persistent *growth* effects.¹⁰

The distinction between temporary or permanent growth effect may seem somewhat semantic if the speed of convergence to the steady state is very slow, as it is the case in most empirical studies focussing on a large set of countries.¹¹ However, consistent with some recent studies focussing on panel data, we find a relatively rapid speed of convergence for the OECD countries and thus, the choice between the two alternative interpretations of the results does matter in drawing policy conclusions. We take a prudent view and interpret the estimated coefficients as indication of temporary effects on growth due to the shift effect on the steady-state path of output per capita.

The econometric technique

The main advantage of pooled cross-country time-series data for the analysis of growth equations is that the country-specific effects can be controlled for, *e.g.* by using a dynamic fixed-effect estimator (DFE). However, this estimator generally imposes homogeneity of all slope coefficients, allowing only the intercepts to vary across countries. The validity of this approach depends critically on the assumption of a common growth rate of technology and a common convergence parameter (see Lee *et al.*, 1997). The first assumption is difficult to reconcile with evidence on multifactor productivity patterns across countries (see *e.g.* Scarpetta *et al.*, 2000). The second assumption is not consistent with the underlying growth model, where the speed of convergence depends, amongst other factors, upon the rate of population growth (see Appendix 1).¹² An alternative approach is to use the mean-group approach (MG) that consists of estimating separate regressions for each country and calculating averages of the country-specific coefficients (*e.g.* Evans, 1997; Lee *et al.*, 1997). While consistent, this estimator is likely to be inefficient in small country samples, where any country outlier could severely influence the averages of the country coefficients (see below).

An intermediate choice between imposing homogeneity on all slope coefficients (DFE) and imposing no restrictions (MG) is the pooled mean group estimator (PMG) that allows intercepts, the convergence parameter (ϕ),¹³ short-run coefficients (b_s) and error variances to differ freely across countries, but imposes homogeneity on long-run coefficients. There are good reasons to believe in common long-run coefficients for the OECD countries, given that they have access to common technologies, and have intensive intra-trade and foreign direct investment, all factors contributing to similar long-run production function parameters. Under the assumption of long-run slope homogeneity, the PMG estimator increases the efficiency of the estimates with respect to mean group estimators (Pesaran *et al.*, 1999). Formally, conditional on the existence of convergence to a steady state path, the long-run homogeneity hypothesis permits

the direct identification of the parameters that affect the steady state path of output per capita ($a_{s,i}/\phi_i = \theta_s$, see below). In other words, with the PMG procedure, the following restricted version of equation [1] is estimated on pooled cross-country time-series data:

$$\Delta \ln y_{i,t} = -\phi_i \left(\ln y_{i,t-1} - \theta_1 \ln sk_{i,t} - \theta_2 \ln h_{i,t} + \theta_3 n_{i,t} - \sum_{j=4}^m \theta_j \ln V_{i,t}^j - a_{m+1} t_i - \theta_{0,i} \right) \\ + b_{1,i} \Delta \ln sk_{i,t} + b_{2,i} \Delta \ln h_{i,t} + b_{3,i} \Delta n_{i,t} + \sum_{j=4}^m b_{j,i} \Delta \ln V_{i,t}^j + \varepsilon_{i,t} \quad (2)$$

Similarly, the investment equation has the general form:

$$\Delta \ln sk_{i,t} = -\rho_i \left(\ln sk_{i,t-1} - \gamma_1 \ln y_{i,t} - \gamma_2 \ln h_{i,t} - \sum_{j=3}^m \gamma_j \ln V_{i,t}^j - \gamma_{0,i} \right) \\ + c_{1,i} \Delta \ln y_{i,t} + c_{2,i} \Delta \ln h_{i,t} + \sum_{j=3}^m c_{j,i} \Delta \ln V_{i,t}^j + \varsigma_{i,t} \quad (3)$$

i.e. the share of business sector investment in GDP is assumed to depend on the level of GDP per capita, human capital and a set of policy and institutional factors.

In both equations, the hypothesis of homogeneity of the long-run policy parameters cannot be assumed *a priori* and is tested empirically in all specifications.¹⁴ Moreover, given the limited degrees of freedom in the country-specific policy-augmented growth regressions, the time trend was not included. Indeed, a sensitivity analysis suggested that a time trend was only statistically significant when human capital is omitted.¹⁵ In addition, a sensitivity analysis suggests the presence in our database of a handful of observations that significantly increase the standard error of the regression and/or affect the estimated coefficients. Thus, these observations have been removed from the sample.¹⁶

REGRESSION RESULTS AND INTERPRETATION

The growth equations were estimated for 21 OECD countries over the period 1971-1998.¹⁷ The countries were chosen because they have continuous annual series for most of the variables used in the growth equations over the bulk of the 1971-98 period. Details on the variables used in the regression are in Box 1. This section presents the core results of the econometric analysis, while a detailed sensitivity analysis is presented in an extended version of this paper (Bassanini *et al.*, 2001).

Box 1. Description of the variables used in the empirical analysis

The baseline equation includes the following explanatory variables (see Appendix 2 for more details):

- *Dependent variable* ($\Delta \ln Y$). Growth in real GDP per head of population aged 15-64 years expressed in (1993) Purchasing Power Parities (PPP).
- *Convergence variable* ($\ln Y_{-1}$). Lagged real GDP per head of population aged 15-64 years, in PPP.
- *Physical capital accumulation* ($\ln Sk$). The propensity to accumulate physical capital is proxied by the ratio of real private non-residential fixed capital formation to real private GDP.
- *Stock of human capital* ($\ln H$) is proxied by the average number of years of schooling of the population from 25 to 64 years of age.
- *Population growth* ($\Delta \ln P$). Growth in population aged 15-64 years.

The auxiliary policy-related variables included in the augmented growth regressions were as follows:

- *Measures of inflation*: i) the rate of growth of the private final consumption deflator ($\ln \pi$); and ii) the standard deviation of the rate of growth in private final consumption deflator ($SD \ln \pi$) – estimated as a centred three-year moving-average.
- *Indicators of government size and financing*: i) the share of general government current nominal tax and non-tax receipts in nominal GDP ($\ln Tax$); ii) the ratio of direct to indirect tax receipts ($\ln(Tax\ dist)$); iii) the ratio of government nominal final consumption expenditure to nominal GDP ($\ln(Gov\ cons)$); and iv) the ratio of government real fixed capital formation to real GDP ($\ln Sk^{gov}$).
- *Measures of R&D intensity*: i) gross domestic expenditure on R&D as a percentage of GDP ($\ln R\&D^{tot}$); ii) business sector expenditure on R&D as a percentage of GDP ($\ln BERD$); and iii) the difference between gross domestic expenditure on R&D and business sector expenditure on R&D as a percentage of GDP ($\ln R\&D^{pub}$).
- *Indicators measuring financial development*: i) private credit of deposit money banks provided to the private sector as a percentage of GDP ($\ln(Priv\ Credit)$); and ii) stock market capitalisation as a percentage of GDP ($\ln(Stock\ Cap)$).
- *Indicators of the exposure of countries to foreign trade*: a weighted average of export intensity and import penetration. It is calculated as follows: $Trade\ Exp = Xi + (1 - Xi) * Mp$, where Xi is the ratio of exports to GDP and Mp is the ratio of imports to apparent consumption (domestic production minus exports plus imports). In the empirical analysis this measure was adjusted for country size as described above ($\ln(Trade\ exp)^{adj}$).

All the auxiliary policy-related variables, with the exception of those related to R&D, have been introduced with a lag to better identify their impact on output.

The role of convergence and capital accumulation in the growth process

Table 2 reports the estimated coefficients and implied parameters for physical and human capital and convergence from different specifications of the growth equation. In all specifications the convergence parameter is significant, suggesting a (conditional) process of convergence and supporting the specification adopted in equation [2]. In the human capital augmented equations, the speed with which countries converge to their specific steady state path of output per capita is, however, higher than what is generally found in growth regressions focussing on larger sets of countries. Estimates reported in Table 2 suggest that, following a change in a steady-state variable, it takes about 4-5 years to go half way to the new steady state output per capita. With rapid convergence, a policy change will have only a temporary impact on growth, but its potential effect on living standards will be quickly felt. Hence, observed changes in factor inputs as well as in policies over past decades are likely to have significantly affected growth patterns and are of importance in the assessment of cross-country differences.

In all specifications where they are included, the coefficients on both physical and human capital appear with the expected sign and are highly significant. There is, however, some variability in the estimated coefficients that implicitly underlines the importance of model specification. The three right-hand-side specifications in Table 2 indicate an implied value for the capital share of about 20 per cent, which is broadly consistent with National Accounts data. The estimated share of human capital is, however, quite large (60 per cent in the preferred specifications).¹⁸ This implies that one extra year of average education (corresponding to a rise in human capital by about 10 per cent) would lead to an average increase in steady-state output per capita by about 4-7 per cent.¹⁹ These values contrast with (many) growth studies that have found no or very limited effects of human capital on growth (see, for example, Benhabib and Spiegel, 1994; Barro and Sala-i-Martin, 1995), although the lowest of the present estimates are broadly consistent with estimated returns to schooling in the microeconomic literature (see Psacharopoulos, 1994).²⁰

Taken at face value, the results on human capital might imply significant positive spillovers and a gap between private and social returns to education, or that the human capital indicator is acting as a proxy for other variables (over and above those included as framework conditions), an issue also raised in some microeconomic studies. The first interpretation of the results potentially has important policy implications. Insofar as policy affects the accumulation of human capital (most prominently through education policy), and the spill-over effects are sufficiently large to imply overall non-declining returns over some range, its effect on growth may not be limited to a shift in the steady-state output level but possibly lead to more persistent (although not necessarily irreversible) effects.

Table 2. **The role of convergence and capital accumulation for growth:**
summary of regression results
(Pooled Mean Group Estimates)

Estimated coefficients	Standard equation ¹	Human-capital augmented	Trade-and-policy-augmented equations		
			A ²	B ³	C ⁴
LnSk	0.39 *** (0.11)	0.18*** (0.04)	0.25*** (0.04)	0.23*** (0.04)	0.24*** (0.04)
LnH	...	1.00*** (0.10)	0.41*** (0.13)	0.70*** (0.16)	0.71*** (0.13)
LnY ₋₁	-0.05*** (0.01)	-0.12*** (0.02)	-0.17*** (0.02)	-0.15*** (0.03)	-0.15*** (0.03)
Partial output elasticities: ⁵					
Physical capital	0.28	0.15	0.20	0.19	0.20
Human capital	...	0.85	0.33	0.57	0.57
Half way to convergence ⁶	13.9 years	5.3 years	3.9 years	4.3 years	4.3 years

All equations include a constant country-specific term and control for outliers. The variables are defined in Box 1. Standard errors are in brackets.

*: significant at 10 % level; ** at 5% level; *** at 1 % level.

1. The standard equation includes investment share in physical capital, population growth and lagged output per capita.

2. The equation also includes trade exposure, inflation and standard deviation of inflation.

3. The equation also includes trade exposure, standard deviation of inflation and tax and non-tax receipts.

4. The equation also includes trade exposure, standard deviation of inflation and government consumption.

5. Implied parameters of the production function .

6. Time to cover half way to convergence as implied by the estimated average coefficient of $\ln Y_{-1}$.

The estimated coefficients on broad capital (physical and human) and the speed of convergence, taken together, are not consistent with the standard neo-classical growth model. In particular, a high output elasticity to broad capital should be associated with a very low speed of convergence to the steady state or *vice versa*.²¹ The fact that the output elasticity to broad capital is close to unity suggests that the OECD data are more consistent with an endogenous growth model that allows for conditional convergence. In particular, as discussed in Bassanini and Scarpetta (2002), an endogenous growth model *à la* Uzawa-Lucas with constant returns to scale to “broad” (human and physical) capital can be observationally equivalent to a more standard neo-classical model with diminishing returns to broad capital.

The role of macroeconomic policy and institutions on growth

Tables 3 and 4 report the estimation results relating to the role of variables reflecting macro policy, trade exposure and financial development. Where appropriate, a version of the regression allowing the long-run coefficient of the addi-

Table 3. **Macro policy influences on growth¹**
(Pooled Mean Group Estimators)

Dependent variable: $\Delta \ln Y$	With control for inflation variables	With control for taxes and government expenditures		With control for both inflation and fiscal policy		
Long-Run Coefficients						
LnSk	0.25*** (0.04)	0.36*** (0.04)	0.14*** (0.04)	0.29*** (0.05)	0.23*** (0.04)	0.24*** (0.04)
LnH	0.41*** (0.13)	1.26*** (0.22)	0.92*** (0.13)	0.88*** (0.19)	0.70*** (0.16)	0.71*** (0.13)
$\Delta \ln P$	-5.69*** (1.02)	-3.86 ² (3.82)	-15.70*** (3.96)	-11.01*** (1.57)	-9.76*** (1.31)	-7.87*** (1.21)
$SD \ln f_{-1}$	-0.02*** (0.00)			-0.02*** (0.01)	-0.03*** (0.00)	-0.03*** (0.00)
$\ln f_{-1}$	-0.01*** (0.00)					
$\text{LnSk}^{\text{gov}}_{-1}$		0.07*** (0.03)	0.09*** (0.02)	-0.02 (0.02)		
$\text{Ln}(\text{Gov cons})_{-1}$		0.19*** (0.04)	-0.15** (0.06)	0.04 (0.07)		-0.10** (0.05)
LnTax_{-1}		-0.44*** (0.10)		-0.18** (0.07)	-0.12** (0.05)	
LnTaxDistr		-0.08** (0.04)				
$\text{Ln}(\text{Trade exp}^{\text{adj}})_{-1}$	0.20*** (0.05)	0.20*** (0.05)	0.10* (0.05)	0.14** (0.06)	0.20*** (0.06)	0.22*** (0.06)
Convergence coefficient						
LnY_{-1}	-0.17*** (0.02)	-0.17*** (0.04)	-0.21*** (0.05)	-0.13*** (0.03)	-0.15*** (0.03)	-0.15*** (0.03)
No. of countries	21	17	21	17	18	21
No. of observations	523	427	522	427	444	523
Log likelihood	1 553	1 362	1 541	1 595	1 349	1 556

1. All equations include short-run dynamics and country-specific terms. Moreover, they control for outliers. The variables are defined in Box 1. Standard errors are in brackets.

*: significant at 10 % level; ** at 5 % level; *** at 1 % level.

2. The Hausman test rejected the hypothesis of common long-run coefficient and thus the coefficient was estimated without cross-country restrictions.

tional variable of interest to vary across countries is reported alongside results where the homogeneity of long-run coefficients is imposed. It should be recalled that all the augmented regressions include investment (in physical capital) as an explanatory variable. Hence, the results can be interpreted as showing the effect on output over and above that which may be operating indirectly via investment. Since our results suggest a significant (positive) impact of trade exposure on GDP per capita, whenever possible the equations include this variable in the specification of the growth equation.

Table 4. **Regressions including indicators of financial development¹**
(Pooled Mean Group Estimators)

Dependent variable: $\Delta \ln Y$	with private credit	... and control for inflation	with stock market capitalisation
Long-Run Coefficients			
$\ln Sk$	0.07 (0.06)	0.30*** (0.06)	0.14*** (0.02)
$\ln H$	1.04*** (0.12)	0.99*** (0.14)	0.93*** (0.15)
$\Delta \ln P$	-14.48*** (2.34)	-11.54*** (1.77)	-4.80*** (0.89)
$\ln(\text{Priv credit})_{-1}$	-0.14*** (0.04)	0.04** (0.02)	
$\ln(\text{Stock cap})_{-1}$			0.09*** (0.01)
$SD\ln f_{-1}$		-0.02*** (0.00)	
Convergence coefficient			
$\ln Y_{-1}$	-0.10*** (0.02)	-0.13*** (0.02)	-0.22*** (0.05)
No. of countries	21	21	18
No. of observations	523	523	338
Log likelihood	1449	1498	1058
1. All equations include short-run dynamics and country-specific terms, and control for outliers. Variables are defined in Box 1. Standard errors are in brackets. *: significant at 10 % level; ** at 5% level; *** at 1 % level.			

Overall, the results suggest a significant impact of macro policy settings on output per capita across countries and over time. The regression results suggest that the variability of inflation has an important influence on output per capita: its estimated coefficient is always negative and more than two standard errors from zero (Table 3). This result supports the hypothesis that high variability of inflation adds noise to capital and other markets, with the repercussion likely including an inefficient choice of potential investment projects, with lower average returns of the set of projects actually undertaken. The effect of the level of inflation is less clear-cut:²² in the trade-augmented specifications presented in Table 3, the level of inflation seems to have a negative and significant impact on the steady state level of GDP per capita, but this is not always so when the trade variable is excluded.²³

The hypothesis that the size of government has an impact on growth receives some qualified support (Table 3). The overall measure of tax and non-tax revenue can be used only for a sub-sample of 18 OECD countries, due to data availability.

The sample is further reduced to 17 countries when a distinction between direct and indirect taxes is introduced. The overall tax burden is found to have a negative impact on output per capita.²⁴ Furthermore, controlling for the overall tax burden, there is an additional negative effect coming from an extensive reliance of direct taxes. With control for the overall size and distribution of the tax burden, both government consumption and investment seem to have a positive impact on output per capita. The second specification on fiscal policy considers only the expenditure side of the government budget, and allows extending the sample to 21 countries. In this case, the coefficient on government consumption becomes negative (and statistically significant). This suggests that focusing on one side of the budget and ignoring the other leads to systematic biases associated with the “implicit financing” assumption.²⁵ In particular, given the high within-country correlation between the tax variable and government consumption,²⁶ the coefficient on consumption when the tax variable is not included should be taken to indicate the effect on growth of the “size” of government, rather than the true effect on growth of one specific element of total expenditure.

The equations in the last three columns of Table 3 include both the variability of inflation and the different fiscal policy variables. The key result is the stability of the coefficient for the variability of inflation across all specifications and the continued negative, if somewhat uneven, impact of government size, whether proxied by total tax burden or by government consumption in the last column of the table. By contrast, government investment becomes insignificant as soon as the model is extended and is dropped in the final specifications on the right-hand side of the table.

Some indication of the link between financial development and growth is presented in regressions including indicators of private credit from the banking sector and stock market capitalisation (see also Leahy *et al.*, 2001). Table 4 gives general support to the notion that the level of financial development influences growth, even after controlling for the propensity to invest. This perhaps points to a greater capacity of more developed financial systems to channel resources towards projects with higher returns. The results point to a robust link between stock market capitalisation and growth, while that between private credit provided to the private sector and growth has the wrong sign, when there is no control for inflation variability. This may reflect the fact that the banking credit indicator is related to other monetary variables, including money supply and demand conditions. Indeed, an extended model that also includes inflation variability points to a positive relationship between private credit and growth.

Research and development

The analysis of the determinants of growth can be further extended to include R&D activities, even though the sample is smaller and inference therefore more

tentative. In particular the analysis is restricted to 14-17 countries depending on the specification, and to the period 1981-98 (and for some countries the period is shorter). The shorter time-series significantly restrict the number of variables that could be considered in the regressions. These include, in addition to the R&D variables, the basic controls and trade exposure, whenever possible.²⁷ The indicators of R&D activity used here are expenditures on R&D as collected in national accounts expressed as a percentage of GDP and are thus indicators of the “intensity” of R&D within each country. The results (Table 5) support previous evidence suggesting a significant effect of R&D activity on the growth process.²⁸ Furthermore, regressions including separate variables for business-performed R&D and the R&D performed by other institutions (mainly public research institutes) suggest that it is the former that drives the positive association between total R&D

Table 5. **Regressions including R&D intensity¹**
(Pooled Mean Group Estimators)

Dependent variable: $\Delta \ln Y$	With total R&D	With distinction between business and non-business R&D	With business R&D only
Long-Run Coefficients			
$\ln Sk$	0.31*** (0.03)	0.28*** (0.02)	0.34*** (0.02)
$\ln H$	1.13*** (0.16)	1.76*** (0.05)	0.82*** (0.18)
$\Delta \ln P$	-12.15*** (1.64)	-33.19*** ² (13.94)	-16.43*** (2.02)
$\ln R\&D^{tot}$	0.14*** (0.03)		
$\ln BERD$		0.26*** (0.01)	0.13*** (0.02)
$\ln R\&D^{pub}$		-0.37*** (0.04)	
$\ln(Trade\ exp^{adj})_{-1}$	0.33*** (0.05)		0.32*** (0.05)
Convergence coefficient			
$\ln Y_{-1}$	-0.22*** (0.05)	-0.23** (0.11)	-0.18*** (0.04)
No. of countries	16	15	16
No. of observations	252	236	251
Log likelihood	860	831	849

1. All equations include short-run dynamics and country-specific terms. Variables are defined in Box 1. Moreover, they control for outliers.
Standard errors are in brackets.

*: significant at 10 % level; ** at 5% level; *** at 1 % level.

2. The Hausman test rejected the hypothesis of common long-run coefficient and thus the coefficient was estimated without cross-country restrictions.

intensity and output growth.²⁹ The results also indicate that the coefficient on business-sector R&D is somewhat lower in the regression with the indicator of trade exposure. This suggests possible interactions between R&D and international trade: the estimated impact of R&D on growth may be overestimated if no proper account is taken of the degree of market openness of a country. Nevertheless, the R&D coefficients remain largely significant in this augmented regression.

The negative results for public R&D needs some qualification. Taken at face value they suggest publicly performed R&D crowds out resources that could be alternatively used by the private sector, including private R&D. There is some evidence of this effect in studies that have looked in details at the role of different forms of R&D and the interaction between them.³⁰ In particular, it is found that defence research performed by the public sector does indeed crowd out private R&D, partly by raising the cost of research. However, there are avenues for more complex effects that regression analysis cannot identify. For example, while business-performed R&D is likely to be more directly targeted towards innovation and implementation of new innovative processes in production (leading to improvement in productivity), other forms of R&D (*e.g.* energy, health and university research) may not raise technology levels significantly in the short run, but they may generate basic knowledge with possible “technology spillovers”. The latter are difficult to identify, not least because of the long lags involved and the possible interactions with human capital and associated institutions.³¹

Bearing these caveats in mind, the coefficient on business-performed R&D intensity, if interpreted structurally, suggests that a persistent 0.1 percentage point increase in R&D intensity (about 10 per cent increase with respect to average R&D intensity) would have a long-run effect of about 1.2 per cent higher output per capita under the “conservative” view that changes in R&D do not permanently affect output growth. However, in the case of R&D it is perhaps more appropriate to consider a permanent effect on GDP per capita growth (*i.e.* a fall in R&D intensity is not likely to reduce the steady-state level of GDP per capita but rather reduce technical progress). If the R&D coefficient is taken to represent growth effects, a 0.1 percentage point increase in R&D could boost output per capita growth by some 0.3-0.4 per cent. These estimated effects are large, perhaps unreasonably so, but nevertheless point to significant externalities in R&D activities.

Policy and institutional influences on capital accumulation

To explore whether a given policy influences growth (or, in our interpretation, the steady-state level of output) indirectly *via* its impact on the accumulation of physical capital, investment-share regressions are shown in Table 6. The estimation approach is similar to that of the growth regressions. Following experimentation

Table 6. **Investment regressions**¹
(Pooled Mean Group Estimators)

Dependent variable: $\Delta \ln Sk$				
Long-Run Coefficients				
$Sdinf_{-1}$	-0.02*	-0.01*		
	(0.01)	(0.01)		
Inf_{-1}	-0.02***	-0.03***	-0.02***	-0.03***
	(0.01)	(0.00)	(0.00)	(0.01)
$\ln Sk^{gov}_{-1}$	-0.21***	-0.11**	0.02	-0.05
	(0.06)	(0.04)	(0.03)	(0.03)
$\ln(Gov\ cons)_{-1}$	-0.26*		-0.71***	
	(0.15)		(0.14)	
$\ln Tax_{-1}$		-0.77***		-0.36**
		(0.12)		(0.14)
$\ln(Stock\ cap)_{-1}$			0.14***	0.17***
			(0.01)	(0.02)
$\ln(Priv\ credit)_{-1}$	0.09**	0.06		
	(0.03)	(0.04)		
$\ln(Trade\ exp^{adj})_{-1}$	-0.32***	-0.05	0.05	-0.31***
	(0.12)	(0.08)	(0.10)	(0.09)
Convergence coefficient				
$\ln Sk_{-1}$	-0.15***	-0.22***	-0.27***	-0.26***
	(0.03)	(0.05)	(0.07)	(0.05)
No. of countries	21	18	18	16
No. of observations	531	443	338	301
Log likelihood	936	776	693	601

1. All equations include short-run dynamics and country-specific terms, and control for outliers. Variables are defined in Box 1.
Standard errors are in brackets.
*: significant at 10 % level; ** at 5% level; *** at 1 % level.

with three control variables – lagged output per capita, human capital and lagged trade exposure – the preferred specification includes only a control for trade exposure.

The variability of inflation has a negative coefficient but it is not significant at the standard levels. Interestingly, the coefficient of the level of inflation is strongly significant and negative in the investment equation, in contrast with the weaker result in the growth regressions. These results are consistent with the view that uncertainty about price developments mainly influences growth via distortions in the allocation of resources (as discussed in the context of the growth equation), rather than via discouraging the overall accumulation of physical capital, while high levels of inflation indeed discourage saving and investment. There is also evidence that the “size” of government may be negatively associated with the rate of accumulation of private capital. This can be seen by looking at the coefficients

on either taxes or government consumption (estimation with the latter allows the use of a larger sample). Possible interactions between private and government investment were also tested, but the results do not yield robust results of either a negative or positive link.

The coefficients on financial variables in the investment regressions in Table 6 have the expected signs and are (mostly) significant. As in the growth regressions, the indicator of credit provided by the banking sector appears to be only weakly associated with investment, while the stock market capitalisation has a stronger bearing on investment.

SOME QUANTITATIVE IMPLICATIONS OF THE REGRESSION RESULTS

The estimated coefficients of the different growth regressions can be used to shed light on the role of policy and institutional settings for the growth experience of different countries over the past two decades. Two important caveats need to be borne in mind in this exercise. As discussed above, it is assumed that the policy and institutional variables affect only the *level* of economic efficiency and not the steady-state growth rate. Moreover, the calculations should only be taken as broad indications, given the variability of coefficients across the specifications, and interaction effects that may be important but cannot be taken into account. The estimated coefficients are used to perform three calculations: *i*) the effect of a given change in a policy or institutional variable on steady-state output per capita; *ii*) the decomposition of observed differences in average growth rate across countries over a 20-year period; and *iii*) the decomposition of observed changes in growth rate within each country into its determinants. In the first calculation, the direct effect of a policy change and the indirect effect (via investment) are considered, while in the other two calculations only the direct effect is estimated and investment in physical capital is considered as one of the exogenous components of the decomposition. Hence, in the last two calculations the overall contribution of different policy levers to differences and changes in growth rates may well be different if one takes into account their impact via the investment channel.

The long-run effect of policy and institutional changes

Bearing the above-mentioned caveats in mind, the total effects of policy and institutional changes on GDP per capita can be summarised as follows (see Table 7):

- The point estimate for the variability of inflation suggests that a reduction by 1 percentage point in the standard deviation in inflation – *e.g.* about one and a half times the reduction recorded on average in the OECD countries

Table 7. **Estimated impact of changes in institutional or policy factors on GDP per capita¹**

Variable	Impact on output per working age person (per cent) ²			Order of magnitude with respect to OECD experience (1980s-90s) ³
	Effect via economic efficiency	Effect via Investment	Overall effect	
Inflation rate (fall of 1% point)		0.4 to 0.5	0.4 to 0.5	about 1/4 the observed fall
Variability of inflation (1% point fall in SD of inflation)	2.0		2.0	about 1.5 times the observed fall
Tax burden⁴ (increase of 1% point)	-0.3	-0.3 to -0.4	-0.6 to -0.7	about 2/3 of the observed increase
Business R&D intensity⁴ (increase of 0.1% points)	1.2		1.2	about the increase observed
Trade exposure⁴ (increase of 10% points)	4.0		4.0	about the increase observed

1. The values reported in this table are the estimated long-run effects on output per working-age person of a given policy change. The range reported reflects the values obtained in different specifications of the growth equation.
2. The direct effect refers to the impact on output per capita over and above any potential influence on the accumulation of physical capital. The indirect effect refers to the combined impact of the variable on the investment rate and by that channel, on output per capita.
3. Average change from the 1980 average to the 1990 average in the sample of 21 OECD countries, excluding new Members as well as Iceland, Luxembourg and Turkey.
4. In percentage of GDP.

from the 1980s to the 1990s – could lead to an increase in long-run output per capita by 2 per cent, *ceteris paribus*.

- The effect of the level of inflation mainly works through investment: a reduction of one percentage point – *e.g.* one-fourth of what was recorded in the OECD between 1980s and 1990s – could lead to an increase in output per capita of about 0.4-0.5 per cent, over and above what could also emerge from any accompanying reduction in the variability of inflation.
- Taxes and government expenditures affect growth both directly and indirectly through investment. An increase of about one percentage point in the tax pressure – *e.g.* two-thirds of what was observed over the past decade in the OECD sample – could be associated with a direct reduction of about 0.3 per cent in output per capita. If the investment effect is taken into account, the overall reduction would be about 0.6-0.7 per cent.

- Finally, an increase in trade exposure of 10 percentage points – about the change observed over the past decade in the OECD sample – could lead to an increase in steady-state output per capita of 4 per cent.

Explaining cross-country differences in average growth rates

Differences in growth rates across countries depends on two factors: i) differences in the long-run equilibrium level of GDP per capita; and ii) differences in the initial conditions (the actual level of GDP per capita at the beginning of period considered). A country with the same initial level of GDP per capita but better framework conditions – and thus a higher steady state level of GDP – than another country will grow faster. This is due to the greater distance from its own steady state. Similarly, a country with a lower initial level of GDP than another country but the same steady state level will grow faster, again, because of the wider gap to be closed. Moreover, within any period of time, the steady-state level of GDP per capita could change because of changes in framework conditions, and this third factor has also to be taken into account. Formally, for each country, the decomposition of observed increases in GDP per capita over a time-period of length t into a set of explanatory variables over the same period can be expressed as follows:

$$(\ln y - \ln y_{-t}) = [(1 - \phi)^t - 1] \ln y_{-t} + \sum_j \sum_{m=1}^t (1 - \phi)^{m-1} \phi \cdot \theta_j X_{j, -(m-1)} \quad (4)$$

where $\mathbf{X} = \{\ln sk, \ln h, n, \mathbf{V}\}$ is the vector of independent variables (determinants of the steady state), θ is the vector of their long-run coefficients, ϕ is the convergence coefficient; $-t$ indicates initial conditions and $-m$ indicates that the variable is lagged m periods.³² The first term on the right-hand side identifies the effect of initial conditions on changes in GDP per capita, while the second accounts for the determinants of the steady state. This decomposition can be used for two purposes: i) to compare average growth rates across countries and assess the role of initial conditions and differences in *levels* of determinants of the steady-state; and ii) to track changes in GDP per capita growth rates within each country over sub-periods and identify the impact of *changes* in the determinants of the steady state. For example, the average level of human capital in Spain was below the OECD average over the period 1970-98 and, thus, this contributed to a lower than average growth in GDP per capita (*ceteris paribus*). At the same time, however, human capital has grown rapidly in Spain from the 1970s to the 1980s and 1990s, pushing forward Spain's steady state GDP per capita more than in most of the other countries: this has resulted in an acceleration in GDP per capita growth, again *ceteris paribus*.

The formula is derived from equation [2] above (omitting short-run dynamics for the independent variables), that is:

$$\ln y = (1 - \phi) \ln y_{-1} + \phi \sum_j \theta_j X_{j,0} \quad (5)$$

Iterating equation (5) $t-1$ times yields:

$$\ln y = (1 - \phi)^t \ln y_{-t} + \phi \sum_{m=1}^t \sum_j (1 - \phi)^{m-1} \theta_j X_{j, -(m-1)} \quad (6)$$

Then, equation (4) can be derived by rearranging equation (6).

Table 8 shows the decomposition in relative terms (to the OECD simple average). To maximise the country coverage, we use the growth regression on the right-hand-side of Table 3. This is run on 21 countries and uses government consumption as a proxy for the effect of government “size” on growth. The results suggest that the model fits the data rather well: there are only three countries where the unexplained growth rate differential is large in absolute terms (last column of Table 8).³³ In two of these three countries (Greece and Portugal) the model would have predicted a higher growth rate than that actually recorded. By contrast, the model under-predicts the average growth rate in the United States. In these three countries, additional factors not accounted for in the present analysis therefore played a significant role in shaping the growth process. The results in Table 8 confirm some priors about the driving forces of output growth in the OECD countries. In the English-speaking countries, a relatively low saving/investment rate had a negative impact on growth, *ceteris paribus*. A relatively low level of human capital compared with the OECD average over the period negatively influenced growth in a number of European countries, but especially in Portugal and Spain. In addition, some countries, including Australia, Canada, Ireland and New Zealand, had a somewhat lower per capita growth as a result of a rapidly growing population. Average macro policy conditions also had a bearing in shaping cross-country differences in growth. Thus, the higher than average variability of inflation had a negative impact in Greece and, to some extent in Portugal, while the large “size” of government seems to have negatively influenced growth in Denmark and Sweden. Finally, the relatively low exposure to foreign trade (after controlling for the size of each country) seems to have had a negative impact on growth in Australia and New Zealand, possibly reflecting geographical location, while the reverse occurred in Belgium and Netherlands – though the trade exposure for these two countries reflects, to some extent, the large cross-border transactions amongst themselves – as well as in the United Kingdom.

Table 8. **Decomposition of country deviations from OECD average output per capita growth rates, 1970s-1990s¹**
(annual percentage point growth rates)

	Annual average growth rate	Growth differential	Initial conditions (real GDP/pop)	Investment share (Sk)	Human capital (H)	Population growth (ALnp)	Variability of inflation (SDinfl)	Government consumption (Gov cons)	Trade exposure (Trade exp ^{adj})	Residual country specific effect
Australia	1.68	0.13	-0.37	0.20	0.52	-0.25	0.03	0.01	-0.41	0.40
Austria	1.57	0.02	-0.41	0.07	0.26	0.01	0.05	0.00	0.03	0.01
Belgium	1.66	0.11	-0.53	0.02	-0.15	0.20	0.03	-0.05	0.53	0.06
Canada	1.32	-0.23	-0.90	-0.21	0.62	-0.18	0.04	-0.07	0.14	0.32
Denmark	1.69	0.14	-0.57	0.28	0.21	0.12	0.02	-0.14	-0.05	0.27
Finland	1.82	0.27	0.51	0.05	0.02	0.15	0.00	-0.06	-0.26	-0.14
France	1.35	-0.20	-0.59	-0.09	-0.10	0.07	0.07	-0.08	0.05	0.48
Greece	1.15	-0.40	2.00	0.19	-0.56	-0.07	-0.16	0.17	-0.51	-1.48
Ireland	3.02	1.47	1.54	-0.18	-0.32	-0.18	0.01	0.09	0.17	0.34
Italy	1.73	0.18	0.22	-0.13	-0.69	0.13	0.02	0.01	0.14	0.48
Netherlands	1.26	-0.29	-0.47	-0.03	0.25	0.01	0.06	-0.13	0.52	-0.50
New Zealand	0.53	-1.02	0.34	-0.17	0.31	-0.29	-0.07	0.10	-0.36	-0.87
Norway	1.72	0.17	-0.12	-0.05	0.35	0.07	0.03	-0.06	-0.04	-0.01
Portugal	2.15	0.60	2.56	0.58	-1.20	0.07	-0.10	0.10	0.11	-1.52
Spain	1.28	-0.27	0.73	0.04	-1.12	0.00	0.03	0.07	-0.14	0.11
Sweden	1.20	-0.35	-0.60	-0.10	0.21	0.11	-0.10	-0.17	0.01	0.30
Switzerland	0.81	-0.74	-1.75	0.08	0.59	-0.04	0.00	0.15	0.02	0.21
United Kingdom	1.63	0.08	0.05	-0.21	0.17	0.15	-0.03	-0.02	0.31	-0.34
United States	1.93	0.38	-1.62	-0.34	0.63	-0.09	0.07	0.09	-0.25	1.89

1. Decomposition of cross-country differences in annualised growth rate (in %) from the average output per capita of 1974-77 to the average output per capita of 1994-97. See main text for details on the decomposition.

The role of policy and institutions in shaping the growth process over the past two decades

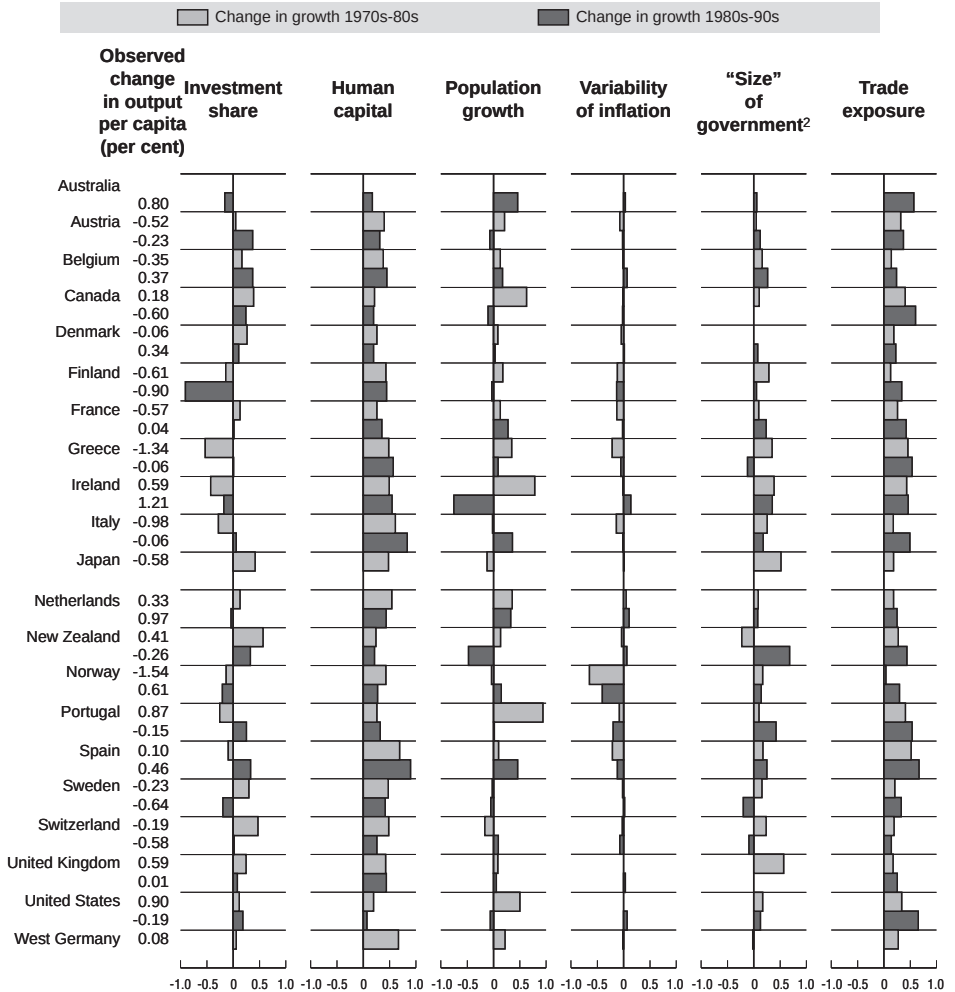
Policy and institutional settings have changed significantly over the past decades and it is also important to examine the possible impact of these changes on the growth path of each country (Figure 4). The improvement in human capital has been one of the key factors behind the growth process of the past decades in all OECD countries, but especially so in Germany, Italy, Greece, Netherlands (mainly in the 1980s) and Spain, where the increase in human capital accounted for more than half a percentage point acceleration in growth (in both the 1980s and 1990s) with respect to the previous decade (even though the level of human capital remained low compared to other countries). The contribution stemming from changes in the investment rate is more mixed. Some countries are estimated to have benefited from an increase in the investment rate in the past two decades with respect to the 1970s (*e.g.* Japan, Canada, Austria, Belgium, New Zealand), while others could have had a negative impact from lower investment rates (*e.g.* Italy and Ireland in the 1980s, Finland in the 1990s).

There have also been important changes in policy and institutional settings in each country that have contributed to growth, over and above the changes in factor inputs. Most countries have benefited, especially in the 1990s, from reduced uncertainty due to a lower variability of inflation. The most noticeable examples include the United Kingdom and Japan (in the 1980s) and Portugal and New Zealand (in the 1990s) where about half a percentage point higher annual output per capita growth rate is estimated to be due to the lower variability of inflation, *ceteris paribus*. By contrast, in spite of public spending restraint, especially in the last decade, the rise in the size of government contributed to slow down growth in most countries. Notable exceptions include the United States, Ireland and the Netherlands where a reduction in taxes and expenditures as a share of GDP somewhat boosted output per capita growth in the 1990s. Finally, but not least, the generalised process of trade liberalisation in which all OECD countries have been involved is estimated to have increased growth by up to two-thirds of a percentage point annually over the past decade.

CONCLUDING REMARKS

In broad terms, the estimated growth regressions explain much of the observed growth paths across countries and over time. One striking result of our study is the high speed with which countries seem to converge to their steady-state growth path compared with previous estimates based on a larger set of countries and cross-section data. This implies that observed cross-country

Figure 4. The estimated effect of changes in explanatory variables to changes in output per capita growth rates¹



Notes: The calculations are from decompositions of differences in growth rates based on the results of multivariate regressions. The estimated impact of initial levels of GDP per capita and the component unexplained by the regressions are not shown.

1. The changes in growth are based on differences in average growth in GDP per person of working age over each decade.
2. Government consumption as a percentage of GDP is used as a proxy for the size of government due to data availability. This variable is highly correlated in most countries with the tax and non-tax receipts (as a share of GDP) for which, however country coverage is more limited.

differences in GDP per capita levels may be largely the result of differences in steady-state levels rather than different positions of countries along similar transitional paths. In consequence, differences in investment rates and human capital as well as in R&D, trade exposure, financial structures and macroeconomic conditions and policy settings seem to play an important role for observed GDP per capita patterns across countries. Changes in these factors can be rapidly translated into changes of living standards. The other main results of our study are as follows:

- The estimated partial elasticity of output to physical capital is consistent with the values implied in National Accounts data, even though it is on the low side of the range. By contrast, and if taken at face value, the estimated elasticity of output to human capital points to potential externalities in investment in education, *i.e.* social returns seem higher than private returns at least over the past decades, when education levels were relatively low.
- The evidence suggests that high inflation is negatively associated with the accumulation of physical capital in the private sector and, through this channel, has a negative bearing on output. Moreover, a high variability of inflation affects GDP per capita, possibly because it leads to a shift in the composition of investment towards less risky but also lower return projects.
- In addition, the empirical evidence lends some support to the notion that the overall size of government in the economy may reach levels that hinder growth. Although expenditure on health, education and research clearly sustains living standards in the long term, and social transfers help to meet social goals, all have to be financed. The results suggest that for a given level of taxation, higher direct taxes lead to lower output per capita, while, on the expenditure side, government consumption and government investment tend to have non negative effects on output per capita. Government investment may also influence growth by improving the framework conditions (*e.g.* better infrastructure) in which private agents operate.
- Research and development (R&D) activities undertaken by the business sector seem to have high social returns, while no clear-cut relationship could be established between non-business-oriented R&D activities and growth. There are, however, possible interactions and international spillovers that the regression analysis cannot identify. Moreover, non-business oriented R&D (*e.g.* energy, health and university research) may generate basic knowledge with possible “technology spillovers” in the long run.
- The empirical evidence also confirms the importance of financial markets for growth, both by helping to channel resources towards the most rewarding activities and in encouraging investment. In particular, the degree of

stock market capitalisation is found to be strongly related with both output per capita (while controlling for investment) and with the investment rate.

All in all, the results suggest that differences in GDP per capita across the OECD countries can be largely explained by different policy and institutional settings and that countries can learn from each other on the optimal growth strategy. Notwithstanding persistent differences in living standards, recent policy changes seem to go in the right direction to enhance growth. Most countries have made significant progress towards price stability and avoiding excessive macroeconomic fluctuations. However, while there have been successful efforts to reduce public-sector deficits, the overall tax pressure is still high in a number of them and has risen in the past decade. On the structural side, most OECD countries have recorded significant increases in their human capital, not least because of government interventions. Even if it is possible that there are diminishing social returns to any given increase in education levels, these developments have had (and will have in the future) a positive impact on observed growth patterns. Furthermore, the amount of resources devoted to R&D has generally increased from the 1980s to the 1990s – though there has been some reduction in recent years largely due to falls in defence-related government outlays. Moreover, more resources seem to be channelled directly to the business sector with a greater role played by industries themselves. Beyond these considerations, there remain significant differences in growth rates across countries, possibly due to differences in other framework conditions (*e.g.* regulation in the product and labour markets) which have not been taken into account in this paper.

NOTES

1. It should also be stressed that evidence of convergence in GDP per capita amongst OECD countries is largely concentrated in the post-war period: during most of the 19th century, most OECD countries were falling behind the United Kingdom and during the first half of the 20th century most of them were falling behind the United States. See Maddison (2001).
2. Indeed, new-growth models that incorporate a knowledge-producing sector can be interpreted as incorporating the role that, for example, research universities may play in the growth process. An early example of this type of model is in Uzawa (1965), later examples include Lucas (1988), Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1998).
3. Amongst the several studies that suggest a negative association between uncertainty on the one hand, and investment and growth, on the other, see Dixit and Pindyck, 1994 and Bertola and Caballero, 1994. Amongst the studies that have put into question the negative association between uncertainty and investment are Abel (1983) and Hartman (1972). They suggest that in an economy with no frictions, an increase in uncertainty could even lead to higher investment rates.
4. In most OECD countries, government finances the bulk of expenditure on educational institutions. See OECD (2001) for more details.
5. For example, Coe and Helpman (1995) find significant interaction between import propensities and the ability to benefit from foreign R&D: *i.e.* for a given level of R&D performed abroad, countries with higher import propensity have higher productivity growth. Moreover, small countries benefit more from R&D performed abroad than from domestic R&D. Sachs and Warner (1995) claim trade openness as being an important constraint to convergence for many of the world's economies. Using aggregate data on trade between (mainly) OECD countries, Ben-David and Kimhi (2000) find evidence to support the idea that increasing trade between pairs of countries is associated with an increased rate of convergence.
6. See *e.g.* Frankel and Romer (1999) and Baldwin (2000).
7. Where data for a large number of countries was available, growth regressions have typically taken averages over long time periods (*e.g.* 20 years). Other studies have taken averages over five-year periods (see *e.g.* Islam, 1995; Caselli *et al.*, 1996). This choice, however, implies a potential loss of information. Moreover, the lack of synchronisation in country business cycles does not purge five-year averages from cyclical influences.
8. This is, for example, the case in one-sector models of endogenous growth in which capital is not characterised by diminishing returns (see *e.g.* Romer, 1986; Rebelo, 1991).
9. This is the case in models of endogenous growth, which explicitly consider different types of capital goods (*e.g.* physical and human), each characterised by its own accu-

mulation process (*e.g.* investment and education). See Uzawa (1965); Lucas (1988); Barro and Sala-i-Martin (1995).

10. As discussed in detail by Klenow and Rodríguez-Clare (1997), most cross-section studies of growth (*e.g.* Barro and Sala i Martin, 1995) estimate a regression of the general form:

$$\Delta \ln y_{i,T} = -b \cdot \ln y_{i,0} + \text{policy and institutional variables} + u_{i,T}$$

where $\Delta \ln y$ is the growth rate of GDP per capita over a period of time (0 to T); $\ln y_0$ is the log of GDP per capita at time 0; and i identifies each country. They derive this equation from a simple neo-classical growth model:

$$\Delta \ln y_{i,T} = g_i + \beta \cdot (\ln y_{i,0}^* - \ln y_{i,0}) + \varepsilon_{i,T}$$

where $\ln y_{i,0}^*$ is the log level of GDP per capita on the country's steady-state path. However, it is not clear whether the policy and institutional variables in the first equation are proxying for differences in country steady-state GDP per capita levels ($\ln y_{i,0}^*$) (consistent with conditional convergence) or for cross-country differences in the long-run growth rates g_i .

11. Estimates of the speed of convergence to steady-state output vary in the literature: while most studies estimated values around 2-3 per cent per year (Mankiw, *et al.*, 1992; Barro and Sala-i-Martin, 1995) – which implies that an economy spends about 20-30 years to cover half of the distance between its initial conditions and its steady state – a few have found values of 10 per cent or more for the OECD countries (*e.g.* Caselli *et al.*, 1996), which imply less than nine years to cover half of the distance.
12. Under slope heterogeneity, estimates of a uniform convergence parameter are affected by a downward heterogeneity bias (Pesaran and Smith, 1995).
13. In a theoretical growth model, ϕ is a function of population growth ($n_{i,t}$) and technological progress ($g_{i,t}$) and thus could vary across countries and over time. For the purpose of the econometric analysis, time homogeneity had to be imposed, but the parameters are allowed to vary across countries.
14. If the homogeneity assumption was not retained, the reported coefficient in the tables is the simple average of country-specific coefficients.
15. This result may arise because for a number of OECD countries the average number of years of schooling has increased steadily over the sample period. Our sensitivity analysis (Bassanini and Scarpetta, 2001) suggests that a time trend is only statistically significant when human capital is omitted and is not statistically significant at standard confidence levels when human capital is included. Consequently, in the retained specification, human capital was included while the time trend was dropped.
16. The identification of outliers is based on the analysis of *studentised residuals* and *leverage values* (see Belsley *et al.*, 1980). The *studentised residuals* are obtained by considering a mean-shift outlier model in which the basic equation is augmented by a dummy variable that has the *i*th element equal to one and all other elements zero. The studentised residual is the t-statistics of the dummy variable. The *leverage point* is identified by the diagonal elements of the least-squared projection matrix, also called the *hat* matrix. It proxies the distance between the *i*th observation and the centre of the data. The outliers that have been removed from the sample are those with a studentised residual greater than 2.5 and a leverage point above the cut-off value suggested by Belsley *et al.* (1980). Eight outliers have been identified using dynamic fixed effect

estimators: the 1972-73 observations for New Zealand; 1995-97 for Ireland; 1972 for Finland and 1972-73 for Portugal.

17. The country sample include: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (western), Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and United States.
18. However, it should also be stressed that the evidence on the returns to education reflects average rates of return based on historical data, whereas future marginal returns may not be as high.
19. It should be stressed that these conclusions do not depend on a particular specification of the growth equation: inclusion of a linear time trend, introduction of country-specific time dummies, sample variations across countries and over time, all lead to similar results (Bassanini and Scarpetta, 2001).
20. One of the reasons behind the difference in the present results on human capital has to do with the quality of the data used in this paper compared with previous attempts. We used a revised version of the Barro and Lee (1996) database as assembled by de la Fuente and Doménech (2000) in combination with more recent OECD data (see Appendix 2). Revisions were largely aimed at removing time and cross-country inconsistencies in the original database. Using a similar proxy, de la Fuente and Doménech (2000) also found a strongly significant coefficient for human capital in level and growth equations.
21. The neo-classical model makes precise predictions on the value of the estimated long-run parameters of the human and physical capital, as well as on population growth. Furthermore, the speed of convergence $[-\log(1-\hat{\rho})]$ can be expressed formally as a function of the rate of technological progress, the rate of growth of the population, the depreciation rate of physical and human capital as well as output elasticities to human and physical capital, the latter being derived from the coefficients on sk and h in equation [2] above. The estimated speed of convergence to the steady state path of output per capita in our equations is too high with respect to what would be implied by the estimated value of the elasticity of output with respect to "broad" (physical and human) capital, which would imply very slow convergence (see Bassanini and Scarpetta, 2001).
22. The uncertain result for the level of inflation is in line with Alexander (1997) who also used an OECD sample. However, other papers that did not include the variability of inflation in the growth equation found more solid links between the level of inflation and growth, even in the OECD samples. See, for example, Andres and Hernando (1997) and Englander and Gurney (1994).
23. The cross-country homogeneity restriction on the coefficient of the level of inflation is rejected at the 5 per cent level in the model that does not include the trade exposure variable and, once allowed to vary across countries, this coefficient becomes statistically insignificant.
24. The latter result is consistent with that of Folster and Henrekson (1998, 2000) who focus on the link between the "size" of government and growth in OECD countries. However the conclusions reached by Folster and Henrekson have been questioned by some researchers, notably Agell *et al.* (1997).
25. See Helms (1985); Mofidi and Stone (1989); Kneller *et al.* (1998).
26. The cross-country correlation between the government consumption variable and the tax and non-tax receipts variable is greater than 0.5 in each year of the sample and

always greater than 0.6 after 1976. In time-series the correlation is greater than 0.9 in 9 out of 18 countries and greater than 0.7 in all but three countries (Belgium, the Netherlands, and the United States).

27. The growth regression maintains its basic properties when estimated over the smaller sample used in the R&D regressions. The coefficients on both physical and human capital maintain their sign and statistical significance, although the convergence is higher than in the regression estimated over the larger sample. This latter result is not driven by the small country sample but rather by the shorter time period over which the model is estimated.
28. In terms of previous evidence, Fagerberg (1994), for example, found a patent-based index significant in growth regressions; and Englander and Gurney highlighted R&D expenditure as a robust variable in their growth regressions.
29. Park (1995) also found private-sector R&D more important than public R&D in OECD-based growth regressions.
30. Lichtenberg (1988) finds that non-competitive R&D procurement tends to crowd out private R&D investment, while competitive procurement stimulates private R&D investment. See David *et al.* (1999) for a survey. By contrast, Guellec and Van Pottelsberghe (1997, 2001) support the complementarity hypothesis.
31. Given the short time period that can be used in this sample, lagging the R&D variable would have induced an excessive loss of degrees of freedom.
32. Taking for simplicity the case of a hypothetical independent variable X that is constant over time, equation [4] suggests that the weight of this variable is greater (in absolute terms) the longer the time span over which the decomposition is performed. This property generalises to time-varying independent variables. Hence, to be comparable across countries and/or over time, growth decompositions undertaken according to equation [4] have to refer to the same time span t .
33. A positive country-specific effect implies that actual growth was higher than predicted by the model, and *vice versa*.

Appendix 1

THE POLICY-AND-INSTITUTIONS AUGMENTED GROWTH MODEL

The growth equation

Following a standard approach (see *e.g.* Mankiw *et al.*, 1992; and Barro and Sala-i-Martin, 1995), the standard neo-classical growth model is derived from a constant returns to scale production function with two inputs (capital and labour) that are paid their marginal products. Production at time t is given by:

$$Y(t) = K(t)^\alpha H(t)^\beta (A(t)L(t))^{1-\alpha-\beta} \quad (\text{A1})$$

where Y , K , H and L are respectively output, physical capital, human capital and labour, α is the partial elasticity of output with respect to physical capital, β is the partial elasticity of output with respect to human capital and $A(t)$ is the level of technological and economic efficiency. It can be assumed that the level of economic and technological efficiency $A(t)$ has two components: economic efficiency $I(t)$ dependent on institutions and economic policy, and the level of technological progress $\Omega(t)$ (see amongst others, Cellini *et al.*, 1999 for a similar formulation). In turn, $I(t)$ can be written as, *e.g.* a log-linear function of institutional and policy variables, while $\Omega(t)$ is assumed to grow at the rate $g(t)$.

The time paths of the right-hand side variables are described by the following equations (hereafter dotted variables represent derivatives with respect to time):

$$\begin{aligned} \dot{k}(t) &= s_k(t)A(t)^{1-\alpha-\beta}k(t)^\alpha h(t)^\beta - (n(t)+d)k(t) \\ \dot{h}(t) &= s_h(t)A(t)^{1-\alpha-\beta}k(t)^\alpha h(t)^\beta - (n(t)+d)h(t) \\ A(t) &= I(t)\Omega(t) \\ \ln I(t) &= p_0 + \sum_j p_j \ln V_j(t) \\ \dot{\Omega}(t) &= g(t)\Omega(t) \\ \dot{L}(t) &= n(t)L(t) \end{aligned} \quad (\text{A2})$$

where $\bar{k} = K/L$, $\bar{h} = H/L$, $\bar{y} = Y/L$, stand for the capital labour ratio, average human capital and output per worker respectively; s_k and s_h stand for the investment rate in physical and human capital respectively; and d stands for the (constant) depreciation rate. Under the assumption

that $\alpha + \beta < 1$ (i.e. decreasing returns to reproducible factors), this system of equations can be solved to obtain steady-state values of k^* and h^* defined by:

$$\begin{aligned}\ln k^*(t) &= \ln A(t) + \frac{1-\beta}{1-\alpha-\beta} \ln s_k(t) + \frac{\beta}{1-\alpha-\beta} \ln s_h(t) - \frac{1}{1-\alpha-\beta} \ln(g(t) + n(t) + d) \\ \ln h^*(t) &= \ln A(t) + \frac{\alpha}{1-\alpha-\beta} \ln s_k(t) + \frac{1-\alpha}{1-\alpha-\beta} \ln s_h(t) - \frac{1}{1-\alpha-\beta} \ln(g(t) + n(t) + d)\end{aligned}\quad (A3)$$

Substituting these two equations into the production function and taking logs yields the expression for the steady-state output in intensive form. The latter can be expressed either as a function of s_h (investment in human capital) and the other variables or as a function of h^* (the steady-state stock of human capital) and the other variables. Since in this paper human capital is proxied by the average years of education of the working age population, a formulation in terms of the stock of human capital was retained. The steady-state path of output in intensive form can be written as:*

$$\begin{aligned}\ln y^*(t) &= \ln \Omega(t) + p_0 + \sum_j p_j \ln V_j(t) \\ &\quad + \frac{\alpha}{1-\alpha} \ln s_k(t) + \frac{\beta}{1-\alpha} \ln h^*(t) - \frac{\alpha}{1-\alpha} \ln(g(t) + n(t) + d)\end{aligned}\quad (A4)$$

However, the steady-state stock of human capital is not observed. It can be shown that an expression for h^* as a function of actual human capital is:

$$\ln h^*(t) = \ln h(t) + \frac{1-\psi}{\psi} \Delta \ln(h(t)/A(t))\quad (A5)$$

where ψ is a function of (α, β) and $n + g + d$.

Equation [A4] would be a valid specification in the empirical cross-country analysis only if countries are in their steady states or if deviations from the steady state are independent and identically distributed. If observed growth rates include out-of-steady-state dynamics, then the transitional dynamics have to be modelled explicitly. A linear approximation of the transitional dynamics can be expressed as follows (Mankiw *et al.*, 1992):

$$\begin{aligned}\Delta \ln y(t) &= -\phi(\lambda) \ln(y(t-1)) + \phi(\lambda) \frac{\alpha}{1-\alpha} \ln s_k(t) + \phi(\lambda) \frac{\beta}{1-\alpha} \ln h(t) + \sum_j p_j \phi(\lambda) \ln V_j(t) \\ &\quad + \frac{1-\psi}{\psi} \frac{\beta}{1-\alpha} \Delta \ln h(t) - \phi(\lambda) \frac{\alpha}{1-\alpha} \ln(g(t) + n(t) + d) \\ &\quad + \left(1 - \frac{\phi(\lambda)}{\psi} \right) g(t) + \phi(\lambda)(p_0 + \ln \Omega(0)) + \phi(\lambda) g(t)t\end{aligned}\quad (A6)$$

* Strictly speaking, equation (A4) is written under the simplifying assumption that policy and institutional variables do not change persistently in the long-run. If this is not the case, $\ln(g + n + d)$ must be augmented by a term reflecting the rate of change of policy and institutional variables. As the estimable equation is linearised and contains short-run dynamics anyway, this term will be omitted hereafter for simplicity.

where $\lambda = (1 - \alpha - \beta)(g(t) + n(t) + d)$. Adding short-term dynamics to equation (A6) yields:

$$\begin{aligned} \Delta \ln y(t) = & a_0 - \phi \ln y(t-1) + a_1 \ln s_k(t) + a_2 \ln h(t) - a_3 n(t) + a_4 t + \sum_j a_{j+4} \ln V_j \\ & + b_1 \Delta \ln s_k(t) + b_2 \Delta \ln h(t) + b_3 \Delta \ln n(t) + \sum_j b_{j+4} \Delta \ln V_j + \varepsilon(t) \end{aligned} \quad (A7)$$

Equation (A7) represents the generic functional form that has been empirically estimated in this paper. Estimates of steady state coefficients as well as of the parameters of the production function can be retrieved on the basis of the estimated coefficients of this equation by comparing it with equation (A6). For instance, an estimate of the elasticity of steady state output to the investment rate (that is the long-run effect of the investment rate on output) is given by $\hat{a}_1 / \hat{\phi}$, where $\hat{\cdot}$ identifies estimated coefficients. Conversely, an estimate of the share of physical capital in output (the parameter α of the production function) can be obtained as $\hat{a}_1 / (\hat{\phi} + \hat{a}_1)$.

Appendix 2

THE DATA

The data used in this paper are from the following sources:

- Data on GDP, working-age population, gross fixed capital formation, general government current nominal tax and non-tax receipts, direct and indirect taxes, government nominal final consumption and imports and exports are from the OECD *Analytical Data Base* (ADB). Purchasing Power Parity benchmarks for 1993 are from the OECD Statistics Department. In the case of Norway, data refer to the mainland economy. In the case of Greece and Portugal the ratio between total gross fixed capital formation and total real GDP was used as a proxy for the investment rate (*i.e.* the ratio of private non-residential fixed capital formation to business sector real GDP), due to data availability.
- Data on Research and Development (R&D) are from the OECD *Main Science and Technology Indicators* (MSTI) database. A few missing observations were obtained by interpolation.
- Data on human capital are calculated on the basis of raw data on education attainment. In particular: three educational groups were considered: below upper secondary education (ISCED 0 to ISCED 2); upper secondary education (ISCED 3); and tertiary education (ISCED 5 to ISCED 7). Data on education attainment up to the early 1980s are interpolated from five-year observations from De la Fuente and Doménech (2000), while later observations are from matched OECD sources (*Education at a Glance*, various issues). The cumulative years of schooling by educational level – required to estimate the average number of years of total schooling used in the empirical analysis – are from the OECD *Education at a Glance* – 1997 (OECD, 1998).
- The indicators measuring financial market developments are from the World Bank's financial development database (see Beck *et al.*, 1999). For more details on the pros and cons of these two indicators and for the motivation of their inclusion, see Leahy *et al.* (2001).

The definition of each variable is provided in Box I of the main text. The exact country coverage of the variables is presented in Table A2.1, while the basic statistics are in Table A2.2.

Table A2.1. Details on data availability

Variable	Start date	End date	Exceptions
Real GDP per person of working age (Y)	1971	1998	1971-1997 for Portugal and Spain; 1971-1996 for Greece; 1971-1994 for Western Germany.
Accumulation of physical capital (Sk)	1971	1998	1971-1997 for Austria, Belgium, Ireland, Italy, New Zealand, Portugal, Spain and Sweden; 1971-1996 for the United Kingdom; 1971-1995 for Switzerland; 1971-1990 for Western Germany; 1975-1998 for Australia.
Human capital (H)	1971	1998	1971-1990 for Western Germany and Japan
Growth of working age population ($\Delta \ln P$)	1971	1998	1971-1997 for Spain; 1971-1996 for Greece; 1971-1994 for Western Germany.
Standard deviation of inflation (Sd $\ln \pi_{-1}$)	1971	1998	
Inflation (Infl $_{-1}$)	1971	1998	
Government consumption (Gov $^{cons}_{-1}$)	1971	1998	
Accumulation of physical capital by government (Sk $^{gov}_{-1}$)	1971	1998	1971-1997 for Switzerland; 1971-1996 for Portugal; 1971-1995 for Western Germany.
Tax and non-tax receipts as a share of GDP (Tax $_{-1}$)	1971	1998	1971-1997 for Portugal; 1971-1995 for Western Germany; 1978-1998 for Ireland; 1987-1998 for New Zealand; 1988-1998 for the United Kingdom; no data for Switzerland.
Ratio of direct/indirect taxes (Tax distr $_{-1}$)	1971	1998	1971-1998 for Switzerland; 1971-1996 for Portugal; 1971-1995 for Western Germany; 1978-1998 for Ireland; 1987-1998 for New Zealand; 1988-1998 for the United Kingdom.
Total R&D expenditure as a share of GDP (R&D ^{tot})	1981	1998	1981-1997 for Greece, Ireland, the Netherlands, Sweden and the United Kingdom; 1981-1996 for Australia and Switzerland; 1982-1997 for Portugal; 1983-1995 for Belgium; 1989-1997 for New Zealand; no data for mainland Norway.
Business-performed R&D expenditure as a share of GDP (BERD)	1981	1998	1981-1997 for Australia, Greece, Ireland, the Netherlands, Sweden and the United Kingdom; 1981-1996 for Switzerland; 1981-1995 for Belgium; 1981-1993 for Austria; 1982-1997 for Portugal; 1989-1997 for New Zealand; no data for mainland Norway.
Non-business performed R&D expenditure as a share of GDP (R&D ^{pub})	1981	1998	1981-1997 for Greece, Ireland, the Netherlands, Sweden and the United Kingdom; 1981-1996 for Australia and Switzerland; 1981-1993 for Austria; 1982-1997 for Portugal; 1983-1995 for Belgium; 1989-1997 for New Zealand; no data for mainland Norway.

Table A2.1. **Details on data availability** (*cont.*)

Variable	Start date	End date	Exceptions
Private credit as a share of GDP (<i>Priv credit₋₁</i>)	1971	1998	1971-1997 for Ireland and Portugal; 1971-1992 for Western Germany.
Stock market capitalisation as a share of GDP (<i>Stock cap₋₁</i>)	1977	1998	1977-1997 for Belgium and Switzerland; 1977-1991 for Western Germany; 1979-1995 for Portugal; 1982-1998 for Norway; 1984-1998 for Finland; 1986-1998 for New Zealand; 1996-1998 for Ireland.
Trade exposure (<i>Trade exp₋₁</i>)	1971	1998	

Table A2.2. **Basic Statistics**

Variables (in per cent)	Sample mean (absolute values or per cent)	Standard deviation
Y (1993 PPP \$) ¹	23 951.0	5 783.0
Sk (%)	17.11	4.46
H (years) ²	10.15	1.69
ΔLnP (%)	0.79	0.62
SdinfL ₋₁	1.51	1.27
Infl ₋₁ (%)	6.87	4.89
Gov cons ₋₁ (%)	18.46	5.11
Sk ^{gov} ₋₁ (%)	3.70	2.06
Tax ₋₁ (%)	39.62	9.19
Tax distr ₋₁ (%)	11 2.44	43.44
R&D ^{tot} (%)	1.72	0.80
BERD (%)	1.05	0.64
R&D ^{pub} (%)	0.66	0.21
Priv credit ₋₁ (%)	56.98	29.50
Stock cap ₋₁ (%)	33.79	28.86
Trade exp ₋₁ (%)	43.25	18.18

1. In 1995 US\$ (Expressed in 1993 EKS PPPS).

2. Average years of education.

BIBLIOGRAPHY

- ABEL, A.B. (1983),
“Optimal investment under uncertainty”, *American Economic Review*, No. 73 (1), pp. 228-233.
- AGELL, J., T. LINDH and H. OHLSSON (1997),
“Growth and the public sector: A critical review essay”, *European Journal of Political Economy*, No. 13, pp. 33-52.
- AGHION, P. and P. HOWITT (1992),
“A model of growth through creative destruction”, *Econometrica*, 60, pp. 323-51.
- AGHION, P. and P. HOWITT (1998),
Endogenous Growth Theory, The MIT Press, Cambridge, Mass.
- AHN, S. and P. HEMMINGS (2000),
“Policy influences on economic growth in OECD countries: an evaluation of the evidence”, *OECD Economics Department Working Papers*, No. 246.
- ALEXANDER, R.J. (1997),
“Inflation and economic growth: evidence from a growth equation”, *Applied Economics*, 29(2), pp. 233-38.
- ANDRES, J. and I. HERNANDO (1997),
“Does inflation harm economic growth? Evidence for the OECD”, *NBER Working Paper* No. 6062.
- ASCHAUER, D.A. (1989),
“Is public expenditure productive?”, *Journal of Monetary Economics*, 23, pp. 177-200.
- BALDWIN, R. (2000),
“Trade and growth: still disagreement about the relationship”, *OECD Economics Department Working Papers*, No. 264.
- BARRO, R.J. (1991),
“Economic growth in a cross section of countries”, *Quarterly Journal of Economics*, 106(2), pp.407-433, May.
- BARRO, R.J. and J.W. LEE (1996),
“International measures of schooling years and schooling quality”, *American Economic Review, Papers and Proceedings*, 32(3), pp. 363-394.
- BARRO, R.J. and X. SALA-I-MARTIN (1995),
Economic Growth, McGraw-Hill, New York.
- BASSANINI, A. and S. SCARPETTA (2002),
“Does human capital matter for growth in OECD countries? A Pooled Mean Group Approach”, *Economics Letters* 12, forthcoming.

- BASSANINI, A., S. SCARPETTA and P. HEMMINGS (2001),
 "Economic growth: the role of policies and institutions. Panel data evidence from OECD countries", *OECD Economics Department Working Papers*, No. 283.
- BECK, T., A. DEMIRGUC-KUNT and R. LEVINE (1999),
 "A new database on financial development and structure", World Bank, June, *mimeo*.
- BELSLEY, D.A., E. KUH and R.E. WELSCH (1980),
 Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, John Wiley and Sons, New York.
- BEN-DAVID, D. and A. KIMHI (2000),
 "Trade and the rate of convergence", *NBER Working Paper* No. 7642.
- BENHABIB, J. and M. SPIEGEL (1994),
 "The role of human capital in economic development: Evidence from aggregate cross-country data", *Journal of Monetary Economics* 43, pp. 143-174.
- BERTOLA, G. and R.J. CABALLERO (1994),
 "Irreversibility and aggregate investment", *Review of Economic Studies*, 61, pp. 223-246.
- BRUNO, M. and W. EASTERLY (1998),
 "Inflation crises and long-run growth", *Journal of Monetary Economics* 41 (1998), pp. 3-26.
- CAMERON, G. (1998),
 "Innovation and growth: A survey of the empirical evidence", *mimeo*.
- CASELLI, F., G. ESQUIVEL and L. FERNANDO (1996),
 "Reopening the convergence debate: a new look at cross-country growth empirics", *Journal of Economic Growth*, 1, pp. 363-389.
- CELLINI, R., M. CORTESE and N. ROSSI (1999),
 "Social catastrophes and growth", *mimeo*.
- COE, D.T. and E. HELPMAN (1995),
 "International R&D spillovers", *European Economic Review*, 39, pp. 859-87.
- DAVID, P.A., B.H. HALL, and A.A. TOOLE (1999),
 "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence", *mimeo*.
- DE LA FUENTE, A. and R. DOMENECH (2000),
 "Human capital in growth regressions, how much difference does data quality make?", CSIC, Campus de la Universidad Autonome de Barcelona, *mimeo*.
- DIXIT, A.K. and R.S. PINDYCK (1994),
Investment under uncertainty, Princeton University Press, Princeton, N.J.
- DURLAUF, S.N. and D.T. QUAH (1999),
 "The new empirics of economic growth", in J. Taylor and M. Woodford (eds), *Handbook of Macroeconomics*, North-Holland, Amsterdam.
- EDEY, M. (1994),
 "Costs and benefits from moving from low inflation to price stability", *OECD Economic Studies*, No. 23, Winter.
- ENGLANDER, S. and A. GURNEY (1994),
 "Medium-term determinants of OECD productivity", *OECD Economic Studies*, No. 22, Spring.
- EVANS, P. (1997),
 "How fast do economies converge", *The Review of Economics and Statistics*, 79(2), pp. 219-25.

- FAGERBERG, J. (1994),
 "Technology and international differences in growth rates", *Journal of Economic Literature*, 32(3), 1147-75, September.
- FÖLSTER, S. and M. HENREKSON (1998),
 "Growth and the public sector: A critique of the critics", *European Journal of Political Economy*, 15, pp. 337-58, June.
- FÖLSTER, S. and M. HENREKSON (2000),
 "Growth effects of government expenditure and taxation in rich countries", *Stockholm School of Economics, Working Paper No. 391*, June.
- FRANKEL, J.A., and D. ROMER (1999),
 "Does trade cause growth: cause and effect?", *American Economic Review*, 89, pp. 379-99.
- GROSSMAN, G.M. and E. HELPMAN (1991),
Innovation and Growth in the Global Economy, MIT Press, Cambridge, Massachusetts.
- GUELLEC, D. and B. VAN POTTELSBERGHE (1997),
 "Does government support stimulate private R&D?" *OECD Economic Studies*, No. 29, pp. 95-122.
- GUELLEC, D. and B. VAN POTTELSBERGHE (2001),
 "R&D and productivity growth: panel data analysis of 16 OECD countries", *OECD Economic Studies*, this issue.
- HARTMAN, R. (1972),
 "The effects of price and cost uncertainty on investment", *Journal of Economic Theory*, 5, pp. 285-266.
- HAUSMAN, J. (1978),
 "Specification tests in econometrics", *Econometrica*, 46, pp. 1251-1271.
- HELMS, L. (1985),
 "The effects of state and local taxes on economic growth: a time series-cross section approach", *The Review of Economics and Statistics*, 67, pp. 574-582.
- ISLAM, N. (1995),
 "Growth empirics: a panel data approach", *Quarterly Journal of Economics*, 110, pp. 1127-1170.
- KLENOW, P.J. and A. RODRÍGUEZ-CLARE (1997),
 "Economic growth: a review essay", *Journal of Monetary Economics*, 40 (1997), pp. 597-617.
- KNELLER, R., M. BLEANEY and N. GEMMELL (1998),
 "Growth, public policy and the government budget constraint: evidence from OECD countries", *University of Nottingham, School of Economics, Discussion Paper No. 98/4*.
- LEAHY, M., S. SCHICH, G. WEHINGER, F. PELGRIN and T. THORGEIRSSON (2001),
 "Contributions of financial systems to growth in OECD countries", *OECD Economics Department Working Papers*, No. 280.
- LEE, K., M.H. PESARAN and R. SMITH (1997),
 "Growth and convergence in a multi-country empirical stochastic Solow model", *Journal of Applied Econometrics*, 12, pp. 357-392.
- LICHTENBERG, F.R. (1988),
 "The private R&D investment response to federal design and technical competitions", *American Economic Review*, 78, pp. 550-559.
- LUCAS, R.E. (1988),
 "On the mechanics of economic development", *Journal of Monetary Economics*, 22.

- MADDISON, A. (2001),
"The world economy: a millennial perspective", *OECD Development Centre Studies*, Paris.
- MANKIW, G.N., D. ROMER and D.N. WEIL (1992),
"A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, 107, pp. 407-37, May.
- MOFIDI, A. and J. STONE (1989),
"Do states and local taxes affect economic growth?", *Review of Economics and Statistics*, 71, pp. 686-691.
- NADIRI, M.I. (1993),
"Innovations and technological spillovers", *NBER Working Paper No. 4423*.
- OECD (1998),
Science and Technology Outlook, Paris.
- OECD (2001),
Education at a Glance, Paris.
- PARK, W.G. (1995),
"International R&D spillovers and OECD economic growth", *Economic Inquiry*, Vol. XXXIII, October.
- PESARAN, M.H. and R. SMITH (1995),
"Estimating long-run relationships from dynamic heterogeneous panels", *Journal of Econometrics*, 68, (1), pp. 79-113.
- PESARAN, M.H., Y. SHIN and R. SMITH (1999),
"Pooled mean group estimation of dynamic heterogeneous panels", *Journal of the American Statistical Association*, 94, pp. 621-634.
- PSACHAROPOULOS, G. (1994),
"Returns to investment in education: A global update", *World Development* 22(9), pp. 1325-1343.
- REBELO, S. (1991),
"Long-run policy analysis and long-run growth", *Journal of Political Economy*, 99 (3), pp. 500-512.
- ROMER, P.M. (1986),
"Increasing returns and long-run growth", *Journal of Political Economy*, 94 (5), pp. 1002-1037.
- ROMER, P.M. (1990),
"Endogenous technological change", *Journal of Political Economy* 98(5) Part 2, pp. 71-102.
- SACHS, F. and A. WARNER (1995),
"Economic reform and the process of global integration", *Brookings Papers on Economic Activity*, 1995, (1), pp. 1-118.
- TEMPLE, J. (1999),
"The new growth evidence", *Journal of Economic Literature*, March, 37(1), pp. 112-156.
- UZAWA, H. (1965),
"Optimum technical change in an aggregative model of economic growth", *International Economic Review* 6, pp. 18-31.
- YOUNG, A. (1991),
"Learning by doing and the dynamic effects of international trade", *Quarterly Journal of Economics*, 106(2), pp. 369-406.

GROWTH EFFECTS OF EDUCATION AND SOCIAL CAPITAL IN THE OECD COUNTRIES

Jonathan Temple

TABLE OF CONTENTS

Introduction	58
The growth effects of education: theory.....	59
The growth effects of education: evidence.....	62
Evidence from labour economics.....	63
Growth accounting.....	68
Evidence from growth regressions.....	72
Human capital externalities.....	78
Wider benefits of education	80
A tentative summary of the evidence.....	80
Social capital and growth	81
What is social capital?.....	83
Empirical evidence	84
The future for social capital research.....	88
Summary and conclusions	89
Bibliography	96

Jonathan Temple is from the Department of Economics at the University of Bristol. This article was prepared for the OECD. The views expressed here do not necessarily represent those of the OECD or its Member governments. The author is grateful to Gavin Cameron, Damon Clark, Martine Durand, Jørgen Elmeskov, Tom Healy, John Martin, Mark Pearson and Dirk Pilat for very helpful comments on an earlier draft.

INTRODUCTION

Public and private expenditure on educational institutions accounts for just over 6 per cent of the collective GDP of the OECD Member countries, or roughly \$1 550 billion each year.¹ This figure understates the true opportunity cost of educational investments, since it does not take into account forgone earnings. Overall, it should be clear that the provision of education represents a major commitment of resources within the OECD, and so measuring the associated welfare benefits is an important task.

One aim of this survey is to examine the available evidence on the benefits of education in developed countries. The main focus is restricted to the effects of education on labour productivity, a topic for which there is a considerable body of evidence, admittedly indirect. I will draw on research from two fields in particular: labour economics, and cross-country empirical work on economic growth. An underlying argument will be that, although the labour economics literature does an impressive job of measuring the private returns to education, it remains the case that macroeconomic studies have a complementary role to play.

The emphasis throughout is very much on education, rather than on any broader concept of human capital. The chief omission is any consideration of training. This does not reflect my view of its relative significance, but rather the focus of the present survey on cross-country evidence. The nature of training varies considerably across countries, and in the manufacturing sector is tightly connected to production strategies (Broadberry and Wagner, 1996). It is difficult to capture these differences in ways that lend themselves to empirical modelling. This means that, in explaining productivity differences across OECD countries, the cross-country evidence has little to say about the role of training, despite its potential importance.² This is one area in which answers should be sought from labour economics and detailed comparisons of practices in individual countries, rather than from the cross-country empirical work reviewed here.

A second theme of the survey is the relation between growth and what has come to be known as “social capital”. It is difficult to arrive at a precise definition of this term, and I will discuss this issue in more detail later on. For now, it can be thought of as capturing such things as the extent of trustworthiness, social norms, and participation in networks and associations. In the last few years, some prominent academics and commentators have argued that these qualities of societies

are potentially valuable not only in themselves, but also because they make a contribution to economic success. This is another area in which cross-country evidence may have something worthwhile to contribute, and later in the paper, I will review the small but growing literature on the correlations between measures of social capital and economic performance.

Empirical work on social capital and growth is a very recent development, and with this in mind, I devote the majority of the survey to research on education and growth. The second section provides the theoretical background, and shows that recent models provide some good reasons for seeing education as a central determinant of economic growth. The third section turns to the empirical evidence. It starts with a brief account of research in labour economics, an essential step in understanding where the cross-country evidence may be relatively useful. The rest of the section, perhaps the heart of the survey, covers evidence from growth accounting and growth regressions, recent attempts to measure externalities to education, and some work on wider benefits.

The second part of the paper turns to social capital and growth. The fourth section discusses the definition of social capital, reviews the macroeconomic evidence on its growth effects, and briefly discusses the prospects for further research in this area. The final section rounds off with some possible conclusions.

THE GROWTH EFFECTS OF EDUCATION: THEORY

The aim of this section is to investigate whether formal models shed any light on the claim that education plays a central role in growth.³ Can the possible role of education be given a secure foundation in terms of economic theory? How plausible are the necessary assumptions? Do the models capture the growth effects of education, as it is generally defined and understood, or of something else?

One of the most prominent and influential contributions is that of Lucas (1988), which is in turn related to previous work by Uzawa (1965). In these models, the level of output is a function of the stock of human capital. In the long run, sustained growth is only possible if human capital can grow without bound. This makes it difficult to interpret the Uzawa-Lucas conception of human capital in terms of the variables traditionally used to measure educational attainment, such as years of schooling. Their use of the term “human capital” seems more closely related to knowledge, rather than to skills acquired through education.

One way to relate the Uzawa-Lucas model to the data is to suggest that the quality of education could be increasing over time (Bils and Klenow, 2000). In this view, the knowledge imparted to school children in the year 2000 is superior to the knowledge that would have been imparted in 1950 or 1900, and will make a

greater difference to their productivity in later employment. Even if average educational attainment is constant over time, the stock of human capital could be increasing in a way that drives rising levels of output.⁴

Yet this argument runs into difficulties, even at the level of university education. There may be some degree courses in which the knowledge imparted currently has a greater effect on productivity than before (medicine, computer science, perhaps economics) but there are other, less vocational qualifications for which this argument is less convincing. At the level of primary and secondary schooling, with their focus on basic skills such as literacy and numeracy, the idea that increases in the quality of schooling drive sustained growth seems even harder to support. Finally, note that these models are typically silent on exactly how the increase in the quality of schooling is brought about: individuals can raise the stock of human capital, or knowledge, simply by allocating some of their time to its accumulation.

An alternative class of models places more emphasis on modelling the incentives that firms have to generate new ideas. Endogenous growth models based on the analysis of research and development, notably the landmark contribution of Romer (1990), yield the result that the steady-state growth rate partly depends on the *level* of human capital. The underlying assumption is that human capital is a key input in the production of new ideas. In contrast with the Uzawa-Lucas framework, this opens up the possibility that even a one-off increase in the stock of human capital will raise the growth rate indefinitely. Indeed, in many endogenous growth models, human capital must be above a threshold level for any innovation to take place at all.

In practice, the generality of these results, and the contrast with the Uzawa-Lucas model, should not be overdrawn. The Uzawa-Lucas framework can be seen as a model of knowledge accumulation in a similar spirit to that of Romer, but easier to analyse; and restrictive assumptions are needed to yield the Romer result that the long-run growth rate depends on the level of human capital (Jones, 1995). But even under more general assumptions, a rise in the level of human capital is likely to be associated with a potentially substantial rise in the level of output, brought about through a transitional increase in growth rates.

In most endogenous growth models based on research and development, the stock of human capital is taken to be exogenously determined. More recent papers, notably Acemoglu (1997) and Redding (1996), have relaxed this assumption, and considered what happens when individuals can choose to make investments in education or training, while firms make investments in R&D. For some parameter values, multiple equilibria are possible, since the incentives of workers to invest in human capital, and those of firms to invest in R&D, are interdependent. This provides a way of formalising earlier ideas about the possible existence

of a “low-skill, low-quality trap” in which low skill levels and slow rates of innovation reflect a co-ordination failure (Finegold and Soskice, 1988). The models suggest that, at the aggregate level, greater investment in education or training might raise expenditure on R&D, and *vice versa*.

Another interesting aspect of recent growth models is their suggestion that individuals may under-invest in education. Rustichini and Schmitz (1991) examine this argument in some detail. They present a model in which individuals divide their time between production, original research, and the acquisition of knowledge. Each individual knows that acquiring knowledge (through education) will raise their productivity in subsequent research, but since they do not fully capture the benefits of research, they will tend to spend too little time acquiring knowledge relative to the socially optimal outcome.⁵ Rustichini and Schmitz calibrate a simple model, and find that although policy intervention has only small effects on the allocation of time to education, it can have a substantial effect on the growth rate.⁶

More recently, Romer (2000) has pointed out that models of growth driven by R&D should potentially inform education policy. He notes that, in the models reviewed above, growth is determined by the quantity of inputs used in R&D, not simply expenditure upon it. One reason this point matters is that incentives to encourage R&D, such as tax credits, may be ineffective unless they encourage a greater number of scientists and engineers to work towards developing new ideas. To illustrate this, consider a very simple model, in which a fixed supply of scientists only work in R&D and are the only input to the research process. Then an increase in R&D spending will simply raise the wages of scientists, with no effect on the number of researchers engaged in R&D, or the growth rate.

In a more general and realistic model, there will be some effect of greater R&D spending on total research inputs and therefore growth.⁷ To create a large effect, higher wages for scientists should encourage more individuals to train as scientists. This requires some flexibility on the part of the education system, and in the provision of relevant information to potential students. So the effectiveness of direct subsidies or tax credits for R&D may be enhanced by complementary education policies, aimed at improving or subsidising the supply of research inputs, rather than simply the demand for them.

In summary, the models of the new growth theory are important for several reasons. First, they see human capital as an important input in the creation of new ideas, and this mechanism provides a relatively appealing justification for viewing education as a central determinant of growth rates, even over long time intervals. Second, they sometimes yield the result that the *laissez-faire* outcome delivers slower growth than is socially optimal. Third, the models suggest that policymakers wishing to raise the level of output have several options: not just direct

support for R&D – which may be difficult to implement and monitor – but also subsidies to certain kinds of education, perhaps especially those which could lead to later work in research and development.

THE GROWTH EFFECTS OF EDUCATION: EVIDENCE

As we have seen, theoretical models imply that, in searching for the determinants of growth, policy on education is one of the first places to look. In this section, I will turn to the attempts of economists to quantify education's importance. The main focus will be on the macroeconomic evidence: the body of research which measures, or attempts to measure, the productivity benefits of education using the variation in educational attainment and growth rates across countries.⁸

It would be a mistake, however, to review this evidence without first discussing the work on education and earnings by labour economists. If education affects productivity directly, this tends to imply an observable relationship between an individual's education and their earnings. The evidence for these effects is the best established in the literature and an understanding of its strengths and weaknesses helps place the cross-country evidence in context. This will clarify the areas in which the macro approach may have something worthwhile to contribute, and also point to the areas in which micro evidence is more likely to be fruitful.

With this in mind, the sections below review: studies of the effects of education based on earnings surveys; growth accounting; the evidence from cross-country regressions; recent work on externalities to human capital; the wider benefits of education; and finally attempt to tie together the various pieces of evidence.

The review points out that each approach to measuring the productivity effects of education has its own important weaknesses and areas of uncertainty. Yet taken together, the various methods tend to agree in pointing to quite substantial effects. As a result, it would be difficult to use the available evidence to construct a case that education is currently over-provided in the OECD as a whole, and perhaps even harder if one acknowledged the wider benefits discussed below.

Broadly speaking, this work might also justify an expansion of educational provision in some countries, especially those where current policies imply relatively low levels of attainment in future years. A full analysis of policy questions, however, would need both to acknowledge the potential importance of training, and to investigate how a given quantity of educational spending is best allocated; these topics are beyond the scope of the present review.⁹

Before turning to the various strands of evidence in more detail, it may be helpful to clarify the concepts of productivity that the different approaches have

in mind. At the level of individuals, output per worker hour seems the most relevant measure of productivity, not least because one benefit of an increase in hourly productivity may be that individuals choose to work fewer hours. In examining productivity differences across countries, however, there are sometimes disadvantages in using output per worker hour as the basis for comparison. This measure of productivity is affected by labour force participation rates, and other aspects of labour market institutions. Further discussion and some recent evidence can be found in Scarpetta *et al.* (2000).

It is also worth pointing out that, for some purposes, policy-makers are interested in output per worker and output per head, as well as output per worker hour. Education may also have indirect effects on these variables, not simply through hourly productivity. For example, education is often thought to affect labour force participation, particularly that of women.¹⁰ It may also affect the non-monetary benefits associated with work and leisure, and so affect working hours. Since cross-country empirical work is typically based on output per capita or output per worker, it will tend to conflate these effects with the direct impact of education on labour productivity that labour economists have sought to quantify.

Evidence from labour economics

This section reviews evidence from labour economics. Rather than attempt to provide a summary of a vast empirical literature, the emphasis will be on how the conventional findings should be interpreted, and to what extent we can infer genuine effects of education on productivity.¹¹

Researchers in this field typically study the link between education and productivity using survey data on the earnings and characteristics of large numbers of individuals. The techniques used to analyse these data have become increasingly sophisticated, and we will see that evidence from “natural experiments” provides measures of the private return to education that are probably quite accurate. There is much greater disagreement on the extent to which labour economists have identified the *social* return to education.¹² For example, educational qualifications may be valued in the labour market because they act as a signal of ability. As a result, private returns to schooling may be high even if education has no effect on productivity. This argument will be discussed further below.

In analysing the private return, the standard empirical approach is to explain the variation in earnings across individuals using regressions, where the explanatory variables include years of schooling, either age or a simple proxy for labour market experience, and other characteristics. The most popular specification draws heavily on the work of Mincer (1974), and earlier contributions on “human

capital earnings functions". The starting point is typically a specification that looks something like this:

$$\ln w = \alpha + \beta_0 S + \beta_1 E + \beta_2 E^2 \quad (1)$$

which relates the natural logarithm of wages (w) to years of schooling (S) and a proxy for labour market experience (E). Under some assumptions, and given the semi-logarithmic formulation, the coefficient on schooling can be interpreted as the private return to education. Empirical estimates of the private return typically have a relatively small standard error and lie somewhere between 5 per cent and 15 per cent, depending on the time and country. If workers are paid their marginal product, these educational wage differentials may also tell us something useful about the effect of education on productivity.

The evidence that earnings are positively associated with schooling is robust and uncontroversial; the obvious difficulty lies in giving this association a causal interpretation. One of the most easily understood problems is that, through lack of suitable data, the regressions inevitably omit some important variables that are likely to be correlated with both schooling and earnings. Family background and traits such as innate ability or determination are notable examples.

The basic problem, from the econometrician's point of view, is that the group of people with a relatively advanced level of educational attainment is not a random selection from the population as a whole. For example, if more able individuals have relatively high earnings regardless of extra education, and also choose to spend more time in school, then the estimated return to schooling overstates the effect of education on productivity. If ability is not observed by employers, then the regression estimate may still capture the private return to schooling, but it will not capture the social return that is ultimately our main interest.

Unfortunately the problems do not stop there. It seems probable that the costs and benefits of education vary across individuals, perhaps substantially. Indeed, this is likely to be the principal cause of the variation in completed schooling that the econometrician uses to identify the effects of education. The heterogeneity will typically mean that the private returns to education vary across individuals. In the unlikely case where the returns vary independently of the explanatory variables, the regressions should still recover an unbiased estimate of the average return. More generally, however, the heterogeneity problem will lead to biased estimates.

The recent focus of the literature on education has been on identifying natural experiments, in the hope that these will allow stronger claims about causality to be made. Researchers look for situations in which the level of schooling varies across individuals for reasons that are likely to be independent of the unobserved characteristics of those individuals (ability, determination, and so on).

The idea is best explained by means of an example. A good starting point is one of the most influential papers, by Angrist and Krueger (1991). The paper starts from the observation that, when it is compulsory to stay in school until a certain age, individuals born earlier in the calendar year will reach the legal minimum age for school-leavers at an earlier stage in their education. As a direct result, there is likely to be a correlation between an individual's quarter of birth and their length of schooling. The correlation means that quarter of birth can potentially be used to identify exogenous variation in schooling – that is, variation independent of unobserved characteristics like ability or determination. In econometric terms, quarter of birth can be used as an instrument for schooling, under the maintained assumption that personal characteristics other than schooling are independent of quarter of birth. Somewhat surprisingly, Angrist and Krueger find that the instrumental variable estimates of the return to schooling are similar to the least squares estimates, supporting the idea that conventional estimates are reasonably accurate.¹³

Another much-discussed natural experiment is provided by identical twins who have different levels of schooling. Given that such twins have the same genes, and will usually share the same family background, the wage differential between twins with different years of schooling may provide useful information on the productivity effect of education. Finally, other natural experiments are provided by the possible connection between the geographical proximity of colleges to individuals, and their choice of schooling (see Card, 1999).

Research of this kind has considerably strengthened the case for productivity effects of education, but even these studies retain an important weakness. It has long been understood that the private return to education may be a poor guide to the social return. The theoretical work of Spence (1973) indicated that educational attainment may be valued by employers mainly because it acts as a signal of innate ability, and not because it has an effect on productivity.

Models of signalling start from the observation that individuals have traits which employers value but do not fully observe at the time of hiring (ability, determination, and so on). If there is a systematic association between these traits and the costs and benefits of education, this may lead to an equilibrium in which high-ability individuals stay in school for longer because this decision signals their ability to employers. This argument provides a plausible reason for a correlation between ability and years of schooling, and suggests that earnings may be correlated with schooling even if schooling has no effect on productivity.

Few doubt that signalling plays some role in explaining educational wage differentials, but its overall importance remains controversial. Weiss (1995) and Quiggin (1999) provide very different perspectives on the theoretical generality and empirical validity of signalling models. There are two main arguments against such models, which note the implications of the assumption that education has no

effect on productivity. First, given the wage premium earned by those with more years of schooling, employers would probably have strong incentives to conduct their own tests of ability and other characteristics, and use this direct information rather than the somewhat indirect signal provided by the schooling decision. This view is supported by evidence that measured performance in school and universities is correlated quite strongly with the outcomes of tests carried out at an earlier stage (see Quiggin, 1999 for references). Yet the argument is not conclusive, mainly because employers may not be able to appropriate the returns to acquiring more information about their employees; other firms could bid away those workers found to have higher ability (Stiglitz, 1975).

The second argument is that, if education does not affect productivity, one would expect to see the educational wage differential decline with job tenure, as employers acquire direct knowledge of the characteristics of their employees. This does not seem to be observed in the data, although this question has not received the sustained attention it probably deserves.

More generally, there is clearly room to develop and test signalling arguments in more detail. This is important not least because, as Weiss (1995) has pointed out, even the results of natural experiments are not necessarily inconsistent with the signalling view of education. To see this, recall that employers may use years of schooling to gain information about characteristics that are not observed at the time of hiring. The results from the Angrist and Krueger quarter-of-birth study and the work on twins can easily be interpreted in terms of these signalling effects, and so one could still defend even the extreme view that productivity is entirely independent of education.

For now, let us assume that employers fully observe all relevant characteristics, and hence do not infer any information about them from schooling decisions. Even in this case, as Card notes, not much is presently known about the mechanisms by which education might contribute to higher wages. The simplest interpretation of the evidence from earnings functions is that more educated individuals are more productive, whatever their chosen occupation. In practice, a college degree is unlikely to make one a noticeably better postman or road sweeper.

Education's role may be to equip workers for the task of working with more advanced technologies, for providing a higher quality of service, or for "learning by doing" in the course of employment. Understanding the mechanisms could be important, and will have implications for the interpretation of earnings functions. For instance, more educated workers may have better access to those jobs in which workers share some of the rents earned by imperfectly competitive firms. If mechanisms like this are at work, there would again be less reason to believe that

the observed correlation between schooling and earnings represents solely a direct productivity effect.

There are other ways in which private and social returns could differ. In some countries, especially poorer ones, the public sector is a major employer of the well-educated. As Pritchett (1996) emphasises, the assumption that wage differentials reflect differences in marginal products is much harder to sustain in this context. If educational credentials are used as a means of determining access to rationed high-paying jobs in the public sector, estimated earnings functions may detect an effect of education even when it has little or no effect on productivity.

The general problem is that estimates of earnings functions capture, at best, the private return to education, yet it is the social return which is of most interest to policy-makers. The two may diverge for a number of reasons, including the possibility that education acts mainly as a signalling device. The arguments discussed above imply that the social return to education is less than the private return, and as we have seen, even just a lower bound on the social returns is difficult to establish.

There are also some reasons to believe that the social return to education could exceed the private return. It is plausible that individuals do not fully capture some of the benefits to society of their schooling, and I will review some of the empirical evidence on externalities and wider benefits below. Another important argument is that educational provision may play a valuable role in allowing a more efficient matching between workers and jobs than would otherwise be possible (Arrow, 1973, Stiglitz, 1975). In other words, even if education does act mainly as a signal, there should not be a presumption that education is therefore socially wasteful.

In summary, there is an ingenious and persuasive body of work which supports the view that private returns to schooling are quite high. Card (1999) concludes that the average marginal return to education is unlikely to be far below the standard regression estimates. The view that this private return originates in a genuine productivity effect is far from universally accepted, however. As Weiss (1995) has argued, even the most recent results can be interpreted as the outcome of signalling effects.

This suggests two lines of enquiry that might be particularly fruitful. The first is further theoretical examination (and perhaps calibration) of signalling models, with a particular focus on the extent to which they can incorporate the direct productivity effects envisaged in the traditional theory of human capital. Second, more evidence on the extent to which educational wage differentials evolve with job tenure could be of great interest in advancing the debate.

Growth accounting

As we have seen, the labour economics literature provides a wealth of evidence on the private returns to schooling. It is necessarily silent, however, on the contribution of education relative to other sources of aggregate growth. Making assumptions similar to those of labour economists, researchers in the growth accounting tradition have set about the complex task of evaluating the overall growth contribution of changing educational attainment. This section will describe the method and review the available evidence.

Growth accounting essentially divides output growth into a component that can be explained by input growth, and a “residual” which captures efficiency change, partly reflecting changes in technology.¹⁴ In explaining the change in output, the change in the quantity of each input is weighted by its marginal product, proxied by its market reward. This principle can be extended to any number of inputs, and where sufficiently detailed data are available, it is possible to disaggregate the labour force into various categories, where each type of worker is weighted by the average wage of that type.

For instance, in analysing the contribution of changes in educational attainment, the researcher disaggregates the labour force by level of schooling, and often by other available characteristics such as age and gender. Changes in the number of employees at each level of schooling are then weighted by their marginal products, proxied by the mean income associated with each schooling level, to give the overall change in an index of “effective” or quality-adjusted labour. This ultimately allows the researcher to quantify the proportion of output growth that can be directly attributed to increases in educational attainment.

Griliches (1997) provides a brief but useful survey of this literature, and points out the two major assumptions, both of which will have a familiar ring to readers of the previous section. First, it is assumed that differences in observed market rewards correspond reasonably closely to differences in marginal products. Secondly, the calculations assume that differences in market rewards across schooling levels originate in schooling, and not in other factors such as native ability or family background that may be correlated with schooling.

The advantage of the first assumption, that market rewards correspond to marginal products, is that it allows the growth accountant to obtain theory-consistent weights using the available data, at least under the assumptions of constant returns to scale and perfect competition. Less restrictive frameworks are possible, but will generally tend to require additional and perhaps controversial assumptions about parameters. It should also be clear that conventional growth accounting will not shed any light on the possible contribution of externalities. This is a major limitation, and an important motivation for the cross-country empirical studies that will be considered further below.

What of the second assumption, that differences in wages originate in schooling? The danger here can be seen from considering an extreme scenario, in which education has absolutely no effect on an individual's productivity, but more able individuals both spend longer in school and earn more while in employment. This scenario clearly implies that educational attainment and earnings are positively correlated. Now consider an exogenous increase in the proportion of individuals with the highest level of education: since the index of labour quality weights the numbers in each education class by the mean income of that class, the index must increase. As a result, the growth accountant will attribute some portion of growth to educational improvement, even though education plays no role in productivity gains. This implies that, as in the labour economics literature, growth accounting can give us some insight into the productivity contribution of education, but the answers are by no means complete or conclusive.

Before describing the results of growth accounting exercises in more detail, it is important to clarify the connection between changes in educational provision and the measured effects. An expansion in provision typically affects only younger cohorts, and so has long-lived effects on educational attainment in the labour force as a whole. Average attainment will continue to increase for some time as older, less educated cohorts retire from employment and are replaced by the more highly qualified. When using growth accounting methods, it is these long-lived effects that are quantified, and one should bear this in mind when interpreting specific findings. The practical implication is that results for recent years are driven by changes in educational provision much further back in time.

Growth accounting exercises vary widely in the extent to which they disaggregate labour input. Nearly all the studies which carry out a detailed disaggregation by level of schooling are restricted to the United States; the classic study is Jorgenson, Gollop and Fraumeni (1987). For the period 1948-79, they find that growth in labour input has contributed about a third of growth in aggregate value added, where the measure of labour input takes into account both hours worked and the quality of labour. Changes in their aggregate index of labour quality are based on changes in the composition of total hours worked by age, sex, education, employment class and occupation. They find that a favourable shift in labour quality is responsible for about a tenth of the growth in value added, or about a fifth of the productivity residual that remains after accounting for the contribution of growth in physical capital (see their Table 9.5).

In interpreting the results of Jorgenson, Gollop and Fraumeni, it is important to note that some of the compositional shifts within the labour force have a *negative* effect on the index of labour quality over the 1948-79 period, which partly offset the benefits of improvements in educational attainment. As previously noted, the calculation of the labour quality index assumes that differences in market rewards reflect genuine differences in marginal products. One consequence is that the

increasing entry of women and young workers into the labour market, mainly into low-paying jobs, has a negative effect on the aggregate index of labour quality.

Over the 1948-79 period, the negative effect on the index of labour quality is more than offset by positive changes in the composition of the labour force by educational attainment and occupation. One implication is that the latter effects are likely to be responsible for more than a fifth of the productivity residual, since the favourable shift in labour quality would have been larger in the absence of the change in composition by age and sex.

In reviewing the evidence as a whole, Griliches (1997) writes that increases in educational attainment seem to have accounted for perhaps a third of the productivity residual in the US over the post-war period. In the 1950s and 1960s, this would correspond to an effect on the annual growth rate of aggregate output of around 0.5 percentage points; during the 1970s productivity slowdown the effect of educational improvement will have been lower, perhaps raising the growth rate by 0.2 or 0.3 percentage points. As discussed above, these effects are inherently transitional ones, driven by long-standing changes in education policy that shift the educational composition of the labour force towards a new steady state.

Other OECD Member countries have also seen important changes in educational attainment in the last fifty years. Englander and Gurney (1994a) note that tertiary education in particular has expanded rapidly in many OECD countries since 1960. More detail on the general trends can be found in OECD (1998, Chapter 2) and OECD (2000a, 2000b).

As yet, however, there are few studies that cover recent experience of other OECD Member countries in the same degree of detail as Jorgenson, Gollop and Fraumeni.¹⁵ The best known studies covering a number of developed countries for recent years are those of Maddison (1987, 1991). Maddison (1991, p. 138) argues that the 20th century saw a fairly steady improvement in educational attainment for the six countries he considers (France, West Germany, Japan, the Netherlands, the UK and the US). One implication is that changing trends in educational attainment are unlikely to provide a satisfactory explanation for the transition from Europe's "Golden Age" of rapid growth (1950-73) to the productivity slowdown after 1973.

For these six countries, Maddison estimates the growth impact of changes in educational attainment by disaggregating the labour force into those with primary, secondary and higher qualifications. He then combines these three different types of labour using weights that are the same across countries and over time. In selecting the weights, he follows Denison (1967) in assuming that observed educational wage differentials overstate the contribution of education to productivity, for the reasons discussed in the previous section. Inevitably, the adjustments made are somewhat arbitrary, but they do serve to highlight the uncertainty inherent in the general approach. The other point to note is that, because of these adjustments,

the estimates of Denison and Maddison are not directly comparable with those of other studies.

With all this in mind, we can turn to Maddison's results on the contribution of increases in labour quality to output growth in France, West Germany, Japan, the Netherlands, the UK and the US. His figures suggest that changes in the quality of the labour force typically added between 0.1 and 0.5 percentage points to annual growth rates between 1950 and 1984 (his Table 20). The Maddison index of labour quality takes into account changes in the male/female composition (though not age composition) of the labour force, as well as changes in educational attainment. In countries where the proportion of women in the labour force has noticeably risen, such as the UK and the US, the contribution of education to growth will be slightly higher than the reported contribution of growth in labour quality.

More recent studies include that of Jorgenson and Yip (1999), who have recently carried out a detailed growth accounting exercise for the G7, and present estimates of growth in labour quality for 1960-95 (their Table 7). These estimates suggest that labour quality has grown particularly quickly in Japan, and to a lesser extent, relatively quickly in France and the US. The Jorgenson-Yip disaggregation of the labour force is slightly finer than that adopted by Maddison, and this makes it harder to assess the role of education within changes in the overall index of labour quality.

A useful survey by Englander and Gurney (1994b) draws together the results of a number of studies for the G7, although some of this evidence is based on regressions rather than growth accounting. Their summary suggests that for the 1960s to 1980s the growth of human capital (sometimes including demographic effects, of the kind discussed above) typically accounts for a tenth to a fifth of growth in total output. For those countries, like the US, where there has been a rapid increase in employment, these figures probably slightly understate the proportion of growth in output per worker that can be attributed to rising attainment.

Another OECD country for which recent and detailed growth accounting results are available is Korea. The most influential contribution is that of Young (1995), who examines and compares the growth performance of four East Asian economies. For the purpose of the present survey, the case of Korea is particularly interesting in that the country has seen a dramatic increase in the educational attainment of the labour force. Between 1966 and 1990, the proportion of the working population with secondary level education or higher roughly trebled, from 27 per cent to 75 per cent. Yet this dramatic expansion does not translate into an equally dramatic effect on the growth rate, at least under the assumptions of growth accounting. For each of the four economies he considers, Young finds that

the improving educational attainment of the workforce raised the annual growth rate of effective labour input by about 1 percentage point (Young 1995, p. 645).

I end this section by noting an essential qualification to the results above, and a possible extension to the conventional approach. All growth accounting results require careful interpretation, because the approach does not tell us everything we need to know about the relevant counterfactual.¹⁶ As an example, consider a claim that X percentage points of growth in a given country is due to a change in the quality of the labour force. This does not imply that, in the absence of the change in labour force quality, the growth rate of output would have been precisely X percentage points lower. The problem is that educational attainment may have other, indirect effects on output through labour force participation, investment, and even R&D and the growth of total factor productivity. Growth accounting does not capture these indirect effects, and so gives only a partial picture of the overall importance to growth of variables like education.

Finally, although accounting decompositions are usually applied to growth rates, the same ideas can be applied to decompositions of output levels. We can then ask questions such as: to what extent do differences in educational attainment explain the variation in GDP per capita across OECD Member countries? Research applying such ideas is just starting to emerge, and Woessmann (2000) discusses the approach in more detail. Working on the assumption that measured private returns to schooling are capturing a genuine productivity effect of education, his analysis suggests that differences in educational attainment account for most of the output variation across OECD Member countries.

Evidence from growth regressions

Although growth accounting exercises are informative and often useful, it is clear that they are not a complete substitute for other forms of investigation, given the necessary assumptions. Griliches (1997, p. S333) writes that “the main, and possibly only, approach to testing the productivity of schooling directly is to include it as a separate variable in an estimated production function”. Such estimates could be at the level of firms or regions, but much of the evidence uses the variation in education across countries, and it is to such estimates that I turn next.

The key attraction of growth regressions is that they provide a way of testing directly for productivity effects of education. This has sometimes been noted in the theoretical literature: Arrow (1973, p. 215) pointed out that the use of macro-economic evidence would be one way of testing the signalling arguments, although he also expressed doubts about the likely reliability of such an approach.

Recent work has led to a better understanding of precisely when and where scepticism might be justified. In what follows, I will review the most important

problems associated with measuring growth effects of education at the macroeconomic level. An underlying theme is that, despite these problems, there are some grounds for optimism that this research can yet make a worthwhile contribution.

This may seem surprising, given that several well-known papers in this field take very different views on the importance of education. The argument below is that a more coherent story is gradually starting to appear, in which the results of cross-country studies increasingly look consistent with the effects identified by labour economists, and which can also explain why some earlier studies failed to detect any significant effect of education using aggregate data.

In the early work in this field, some of the estimated effects looked too large to be credible, as will be discussed further below. One of the best known and most influential contributions to the empirical growth literature is that of Mankiw, Romer and Weil (1992). If taken at face value, their parameter estimates for an OECD sample imply that if human capital investment (as a share of GDP) is increased by a tenth, output per worker will rise by 6 per cent; if investment in human capital is doubled, output per worker will eventually rise by about 50 per cent.¹⁷

Results of this kind are often perceived as rather dubious, since all growth regressions share a number of important statistical problems (Temple 1999a). In the present context, one drawback of most regression studies is their focus on a large sample that includes less developed countries as well as OECD Member countries. One should clearly be rather wary about drawing conclusions for OECD policy based on samples that are often dominated by developing countries. I will usually concentrate on the few studies that include separate estimates of regressions restricted to either OECD Member countries or rich countries.

Researchers have generally used one of two specifications in modelling growth and education. In the first and most common specification, the researcher chooses to regress growth on control variables and the initial level of an education measure, such as the secondary school enrolment rate or (preferably) average years of schooling. The underlying idea is that the stock of human capital could affect subsequent growth in a variety of ways, notably by influencing a country's ability to adopt technology from abroad.¹⁸ The second specification uses the change in educational attainment, not its level, to explain output growth; this approach will be discussed further below.

It has sometimes been argued that in practice, one might expect a negative effect to emerge from regressions based on the level of education, and this potential ambiguity could make the results hard to interpret (Topel, 1999). For example, countries with a low level of education may also be relatively far behind technological leaders like the US, and therefore have more opportunities to catch up and grow quickly. Arguments of this kind are not yet altogether convincing. In this spe-

cific case, one should note that growth regressions usually control for initial output per worker, and this will incorporate a large part of the catch-up effects associated with technological backwardness.

When researchers relate growth to the initial level of education, they typically find an effect of schooling that is both large and precisely estimated, at least when initial output per worker is also included as an explanatory variable (as in Barro, 1991). Yet it is not clear that these results are applicable to OECD Member countries. In an interesting exercise, Englander and Gurney (1994a) re-estimate growth regressions based on four influential papers, including Barro (1991), but restricting the sample to the OECD. Three of the four sets of regressions include human capital variables, typically primary and secondary school enrolment rates.¹⁹ These variables turn out to perform relatively well, but are still far from robust. In further work, it may be valuable to repeat this exercise, drawing on more recent data sets that allow one to use average years of schooling rather than enrolment rates.

A more recent paper that includes results specific to OECD samples is Gemmell (1996). He emphasises the problems of using enrolment rates, and constructs alternative measures of human capital based on attainment at the primary, secondary and tertiary levels. For a sample of 21 OECD countries, he finds a correlation between the number of people with tertiary qualifications and subsequent growth. He also finds some evidence that investment in OECD countries is positively correlated with the extent of secondary schooling in the labour force.

One drawback of most cross-country work is the likelihood of important differences in the nature and quality of schooling across countries, which could undermine the usefulness of international comparisons. Even such things as the length of the school year can show a surprising degree of variation across countries. An alternative data set, which may overcome these problems to some extent, has been introduced by Hanushek and Kimko (2000). They measure educational attainment using scores in international tests of cognitive skills in maths and science. Their results support the idea that education has a substantial effect on growth rates, although the applicability to OECD countries is not clear.

The lack of studies with direct relevance to the OECD is not the only dilemma for those who wish to draw policy conclusions for developed countries. The rather atheoretic approach of the macroeconomic literature on education and growth has attracted a certain amount of criticism. One argument, used by Topel (1999), is that the measured effect of the initial level of human capital is often too large to be credible. The underlying assumption here is that education's effects are mostly accounted for by examining the correlation between education and earnings at the individual level. The models of the new growth theory, reviewed above, indicate that this view of education's role is perhaps too narrow.

Nevertheless, the perspective of labour economics remains of interest. Starting with Pritchett (1996), researchers have noted the implications of traditional earnings functions for analyses at the cross-country level. If an individual's education contributes directly to their productivity, in the manner envisaged by labour economists, we should expect to observe a correlation between the change in output per worker and the *change* in average educational attainment, at least after controlling for other variables. Furthermore, it should be possible to detect this effect regardless of whether or not the initial *level* of educational attainment determines growth.²⁰

This argument has shifted the focus of research towards regressions that relate growth to the change in educational attainment, rather than its level. Several well-known studies have found the correlation to be surprisingly weak; Benhabib and Spiegel (1994) and Pritchett (1996) both come to this conclusion for a large sample of countries.²¹ Benhabib and Spiegel do find a statistically significant correlation between the level of educational attainment and growth for the wealthiest third of the sample (their Table 5, model 2) but no connection between the change in attainment and growth in a larger sample. One reason for this may be the effect of outliers, as discussed in Temple (1999b, 2001).

There are a number of other problems that dictate caution in reading these papers. One is the specification chosen for the relation between years of schooling and output. The specification adopted by Benhabib and Spiegel implicitly assumes that the returns to an extra year of schooling are much higher at low levels of schooling than high levels. As Topel (1999) points out, this runs contrary to the standard semi-logarithmic formulation for earnings functions, which in its simplest form assumes that the returns to an extra year of schooling are independent of the level of schooling. When growth regressions are specified in a way that is more consistent with this idea, the evidence for an effect of education is rather stronger.

Krueger and Lindahl (1999) have argued convincingly that another important problem is likely to be measurement error. The difficulty is that a specification based on an aggregate production function (as in Benhabib and Spiegel) typically seeks to explain growth using the change in educational attainment, but first-differencing the education variable in this way will usually exacerbate the effect of any measurement errors in the data.

To support this argument, Krueger and Lindahl examine the correlation between two different measures of the change in average years of schooling that have been used in the literature. The correlation is low enough to suggest that a substantial component of the measured change in educational attainment is uninformative noise. As a consequence, regressions that use the change in education to explain growth will tend to understate its importance.²²

The case for seeing measurement error as an important part of the story has been considerably strengthened by the impressively careful and detailed work of de la Fuente and Domenech (2000). Unusually, they restrict attention throughout to OECD Member countries. Their close examination of standard data sets reveals that schooling levels for some countries appear implausible; some of the figures for average years of schooling display surprising short-run volatility; and others appear to give a misleading view of trends. Other writers, notably Steedman (1996), have also noted inconsistencies in the way data on human capital are collected and compared.

By drawing on national sources and more recent figures compiled by the OECD, de la Fuente and Domenech compile a new and more reliable data set for years of schooling in OECD Member countries. In their empirical work, they find that changes in output and educational attainment are positively correlated, even in panel estimates that include country and time fixed effects. This supports the idea that, where previous researchers have failed to detect an effect, this may be due to measurement error.

More recently, Bassanini and Scarpetta (2001) have extended the de la Fuente and Domenech database forward in time, and estimated the effect of education over 1971-98 for 21 OECD Member countries using the Pooled Mean Group (PMG) estimator. The key advantage of this approach is that, compared with traditional methods of estimating panel data models, it allows greater flexibility in the short-run dynamics. Using the PMG estimator, Bassanini and Scarpetta's preferred estimate is an elasticity of 0.6 for output per capita in response to additional years of schooling. This implies that, at the sample mean of average schooling of about ten years, an extra year of average schooling would raise output per capita by 6 per cent. This effect is similar in magnitude to that found in microeconomic estimates based on survey data, of the type reviewed earlier.

Engelbrecht (1997) also finds significant effects of education on OECD growth. His empirical model controls for the effects of R&D spending, and is estimated using the education data of Barro and Lee (1993) for the population aged 25 and over. Again, these results suggest that the growth of productivity is related to the change in average years of schooling, as one would expect if microeconomic estimates of the return to schooling are picking up a genuine productivity effect. In a separate set of estimates, Engelbrecht also finds support for the idea that the level of education plays a role in technological catch-up; he finds productivity growth is more rapid where countries have a higher level of average schooling.

Overall, this literature is beginning to suggest that there is a correlation between changes in education and growth, of the kind that most labour economists would expect to observe. This is reassuring, but there are a number of interesting open questions. One obvious question mark surrounds the interpretation

of the earlier results that related growth to the initial level of attainment, rather than the change in attainment. Growth studies for the OECD that allow a role for both possibilities simultaneously are yet to appear. This omission may be inevitable given the small sample size, but it should not lead one to underestimate the possible role for human capital in technological catch-up or the creation of new ideas, either of which could yield a relationship between the level of education and subsequent growth.

There is another reason why the effect of the initial level of education remains of some interest. Studying the relation between the change in output and the change in education remains somewhat vulnerable to the charge that causality runs from output (or anticipated output) to education, and not simply *vice versa*.²³ To a large extent, long-run changes in average educational attainment are driven by government policy. It seems plausible that as output and tax revenues increase, governments will often allocate more resources to education, and attainment will rise for a transitional period.

Yet the argument that panel data results, such as those of de la Fuente and Domenech (2000), are driven by reverse causation is rather less strong than it may appear at first. This is a key advantage of their use of data on average years of schooling in the population, rather than enrolment rates. Given that new entrants are typically a small fraction of the labour force, average attainment will change only very slowly in response to any change in educational provision. It therefore seems rather unlikely that reverse causation explains the panel data findings.

Where does this leave us? Earlier in the survey, we saw the important qualifications that surround microeconomic estimates of the social returns to schooling. Ultimately we would like the cross-country evidence to shed light on the accuracy of these estimates. In practice, we are likely to remain some way short of this goal, at least in the absence of better data. The aggregate evidence is currently too fragile to draw any strong conclusions about the possible extent of social returns.

Even so, the results we have provide some grounds for optimism, and it is reassuring that several recent studies find education to be important, despite the likely presence of measurement error. This suggests that better data, and more sophisticated methods, may yet lead to a steady improvement in the precision of our estimates of the growth effects of education. The prospects for this should not be exaggerated, but there is certainly more reason to be hopeful now than in the early days of the literature, when the various sets of estimates were hard to reconcile into any kind of coherent story.

Another advantage retained by the macroeconomic approach, compared with micro estimates, is that we can explore indirect effects of education, notably those working through investment. These effects are present in the model introduced by Mankiw, Romer and Weil (1992) and may have wider relevance. Two-sector models

of endogenous growth, such as those reviewed in Barro and Sala-i-Martin (1995, Chapter 5), typically yield a steady state in which there is an equilibrium ratio of human capital to physical capital. An immediate consequence is that a rise in educational attainment will eventually be met with a corresponding rise in the stock of physical capital.

Analysing the consequences for welfare is not wholly straightforward. Growth economists have not yet developed and calibrated a model which derives overall output and welfare effects of education based on sensible microfoundations for investment.²⁴ This may explain why the effect is ignored by most interpretations of the empirical literature on education and growth. For now, it is important to be aware that growth regressions and growth accounting, by using capital growth as one of the conditioning variables, may understate the total impact of an increase in educational attainment on output per worker. The probable magnitude of this effect, and its significance for welfare, remain uncertain.

Human capital externalities

One important motivation for looking at the cross-country data is the possible presence of externalities to human capital. As we have seen, however, the empirical growth literature gives rather imprecise answers about the social returns to education. In this section, I will briefly review theoretical work on this topic, and then discuss some innovative recent evidence based on microeconomic data sets.

Interest in human capital externalities was revived by Lucas (1988, 1990). One of his arguments was that, in the absence of such externalities, it is difficult to reconcile observed pressures for migration from poor to rich countries with the absence of massive capital flows in the other direction. He also drew on the work of Jacobs (1969) to argue that such externalities are a natural explanation for the existence of cities.

In more recent work, Acemoglu (1996) has provided an ingenious justification for the presence of externalities. His theory is based on microeconomic foundations, and so is particularly worthy of attention. In his model, firms and workers make investments in physical capital and human capital respectively, before production begins. Production requires a partnership between a firm and a worker, but when firms or workers make their respective investments, they do not know the identity of their future partner. A key assumption of the model is that firms and workers are then brought together via a matching process that is imperfect, perhaps because searching for partners is costly.

Acemoglu shows how the structure of the model yields an important result: an increase in the average level of human capital can have a positive effect on the private return to human capital, at least over some regions. The intuition is as follows: say that a subset of workers decides to acquire more human capital. This will

raise average human capital, and anticipation of this encourages firms to make greater investments in physical capital. Since the matching process is inefficient, the firms that have invested more are not necessarily matched with the workers who have invested more in human capital. As a result, some of the other workers will gain from the increase in average human capital, since they are matched with firms using more physical capital than before; and in this sense the average level of human capital has an external benefit.

Work of this kind has helped to motivate the recent search for externalities, using survey data sets that include individuals who live in different cities or regions. The idea is to estimate human capital earnings functions in the normal way, but including a new variable, the average level of schooling in each individual's city or region. The central idea is that, if there are significant externalities to human capital, individuals should earn more when they work in those cities with a higher average level of schooling. The exercise will miss externalities that work at the national level, perhaps through social structures or institutions, but it remains of considerable interest.

Several studies based on this idea have been carried out for the US. The initial results of Rauch (1993) appeared promising. Consider two otherwise similar individuals living in two different cities, the second city with a population that has an extra year of average schooling. His estimates suggested that an individual living in the second city could expect to gain a wage premium of around 3 per cent, an effect large enough to be worthy of further investigation.

Unfortunately, as Ciccone *et al.* (1999) point out, there is an important argument against interpreting the observed wage premium as solely driven by externalities. Differences in average years of schooling across cities are likely to be associated with differences in the relative supplies of skilled and unskilled labour. These relative supply effects may give rise to an apparent wage premium for average schooling even in the absence of externalities.

The empirical work of Ciccone *et al.* (1999) supports this proposition. When they follow Rauch and do not allow for relative supply effects, they are able to obtain a high and precise estimate of the social return to education. In a more general approach, which builds in a role for supply effects, the measured externalities are greatly reduced; indeed it is not possible to reject the hypothesis that externalities are absent altogether. Related work by Acemoglu and Angrist (1999) also indicates that the overall social returns to education may be close to the private returns, this time using the variation in average schooling across US states to capture the effects of externalities.

Wider benefits of education

So far, the survey has only considered the effects of education on productivity, yet it is clear that the benefits of education are likely to be more far-reaching. The traditional case for education is that it makes a fundamental contribution to personal development, and probably to the health of societies more generally. In thinking about public provision, it is crucial to remember that education may have significant welfare benefits that are not captured in the models and data typically analysed by economists and governments.

These include even the benefits of education that accrue directly to individuals. It is plausible that education has both an immediate consumption benefit and a long-term effect on life satisfaction, other things equal. The difficulty here is that it is much harder to measure well-being in a meaningful way than it is to measure output of goods and services, and economists are only just starting to investigate well-being and its determinants.

In an innovative paper, Blanchflower and Oswald (2000) report estimates of "happiness equations", regressions that relate survey measures of well-being to individual characteristics. They find that educational attainment is associated with greater happiness, even when controlling for family income. Such findings could have important implications for education policy. For example, it is quite possible that the extent of an individual's education has a positive effect on the well-being of others, in which case self-interested individuals may tend to under-invest in education from society's point of view. Alternatively, education may affect happiness because it influences perceptions of status relative to others, in which case the overall welfare benefits of education may be less than the results of Blanchflower and Oswald seem to imply.

Education policy also has implications for society as a whole. Some economists may feel that these wider benefits lie outside the remit of the subject, but this argument would mean departing from the orthodox definition of economics – namely the study of the relation between the allocation of scarce resources and human welfare. Educational provision may affect public health, crime, the environment, parenting, and political and community participation. Some of these effects are discussed in more detail in OECD (1998, Chapter 4), Behrman and Stacey (1997) and Wolfe and Haveman (2000). All of these wider benefits could feed back into economic performance, which reinforces the case for a much broader view of education's role.

A tentative summary of the evidence

At this point, one may be left wondering what the evidence ultimately achieves in terms of lessons for policy. The most useful perspective is probably to combine the various strands of evidence and see whether they form a coherent whole, despite the problems inherent in each.

Labour economists seem to be agreed that the private rate of return to a year's extra schooling is typically between 5 per cent and 15 per cent. Working under similar assumptions, growth accountants find that increases in educational attainment account for perhaps a fifth of growth in output per worker.

Labour economics and growth accounting have a relatively long history, and the strengths and weaknesses of the available evidence are well understood. It is possible that both approaches overstate the social benefits of education, perhaps because of signalling effects. Acting in the other direction, the estimates provided by this research may understate the role of education, because they rarely allow measurement of externalities, or quantify the importance for productivity of an improved matching between workers and jobs, or incorporate the more general mechanisms connecting education and growth that are found in theoretical models.

The great strength of the emerging macroeconomic literature is that, at least in principle, it could provide a direct test of the productivity benefits. As we have seen, however, this field has significant weaknesses of its own. Estimates that are sufficiently accurate and robust to allow confident conclusions are some way off. That may have to wait until growth economists have longer spans of data to work with, and greater skill at matching a variety of possible statistical techniques to the question at hand.

With these caveats in mind, a brief summary of the macroeconomic evidence may be useful. Although in some ways such an exercise is rather premature, it should at least prevent the unwary from jumping to an over-hasty conclusion based on the reading of one or two papers alone. That would be an easy mistake to make: over the last ten years, growth researchers have bounced from identifying quite dramatic effects of education, to calling into question the existence of any effect at all.

More recent research is placed somewhere between these two extremes, but perhaps leaning closer to the original findings that education has a major impact. In examining the studies that have not detected an effect, we have some convincing reasons (measurement error, outliers, incorrect specification) to doubt such results. The balance of recent evidence points to productivity effects of education which are at least as large as those identified by labour economists. This should reassure us that most countries are not over-providing education, but a fuller discussion of policy implications will be deferred until the final section.

SOCIAL CAPITAL AND GROWTH

In this section, I provide some discussion of the emerging idea of “social capital”, and its potential role in the growth process. The literature on social capital is still relatively undeveloped, and in reviewing the empirical work on this topic,

I will draw heavily on a few key papers. The nature of the surrounding discussion is necessarily broader and more speculative than elsewhere in the survey, and this reflects some of the uncertainties currently surrounding the field, which should be borne in mind throughout.

Before describing the underlying ideas in more detail, it may be helpful to place them in the wider context of empirical growth research. The aim is to indicate why social capital might yet be a useful concept, given that views on its importance currently differ greatly – certainly compared with views on education, where there is fairly general agreement that education matters, even if our measurements of its effect are imprecise.

Ideally, researchers studying development and growth would like to find a set of policy interventions sufficient to raise living standards and welfare. It is sometimes argued that this is an impossible goal, partly because the circumstances of each country are unique. A less extreme position is that growth research can give us some insight into possible generalisations, by telling us about the average pattern; at the same time, it should be recognised that any proposed set of “sufficient” conditions will never be universal.

One way of making our generalisations more widely applicable is to discriminate more finely between societies, by introducing extra dimensions into our analysis of the growth process. This cannot be pushed too far, since we only have a limited set of countries, and a limited time span, from which to draw evidence. The central challenge for growth researchers is to identify the dimensions which are most relevant for growth, without endlessly multiplying the possibilities in such a way that we ultimately ask too much of the data. At the moment, the hope appears to be that a coherent picture will ultimately emerge through a gradual accumulation of evidence, as empirical researchers both introduce new variables and indicate that some earlier proposals should be discarded. The fundamental problem here is that the most general model, which in principle would allow us to discriminate easily between the competing hypotheses, has already become too large to be informative (Levine and Renelt, 1992).

In this context, in explaining growth, it makes sense to concentrate on those dimensions of societies which have a strong prior claim on our attention. Among the dimensions recently proposed for further investigation, one stands out as both promising and – in terms of its prior claim – relatively controversial. The concept of “social capital” appears to be a potentially formidable way of discriminating between countries and their growth prospects. It provides a useful way to think about aspects of societies which, though difficult to measure and incorporate into formal models, may be important determinants of long-run economic success. For some economists (not all) the intuition that “society matters” is strong enough to outweigh the current absence of much in the way of a theoretical underpinning.

There is a long academic tradition that something is not fully understood until it can be measured, and the concept of social capital presents serious problems of definition, let alone measurement. But in this respect, it is interesting to note the comment of Lucas (1988, p. 35) about the early days of human capital theory. He wrote that “the idea of human capital may have seemed ethereal when it was first introduced – at least, it did to me – but after two decades of research applications of human capital theory we have learned to ‘see’ it in a wide variety of phenomena”. The possible analogy with the present and future status of social capital should be clear.

Overall, it is easy to see why growth economists and others have started to emphasise social capital only very recently, even though the basic ideas have a long intellectual history. In this part of the survey, I will discuss some of the most recent work, starting with a discussion of the nature of social capital. This provides a necessary backdrop for the next section, which covers the limited cross-country evidence so far available, most of it based on survey evidence on willingness to trust. The implications for policy may seem rather meagre, but it should be remembered that this literature is still in its early stages. Finally, there is a discussion of some of the questions that remain to be answered.

What is social capital?

It is widely acknowledged that social capital needs to be carefully defined, if it is to prove anything more than suggestive in thinking about growth. One of the best known and most representative definitions can be found in the highly influential work of Putnam (1993): “social capital ... refers to features of social organisation, such as trust, norms, and networks, that can improve the efficiency of society by facilitating co-ordinated actions” (p. 167).

As Woolcock (1998) and others have noted, this is useful but comes close to defining social capital in terms of its function, so that it becomes difficult to separate analytically the sources of social capital from its consequences. As an example, social capital in the form of trust may be created by participation in civic associations, but these associations could themselves be regarded as an important form of social capital. The importance of this point is reinforced when one considers that social capital may also have costs: one person's valuable network may be another's restrictive interest group.

Many discussions of social capital, including those of Putnam (1993), Schuller (2000) and Woolcock (2000), associate it with a resource that is useful in achieving common objectives. For example, the suggested definition of Woolcock (2000, p. 5) is that “social capital refers to the norms and networks that facilitate collective action”. This emphasis on collective action may be problematic for economists who wish to make wider use of the idea. As I will discuss later, an

understanding of the formation of social capital is likely to require an understanding of its value as a resource to individuals (Glaeser, 2000). This can easily conflict with a definition of social capital that emphasises its role in collective action, in the usual sense of the latter term. For example, an entrepreneur who gains knowledge from participating in various networks is arguably benefiting from social capital, and this benefit occurs, and may be worthy of analysis, even if the entrepreneur does not share goals, objectives or outcomes with others.

A broader exploration of the term can be found in Woolcock (1998). He proposes a scheme in which it has four dimensions, roughly corresponding to *i*) the extent of horizontal associations; *ii*) the nature of social ties within communities; *iii*) the nature of the relation between civil society and the state; and *iv*) the quality of governing institutions. Independently of the social capital literature, economists have made some progress under category *iv*), in analysing the growth impact of the quality of institutions (for instance Knack and Keefer, 1995). At least for present purposes, it is not clear that bringing this work under the umbrella of social capital will yield extra insight.²⁵ In any case, measuring the benefits of good institutions is arguably a less urgent task than formulating practical advice on how to improve bad ones, and the growth literature does not have much to offer here.

With these points in mind, this survey will mainly restrict itself to recent empirical work that uses the extent of trust in a society as an indicator of its underlying social capital. It should already be clear that this is an imperfect and simplistic way of capturing the ideas of Putnam and others. Trust may be determined by social capital, but also by other aspects of societies; and the extent of trust may be influenced, in very different ways, by all four of the dimensions of social capital identified by Woolcock. Yet a focus on trust has one key advantage: it can potentially be measured in a way that is comparable across countries, as we will see below.

Empirical evidence

The most important macroeconomic evidence on social capital takes the World Values Survey as its starting point. The 1981 survey is based on responses from thousands of individuals across 21 market economies, while the 1990-91 survey covers 28 market economies. Overall, 29 market economies are covered at least once. The selection of respondents is not completely random, but adjustments to take this into account are available.²⁶ Among the issues addressed in the surveys, economists have mainly focused on a question designed to capture willingness to trust. Respondents were asked "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" The percentage of respondents in each nation replying "most people can be trusted" forms a potentially useful index of trust.²⁷ Table 1 shows values for this

Table 1. A measure of trust

OECD Member countries:			
Norway	61.2	Ireland	40.2
Finland	57.2	Korea	38.0
Sweden	57.1	Spain	34.5
Denmark	56.0	Austria	31.8
Canada	49.6	Belgium	30.2
Australia	47.8	Germany	29.8
Netherlands	46.2	Italy	26.3
United States	45.4	France	24.8
United Kingdom	44.4	Portugal	21.4
Switzerland	43.2	Mexico	17.7
Iceland	41.6	Turkey	10.0
Non-OECD countries:			
India	34.3	Nigeria	22.9
South Africa	30.5	Chile	22.7
Argentina	27.0	Brazil	6.7

Source: Knack and Keefer (1997).

index, TRUST, for those OECD countries covered in the survey, and also for a small selection of less developed countries.

Clearly, measurement error is potentially a major problem in using such data. For the 20 countries with TRUST values for both 1981 and 1990, the correlation between the two is 0.91, which suggests a limited role for transitory measurement errors. This leaves open the possibility, however, that the variable is an accurate measure of something other than the extent of trust. Knack and Keefer (1997) report on an interesting experiment, that provides independent evidence on the possible validity of the TRUST measure. In the experiment, a large number of wallets containing \$50 were deliberately "lost" in a number of cities. The percentage of "lost" wallets that were returned to their owners in each country has a correlation with TRUST of 0.67, providing a tentative indication that people are genuinely more trustworthy in countries with high values of the TRUST index.

Knack and Keefer (1997) also construct a second index, CIVIC, designed to capture the strength of norms of civic co-operation. The index is constructed by averaging across five questions, addressing the attitudes of the respondents to such things as fraudulent benefit claims and avoidance of fares on public transport. Perhaps surprisingly, this index shows relatively little variation across OECD countries, although it is positively correlated with TRUST. In what follows I will concentrate on the empirical evidence relating to the TRUST variable; Knack and Keefer note that results are broadly similar when CIVIC is used in its place.

Before turning to the evaluation of the results, it is worth examining the data in Table 1 more closely. Most of the empirical work is based on samples that con-

tain a small number of less developed countries, as well as OECD Member countries; as a result, one might be concerned that poorer countries are responsible for most of the identifying variation in the TRUST variable. The evidence of Table I, however, suggests that there is substantial variation in TRUST across OECD Member countries.²⁸

The index of trust is used by both Knack and Keefer (1997) and La Porta *et al.* (1997). Both these studies report cross-country regressions that relate a wide variety of dependent variables to trust and a number of controls. In many cases, the results should be regarded as indicating the existence of associations, rather than establishing a causal relationship.

The evidence for causality is arguably strongest in the regressions seeking to explain growth in output per head. La Porta *et al.* (1997) report that the trust index is weakly associated with growth over 1970-93, although the explanatory power of their growth regression is low and the sample includes some countries that were centrally planned during this period. This suggests that one should be quite careful about drawing conclusions for OECD countries. Knack and Keefer exclude socialist countries and focus on a shorter period, 1980-92. They find stronger results. Controlling for initial income per head, a human capital variable, and the relative price of investment goods, they find that a one standard deviation change in the trust index is associated with a change in the growth rate of 0.56 of one standard deviation. In alternative terms, a level of TRUST that is 10 percentage points higher (slightly less than one standard deviation) is associated with an annual growth rate that is higher by 0.8 percentage points.

As always in the empirical growth field, one should be careful not to regard these growth effects as ones that will persist indefinitely. It would perhaps be rather implausible to assert that countries will grow at permanently different rates, simply because of differing levels of trust. The correlations highlighted by La Porta *et al.* (1997) and Knack and Keefer (1997) are better seen as indications of a possible role for social capital in determining the steady-state level of income. In other words, changes in social capital might affect growth rates, but only for a transitional period. This qualification is also true of almost any other variable that one might use to explain growth, and it should be remembered that transitional effects may easily be large enough to be worth considerable attention.

Knack and Keefer carry out a number of robustness tests. When influential outliers are deleted, or growth analysed over longer periods (1960-92 and 1970-92), the point estimate of the growth effect is roughly halved, but remains statistically significant (see their Table II). They do note that, over the longer time span, the effect of TRUST is not always robust to the inclusion of other explanatory variables in the growth equation.

The evidence suggests that the effect of TRUST is large enough to be worthy of further investigation. It is important to note, however, that results are typically less strong when attention is restricted to a sample of OECD countries. Also using World Values Survey data, Helliwell (1996) found a negative effect of trust on growth in a sample of 17 OECD Member countries. Knack (2000) reports that in a sample restricted to 25 OECD Member countries, the effect of trust is imprecisely measured, and the hypothesis that it has no effect cannot be rejected at conventional significance levels.

These are quite small samples, so in a sense it is not surprising that trust is insignificant when attention is restricted to the OECD. Knack (2000) makes two additional points in relation to the OECD results. First, as in Knack and Keefer (1997), there is evidence that the effect of trust is greater in low-income countries, based on an interaction term in the growth regressions. Even if one is sceptical that trust matters for the high-income members of the OECD, it may still play an important role in poorer countries like Mexico and Turkey. Secondly, Knack (2000) reports a statistically significant and positive correlation between investment and TRUST within an OECD sample, supporting the idea that trust plays some role even for richer nations.

Both La Porta *et al.* (1997) and Knack and Keefer (1997) report evidence on other interesting associations between TRUST and indicators of performance. La Porta *et al.* find strong positive associations between TRUST and a number of measures of government performance, including the effectiveness of the judiciary and the quality of the bureaucracy (their Table 2). Knack and Keefer present very similar results (their Table V). They also provide some evidence that the effect of trust works through raising the share of investment in GDP.

These results are intriguing, but one should be careful to avoid jumping to strong conclusions about the importance of trust, or other aspects of social capital. A fundamental problem is that the extent of trust may well be determined by, or correlated with, other aspects of societies that are omitted from the growth regressions. For instance, it may be that corruption or weak legal enforcement lowers trust and, for quite independent reasons, the growth rate. As Knack and Keefer note, one could even tell a story in which trust is a product of optimism in societies that are performing well in economic terms.

One obvious variable that might be correlated with social capital is educational attainment, and this is particularly interesting from the point of view of the present survey. La Porta *et al.* (1997, p. 336) argue that trust has a positive effect on educational achievement, but it should be clear that causality may run in the opposite direction. Knack and Keefer report a strong correlation ($r = 0.83$) between TRUST and an estimate of average years of schooling for 1980, and note that "education may strengthen trust and civic norms, for example, if ignorance breeds

distrust, or if learning reduces uncertainty about the behaviour of others, or if students are taught to behave co-operatively" (p. 1270). If we see trust as endogenous to the extent and quality of education, we have the beginnings of a story about externalities to education, of the kind briefly discussed earlier.

The future for social capital research

Given that interesting and suggestive evidence for the importance of social capital has been compiled in so short a time, further research on social capital appears to have a bright future. To live up to this promise, however, there are at least two potentially difficult questions that will need to be addressed. The first question concerns the origins and formation of social capital; the second, the precise mechanisms by which social capital, once formed, gives rise to particular microeconomic and macroeconomic outcomes.

It should be clear that, to incorporate the ideas of the social capital literature in policy advice, we will often need to understand how social capital is created, and how it might sometimes be undermined. In line with the usual methods of economists, Glaeser (2000) has convincingly argued that we need to give more attention to the value of social capital as a resource for individuals, as well as for communities as a whole. It seems unlikely that social capital is best understood as simply an unintended by-product of other decisions. With this in mind, we need a model that captures the incentives of individuals to form or undermine social capital, and which also shows how these incentives are affected by policy. Without such a model, our knowledge of policy implications will remain incomplete, however strong our intuition and evidence that social capital matters.

It can seem that social capital resists the usual methods of analysis of economists, given that it is usually understood to be a property of groups rather than individuals. The Glaeser argument works well for the "networks" aspect of social capital, since participation in networks can be modelled as the outcome of individual investment decisions; the argument is less clearly applicable to other aspects of social capital, such as social norms. Yet even for social norms, such as the value of trustworthiness, it is possible to analyse their creation and evolution in terms of individual decisions to observe (or not to observe) the prevailing norm. Economists have recently started to give greater attention to constructing models in which social norms are endogenous, and it seems probable that this work will yield some valuable insights.²⁹

A second, and related, question concerns the precise mechanisms by which social capital, once in place, affects economic outcomes. Again, formal modelling may be useful. For example, Zak and Knack (1999) present a model in which agents divide their time between production and verifying the actions of those they transact with. Their model captures the simple idea that in low trust societ-

ies, some resources and time are diverted to verification, and this results in lower output.

It will be very difficult to discriminate between alternative theoretical models using macroeconomic data, and the prospects for further cross-country empirical research appear limited. Studies based at the level of firms or regions may ultimately be more informative, and some interesting work has already started to appear. Guiso *et al.* (2000) argue that one of the best testing grounds for the importance of social capital may lie in the financial sector, since it is here that trust may be especially relevant to economic activity. They study this effect within Italy, using a measure of civic engagement (essentially voter turnout in certain elections) as a proxy for social capital, as in Putnam (1993). Using large samples of households and firms, they find that their measure of civic engagement helps explain variation in financial practices across Italian regions, even when controlling for different levels of development.

Such studies are likely to play an increasingly important role in the wider debate on the importance of social capital. Sceptics will remain unconvinced by the economic importance of trust and other aspects of societies (networks, norms, participation) until we have a more complete and detailed story describing their connection to economic outcomes, supported by reliable evidence.

SUMMARY AND CONCLUSIONS

This section rounds off the paper with a discussion of how these areas of research might inform future policy. I will look at education first, but care is needed here. Griliches (1997, p. S339) notes that for academic economists, an emphasis on the importance of education for economic growth “may be somewhat self-serving” and occasionally in the literature one does come across a paper which echoes to the sound of grinding axes. This is particularly true in reading opposing assessments of the signalling argument, where the lack of reliable evidence seems to encourage, rather than discourage, strong views.

In assessing the empirical evidence for productivity benefits of education, it is quite possible that an overall judgement is frequently contaminated by a keen awareness of wider benefits of education that are not captured in economic data. After all, one could probably construct a viable case for much educational expenditure entirely based on its implications for personal development, independent of any productivity effects. It is worth quoting Weiss (1995, p. 151):

“Education does not have to be justified solely on the basis of its effect on labour productivity. This was certainly not the argument given by Plato or de Tocqueville and need not be ours. Students are not taught civics, or art, or

music solely in order to improve their labour productivity, but rather to enrich their lives and make them better citizens.”

Most economists, appropriately enough for practitioners of the “dismal science”, have concentrated on examining a rather more narrow case for education, in terms of its contribution to productivity growth. As we have seen, the weight of evidence points to significant productivity effects, but the degree of uncertainty is large, and even a lower bound is surprisingly difficult to establish.

The evidence from labour economics has the greatest weight of experience, time and academic firepower behind it, and this suggests that it would be a mistake to summarise the macroeconomic results in isolation. Although a reconciliation of these two literatures is in its early stages, the correlation across countries between measures of human capital and growth is arguably robust enough to support the belief that earnings functions pick up genuine productivity effects, and not simply the effects of signalling or omitted characteristics.

That is reassuring, but it leaves many questions open for policy-makers. There is likely to be pervasive heterogeneity in rates of return across individuals, let alone across countries. A greater understanding of the pattern of heterogeneity will lead to better policy decisions, but on this subject the macroeconomic literature surveyed here is necessarily silent. Evidence compiled by labour economists will be far more useful in this respect.³⁰

Other limitations of the macroeconomic evidence are worth noting. Growth regressions are best thought of as picking up an average effect of schooling, and should certainly not be used to conclude that every OECD Member country is currently under-providing education. Indeed, the results from growth accounting exercises suggest that, although increases in educational provision can yield a worthwhile increase in the growth rate, one should not necessarily expect an effect that is large relative to current rates of growth. For policy-makers who wish to raise the growth rate, policy on education remains a natural place to look, but it is by no means a panacea.

Not only that, one cannot altogether dismiss the possibility of “over-education” in some countries, at some levels of education. For example, it is plausible that there are some people for whom conventional academic education beyond a certain point is ultimately less useful than vocational training. This draws attention, once again, to the way in which the unbalanced nature of the existing evidence may mislead. We need to learn more about the relative merits of schooling and training for different individuals. This will require estimates of how returns to schooling vary with personal characteristics, and should also involve some consideration of wider benefits.

In thinking broadly about the over-education question, it is interesting to consider the evolution of educational wage differentials since the late 1970s. Even

though the relative supply of skilled labour has increased, there has been a substantial and well-documented increase in educational wage differentials in the UK and the US, with less pronounced changes in other OECD countries.³¹ It seems difficult to explain the evidence for the UK and the US without a dominant role for a shift in the relative demands for different types of labour, favouring the more educated.

Much research has focused on the origins of the change in relative labour demand, but for policy-makers an equally important question is whether this change is generating a rising return mainly to education, or to other characteristics such as innate ability or initiative. Clearly the policy implications are very different under the alternative scenarios, yet disentangling the two effects is difficult. Existing research often finds that it is the return to ability which is rising, but the work of Cawley *et al.* (1998) suggests that, due to some important identification problems, these results are not robust to small changes in assumptions.

It is also the case that over-education may take time to appear in the data. One problem here is that average attainment typically evolves only slowly, and so measured returns to education will also change only slowly, and are not necessarily informative on the desirability of current provision. This means that evidence on current wage differentials needs to be supplemented by other approaches, including those discussed in Harmon *et al.* (2000).

The recent shifts in differentials also remind us that policy on education has distributional consequences (Topel, 1997). Given that trade does not seem to equalise factor prices across countries, any increase in the relative supply of skilled labour is likely to lower the wage premium for the possession of skills. In turn this could make an important contribution to reducing income inequality.³²

To summarise, can we justify the massive amount of resources allocated to education by OECD Member countries, around \$1 550 billion in total each year? On the available evidence, including recent changes in wage dispersion, the arguments for cutting back on this provision seem rather weak. In deciding if provision should be expanded, perhaps the key open question is the validity of the signalling arguments. More evidence on the signalling debate would be extremely helpful in judging the benefits of expanding higher education, one of the main changes in provision within the OECD since the 1960s.

In exploring some of the details of such arguments, empirical evidence is not the only way forward. Theory and calibration exercises may also shed light on these issues. An example is the interesting implication of new growth theory that individuals may under-invest in education, because those who later go into research careers do not capture all the benefits of the new ideas that they help to create. This provides the beginnings of an argument for subsidising education in engineering and science, at least at those levels (perhaps PhDs, or post-doctorates) where a

high proportion of individuals subsequently go into research and development activity. Romer (2000) has recently presented specific policy proposals along these lines.

The literature on social capital and growth is at an earlier stage than the macroeconomic evidence on education, and the policy implications are less clear. Indeed, one weakness of the social capital literature, at least in relation to richer countries, is that it is currently difficult to see what policy conclusions could ever be drawn. What can a policy-maker in Mexico or Turkey actually do, confronted with the evidence from the World Values Survey that they govern a low-trust society? Standard recommendations, such as attempting to eliminate corruption and improve the legal system, are nothing new, and make good sense quite independently of any emphasis on social capital.

Perhaps the best answer lies in drawing an analogy with the introduction of human capital theory into economics. In its early stages, as Lucas (1988) makes clear, human capital seemed a rather ethereal concept, and presumably one with little immediate message for education policies. Work on social capital is still in its early stages, and as we learn more about what it is, where it comes from, and what it does, there may ultimately be implications and conclusions that leave our successors wiser in ways that we can only guess at.

NOTES

1. The expenditure share is taken from OECD (2000a) and relates to 1997. Collective GDP is based on total GDP in 2000 for 29 of the 30 current OECD Member countries, at current prices and exchange rates, where the Slovak Republic is the country excluded. The GDP figure is taken from the national accounts statistics available online at www.oecd.org/.
2. One consequence of this omission is that I will have to ignore the interaction between education and training. To the extent that education is about “learning how to learn”, it may have consequences for the value of subsequent on-the-job training. Some international comparisons of training programmes can be found in OECD (1998, Chapter 3). The work of Van Ark and Pilat (1993) includes an examination of the role of vocational skills in explaining productivity differences across Germany, Japan and the United States.
3. A more detailed and rigorous summary can be found in Aghion and Howitt (1998, Chapter 10).
4. Note that this effect is potentially independent of other benefits of increased knowledge, such as increases in the quality of capital goods, or more general forms of technical progress.
5. The assumption that it is difficult to fully capture the benefits of research is uncontroversial. The presence of substantial research spillovers is intuitively plausible, and supported by empirical evidence. Griliches (1992) provides a survey.
6. A complete welfare analysis of policy intervention would need to consider the effects on the level of the output path, as well as its growth rate.
7. For example, some individuals outside the R&D sector, but originally trained as scientists, may switch into R&D careers in response to higher wages.
8. Sianesi and Van Reenen (2000) also provide a review of the macroeconomic literature on education and growth, with extra detail on individual papers. Scarpetta *et al.* (2000) analyse the recent growth performance of OECD Member countries in more general terms.
9. An innovative paper by Judson (1998) investigates whether educational spending is allocated efficiently. It seems likely that future research will give increasing emphasis to this topic.
10. The interaction between growth, human capital and female labour force participation is discussed in more detail by Mincer (1996). For evidence on female labour force participation in the OECD, see OECD (1998, Chapter 4).

11. Card (1999) and Harmon *et al.* (2000) provide excellent and detailed reviews of the various issues. Another useful survey is that of Ashenfelter *et al.* (1999), which includes a detailed investigation of problems associated with publication bias.
12. Different authors use the term “social return” in different ways. Here I will use it to denote the overall return on an individual's education from society's point of view, rather than (say) the private return simply adjusted for taxation and direct costs of education.
13. Space precludes a full discussion of the interpretation of these results, and their robustness. For a more detailed review, see Harmon *et al.* (2000).
14. Note that, depending on the approach adopted, some forms of technical change will be treated as changes in the quality of capital inputs, and will not appear in the residual.
15. The KLEMS project is seeking to extend this form of analysis to other major economies. See: www.conference-board.org/economics/klems/index.htm
16. Barro and Sala-i-Martin (1995, p. 352) make this point in greater detail.
17. Other papers which extend these findings for the OECD sample, and at least implicitly examine their robustness, include Nonneman and Vanhoudt (1996), Temple (1998) and Vasudeva Murthy and Chien (1997).
18. There is also important work on human capital as a determinant of technological catch-up using data at the sectoral level. For example, Cameron, Proudman and Redding (1998) investigate the role of human capital and openness to trade in explaining catch-up by UK manufacturing sectors.
19. Of the two measures, only the secondary school enrolment rate is likely to be relevant in explaining growth within the OECD. Englander and Gurney (1994a) report that average secondary enrolment in the OECD was about 70 per cent in 1960, so there may be enough variation across countries for regression evidence to be informative.
20. Problems in discriminating between the two effects are discussed in Cannon (2000).
21. This finding is also associated with a number of panel data studies using fixed effects, but these results should almost certainly be discounted. Researchers using panels typically do not allow for lags in the effect of variables like enrolment rates. In any case, given the way the education data are constructed, the time series variation will sometimes be too noisy to draw sensible conclusions.
22. Note, though, that measurement error in other explanatory variables (notably physical capital) could bias the coefficient on education in the opposite direction.
23. The two-way interaction between growth and education is discussed in more detail by Mincer (1996) and Bils and Klenow (2000). Bils and Klenow argue that the direction of causality may be uncertain even when attention is restricted to the growth effect of the initial level of education.
24. Although some theoretical work has started to appear: Masters (1998) analyses the efficiency of investments in human and physical capital in a bilateral search context.
25. Following Abramovitz (1986), Temple and Johnson (1998) argue in favour of the use of the term “social capability” when referring to social arrangements and institutions defined more broadly. There are likely to be important benefits, in terms of clarity and rigour, from keeping the term “social capital” narrowly defined, as Putnam has advocated.
26. For a more detailed discussion, see Knack and Keefer (1997).

27. Note that someone's response to the survey question may tell us more about their own trustworthiness, rather than a view of trust in their country as a whole. Even then, the pattern of responses may form a useful guide to the prevalence of trust in a particular country (see for instance Glaeser, 2000).
28. The scatter plots presented in Knack and Keefer are also reassuring in this respect, as they suggest that the partial correlations between growth, investment and TRUST that will be discussed later are not simply driven by the inclusion of a few less developed countries.
29. Many references to research in this area can be found in Zak and Knack (1999).
30. A special issue of the *Journal of Labour Economics* (November, 1999) includes studies that measure the returns to schooling for a variety of OECD Member countries, and thus sheds light on the possible heterogeneity across countries.
31. See for instance Katz and Autor (1999), pp. 1501-1503. Evidence on recent trends in wage dispersion more generally can be found in OECD (1996, Chapter 3).
32. One has to be careful in making this argument, even in a simple model with just two types of labour. Inequality depends not only on the skill premium, but also on the relative supplies of skilled and unskilled labour.

BIBLIOGRAPHY

- ABRAMOVITZ, M. (1986),
"Catching up, forging ahead, and falling behind", *Journal of Economic History*, 46, 385-406.
- ACEMOGLU, D. (1996),
"A microfoundation for social increasing returns in human capital accumulation", *Quarterly Journal of Economics*, 111, 779-804.
- ACEMOGLU, D. (1997),
"Training and innovation in an imperfect labour market", *Review of Economic Studies*, 64, 445-464.
- ACEMOGLU, D. and J. ANGRIST (1999),
"How large are the social returns to education? Evidence from compulsory schooling laws", *NBER Working Papers*, No. 7444.
- AGHION, P. and P. HOWITT (1998),
Endogenous Growth Theory. MIT Press, Cambridge.
- ANGRIST, J.D. and A.B. KRUEGER (1991),
"Does compulsory school attendance affect schooling and earnings?", *Quarterly Journal of Economics*, 106, 979-1014.
- ARROW, K.J. (1973),
"Higher education as a filter", *Journal of Public Economics*, 2, 193-216.
- ASHENFELTER, O., C. HARMON and H. OOSTERBEEK (1999),
"A review of estimates of the schoolings/earnings relationship, with tests for publication bias", *Labour Economics*, 6(4), 453-470.
- BARRO, R.J. (1991),
"Economic growth in a cross section of countries", *Quarterly Journal of Economics*, 106(2), 407-443.
- BARRO, R.J. and J.-W. LEE (1993),
"International comparisons of educational attainment", *Journal of Monetary Economics*, 32, 363-394.
- BARRO, R.J. and X. SALA-i-MARTIN (1995),
Economic Growth. McGraw-Hill, New York.
- BASSANINI, A. and S. SCARPETTA (2001),
"Does human capital matter for growth in OECD countries? Evidence from pooled mean-group estimates", *OECD Economics Department Working Papers*, No. 282.
- BEHRMAN, J.R. and N. STACEY (1997),
The Social Benefits of Education. University of Michigan Press, Ann Arbor.

- BENHABIB, J. and M.M. SPIEGEL (1994),
 "The role of human capital in economic development: evidence from aggregate cross-country data", *Journal of Monetary Economics*, 34(2), 143-173.
- BILS, M. and P.J. KLENOW (2000),
 "Does schooling cause growth?", *American Economic Review*, 90(5), 1160-1183.
- BLANCHFLOWER, D.G. and A.J. OSWALD (2000),
 "Well-being over time in Britain and the USA", *NBER Working Papers*, No. 7487.
- BROADBERRY, S.N. and K. WAGNER (1996),
 "Human capital and productivity in manufacturing during the twentieth century: Britain, Germany and the United States", In Van Ark, B. and N. Crafts (eds.), *Quantitative Aspects of Post-War European Economic Growth*, Cambridge University Press, Cambridge.
- CAMERON, G., J. PROUDMAN and S. REDDING (1998),
 "Productivity convergence and international openness", In J. Proudman and S. Redding (eds.) *Openness And Growth*, Bank of England, London.
- CANNON, E. (2000),
 "Human capital: level versus growth effects", *Oxford Economic Papers*, 52, 670-676.
- CARD, D. (1999),
 "The causal effect of education on earnings", in O.C. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. 3A, North-Holland, Amsterdam.
- CAWLEY, J., J. HECKMAN and E. VYTLACIL (1998),
 "Cognitive ability and the rising return to education", *NBER Working Papers*, No. 6388.
- CICCONE, A., G. PERI and D. ALMOND (1999),
 "Capital, wages and growth: theory and evidence", *CEPR Discussion Papers*, No. 2199.
- DE LA FUENTE, A. and R. DOMENECH (2000),
 "Human capital in growth regressions: how much difference does data quality make?", manuscript, CSIC, Campus de la Universidad Autonoma de Barcelona.
- DENISON, E.F. (1967),
Why Growth Rates Differ, Brookings Institution, Washington DC.
- ENGELBRECHT, H.-J. (1997),
 "International R&D spillovers, human capital and productivity in OECD economies: an empirical investigation", *European Economic Review*, 41, 1479-1488.
- ENGLANDER, A.S. and A. GURNEY (1994a),
 "Medium-term determinants of OECD productivity", *OECD Economic Studies*, 22, 49-109.
- ENGLANDER, A.S. and A. GURNEY (1994b),
 "OECD productivity growth: medium-term trends", *OECD Economic Studies*, 22, 111-129.
- FINEGOLD, D. and D. SOSKICE (1988),
 "The failure of training in Britain: analysis and prescription", *Oxford Review of Economic Policy*, 4, 21-53.
- GEMMELL, N. (1996),
 "Evaluating the impacts of human capital stocks and accumulation on economic growth: some new evidence", *Oxford Bulletin of Economics and Statistics*, 58(1), 9-28.
- GLAESER, E. (2000),
 "The formation of social capital", manuscript, Harvard University, March.
- GRILICHES, Z. (1992),
 "The search for R&D spillovers", *Scandinavian Journal of Economics*, 94, S29-S47.

- GRILICHES, Z. (1997),
"Education, human capital, and growth: a personal perspective", *Journal of Labor Economics*, 15(1), S330-S344.
- GUIO, L., P. SAPIENZA and L. ZINGALES (2000),
"The role of social capital in financial development", *NBER Working Papers*, No. 7563.
- HANUSHEK, E.A. and D.D. KIMBO (2000),
"Schooling, labor force quality, and the growth of nations", *American Economic Review*, 90(5), 1184-1208.
- HARMON, C., H. OOSTERBEEK and I. WALKER (2000),
"The returns to education: a review of evidence, issues and deficiencies in the literature", Centre for the Economics of Education, London School of Economics, December.
- HELLIWELL, J. (1996),
"Economic growth and social capital in Asia", *NBER Working Papers*, No. 5470.
- JACOBS, J. (1969),
The Economy of Cities, Random House, New York.
- JONES, C.I. (1995),
"R&D-based models of economic growth", *Journal of Political Economy*, 103, 759-784.
- JORGENSEN, D.W., F.M. GOLLOP and Barbara M. FRAUMENI (1987),
Productivity and US Economic Growth. Harvard University Press, Cambridge.
- JORGENSEN, D. and E. YIP (1999),
"Whatever happened to productivity growth?" manuscript, Harvard University, June.
- JUDSON, R. (1998),
"Economic growth and investment in education: how allocation matters", *Journal of Economic Growth*, 3, 337-359.
- KATZ, L.F. and D.H. AUTOR (1999),
"Changes in the wage structure and earnings inequality", in O.C. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. 3A, North-Holland, Amsterdam.
- KNACK, S. (2000),
"Trust, associational life and economic performance in the OECD", manuscript, The World Bank.
- KNACK, S. and P. KEEFER (1995),
"Institutions and economic performance: cross-country tests using alternative institutional measures", *Economics and Politics*, 7(3), 207-227.
- KNACK, S. and P. KEEFER (1997),
"Does social capital have an economic payoff? A cross-country investigation", *Quarterly Journal of Economics*, 112(4), 1251-1288.
- KRUEGER, A.B. and M. LINDAHL (1999),
"Education for growth in Sweden and the world", *NBER Working Papers*, No. 7190.
- LA PORTA, R., F. LOPEZ-DE-SILANES, A. SHLEIFER and R. VISHNY (1997),
"Trust in large organisations", *American Economic Review*, 87(2), 333-338.
- LEVINE, R. and D. RENELT (1992),
"A sensitivity analysis of cross-country growth regressions", *American Economic Review*, 82(4), 942-963.
- LUCAS, R.E. (1988),
"On the mechanics of economic development", *Journal of Monetary Economics*, 22, 3-42.

- LUCAS, R.E. (1990),
 "Why doesn't capital flow from rich to poor countries?", *American Economic Review*, 80(2), 92-96.
- MADDISON, A. (1987),
 "Growth and slowdown in advanced capitalist economies: techniques of quantitative assessment", *Journal of Economic Literature*, 25, 649-698.
- MADDISON, A. (1991),
Dynamic Forces in Capitalist Development, Oxford University Press, Oxford.
- MANKIW, N.G., D. ROMER and D. WEIL (1992),
 "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, 107, 407-437.
- MASTERS, A.M. (1998),
 "Efficiency of investment in human and physical capital in a model of bilateral search and bargaining", *International Economic Review*, 39, 477-494.
- MINCER, J. (1974),
Schooling, Experience and Earnings, Columbia University Press, NY.
- MINCER, J. (1996),
 "Economic development, growth of human capital, and the dynamics of the wage structure", *Journal of Economic Growth*, 1(1), 29-48.
- NONNEMAN, W. and P. VANHOUDT (1996),
 "A further augmentation of the Solow model and the empirics of economic growth for OECD countries", *Quarterly Journal of Economics*, 111, 943-953.
- OECD (1996),
Employment Outlook, OECD, Paris.
- OECD (1998),
Human Capital Investment: An International Comparison, Centre for Educational Research and Innovation, OECD, Paris.
- OECD (2000a),
Education at a Glance, OECD, Paris.
- OECD (2000b),
OECD Economic Outlook 68. OECD, Paris.
- PRITCHETT, L. (1996),
 "Where has all the education gone?", *World Bank Policy Research Department Working Papers*, No. 1581.
- PUTNAM, R. (1993),
Making Democracy Work, Princeton University Press, Princeton.
- QUIGGIN, J. (1999),
 "Human capital theory and education policy in Australia", *Australian Economic Review*, 32(2), 130-144.
- RAUCH, J. (1993),
 "Productivity gains from geographic concentration of human capital: evidence from the cities", *Journal of Urban Economics*, 34(3), 380-400.
- REDDING, S. (1996),
 "The low-skill, low-quality trap: strategic complementarities between human capital and R&D", *Economic Journal*, 106, 458-470.

- ROMER, P.M. (1990),
 "Endogenous technological change", *Journal of Political Economy*, 98(5), S71-S102.
- ROMER, P.M. (2000),
 "Should the government subsidize supply or demand in the market for scientists and engineers?", *NBER Working Papers*, No. 7723.
- RUSTICHINI, A. and J.A. SCHMITZ (1991),
 "Research and imitation in long-run growth", *Journal of Monetary Economics*, 27(2), 271-292.
- SCARPETTA, S., A. BASSANINI, D. PILAT and P. SCHREYER (2000),
 "Economic growth in the OECD area: recent trends at the aggregate and sectoral level", *OECD Economics Department Working Papers*, No. 248.
- SCHULLER, T. (2000),
 "The complementary roles of human and social capital", manuscript, Birkbeck College, University of London, March.
- SIANESI, B. and J. VAN REENEN (2000),
 "The returns to education: a review of the macroeconomic literature", Centre for the Economics of Education, *London School of Economics Discussion Papers*, No. 6.
- SPENCE, A.M. (1973),
 "Job market signaling", *Quarterly Journal of Economics*, 87, 355-374.
- STEEDMAN, H. (1996),
 "Measuring the quality of educational outputs: a note", *CEP Discussion Papers*, No. 302, London School of Economics.
- STIGLITZ, J.E. (1975),
 "The theory of screening, education, and the distribution of income", *American Economic Review*, 65, 283-300.
- TEMPLE, J.R.W. (1998),
 "Equipment investment and the Solow model", *Oxford Economic Papers*, 50(1), 39-62.
- TEMPLE, J.R.W. (1999a),
 "The new growth evidence", *Journal of Economic Literature*, 37(1), 112-156.
- TEMPLE, J.R.W. (1999b),
 "A positive effect of human capital on growth", *Economics Letters*, 65(1), 131-134.
- TEMPLE, J.R.W. (2001),
 "Generalizations that aren't? Evidence on education and growth", *European Economic Review*, 45(4-6), 905-918.
- TEMPLE, J.R.W. and P.A. JOHNSON (1998),
 "Social capability and economic growth", *Quarterly Journal of Economics*, 113, 965-990.
- TOPEL, R. (1997),
 "Factor proportions and relative wages: the supply-side determinants of wage inequality", *Journal of Economic Perspectives*, 11(2), 55-74.
- TOPEL, R. (1999),
 "Labor markets and economic growth", in O.C. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. 3C, North-Holland, Amsterdam.
- UZAWA, H. (1965),
 "Optimal technical change in an aggregative model of economic growth", *International Economic Review*, 6, 18-31.

- VAN ARK, B. and D. PILAT (1993),
"Productivity levels in Germany, Japan, and the United States: differences and causes",
Brookings Papers on Economic Activity: Microeconomics 2, Washington D.C., December.
- VASUDEVA MURTHY, N.R. and I.S. CHIEN (1997),
"The empirics of economic growth for OECD countries: some new findings", *Economic Letters*, 55, 425-429.
- WEISS, A. (1995),
"Human capital vs. signalling explanations of wages", *Journal of Economic Perspectives*, 9(4), 133-154.
- WOESSMANN, L. (2000),
"Specifying human capital: a review, some extensions, and development effects", *The Kiel Institute of World Economics Working Papers*, No. 1007.
- WOLFE, B. and R. HAVEMAN (2000),
"Accounting for the social and non-market benefits of education", manuscript, University of Wisconsin-Madison.
- WOOLCOCK, M. (1998),
"Social capital and economic development: toward a theoretical synthesis and policy framework", *Theory and Society*, 27, 151-208.
- WOOLCOCK, M. (2000),
"The place of social capital in understanding social and economic outcomes", manuscript, The World Bank.
- YOUNG, A. (1995),
"The tyranny of numbers: confronting the statistical realities of the East Asian growth experience", *Quarterly Journal of Economics*, 110(3), 641-680.
- ZAK, P.J. and S. KNACK (1999),
"Trust and growth", manuscript, Claremont Graduate University, August.

R&D AND PRODUCTIVITY GROWTH: PANEL DATA ANALYSIS OF 16 OECD COUNTRIES

Dominique Guellec and Bruno van Pottelsberghe de la Potterie

TABLE OF CONTENTS

Introduction	104
The model and data	106
Estimation results	111
Tentative policy conclusions	117
<i>Appendix 1.</i> Calculation and data sources of the technological variables	120
<i>Appendix 2.</i> Regressions in log-levels.....	122
<i>Appendix 3.</i> Stability of the estimated parameters.....	124
Bibliography	125

Economic Analysis and Statistics Division, Directorate for Science, Technology and Industry; and Brussels University (ULB), Solvay Business School, (Solvay Chair of Innovation), Centre Emile Bernheim and DULBEA (CP 145/01, 21 avenue F.D. Roosevelt, B-1050 Brussels. Email: bruno.vanpottelsberghe@ulb.ac.be), respectively.

INTRODUCTION

The recent acceleration of multifactor productivity (MFP) in several OECD Member countries, after two decades of slow growth, is often explained by a surge in the pace of technical change. That is consistent with both economic theory and anecdotal evidence. Economic theory (Solow, 1957; Romer, 1990) points to technical change as the major source of productivity growth in the long run. Anecdotal evidence suggests that new technology (especially information technology) has substantially contributed to recent improvement in the productivity of firms.¹

This study presents estimates of the contribution of technical change to MFP growth in major OECD countries over the 1980-98 period. It contributes to the existing literature in this field of analysis in two ways. First the major sources of new technology are taken into account simultaneously: domestic business R&D, public R&D and foreign business R&D. Second, an attempt is made to differentiate across countries the impact of the three sources of knowledge on output growth. The results are intended to provide insights into the following:

- The contribution of business generated technology to productivity growth.
- The importance of foreign flows of technology (“international spillovers”) as compared with domestic technology.
- The contribution of government and university research to productivity growth.
- The role of absorptive capabilities of firms for technology coming from public and foreign sources.
- How the impact of the various sources of new technology has evolved over time.
- Which country-specific factors influence the effect of these various sources of technology.

The analysis is performed at the aggregate (macroeconomic) level for 16 OECD countries over the period 1980-98 using annual data. This study complements a previous work published by the OECD (Bassanini *et al.*, 2000) which took into account more factors of productivity growth (*e.g.* human capital) but did not investigate in detail the effects of technology on MFP. Results from these two studies are consistent with each other.

Research and development, resulting in new goods, new processes and new knowledge, is a major source of technical change. As defined by the *Frascati Manual* (OECD, 1993, p. 29), R&D “comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge and the use of this stock of knowledge to devise new applications”. There are different types of R&D, however, and the effect of R&D on productivity may work through various channels. In order to capture the links between R&D and productivity it is necessary to take these aspects into account. R&D is not the only source of new technology: in modern, industrial economies, other activities, such as learning by doing or design are conducted in most cases on the basis of new technology coming out of R&D (*e.g.* changes in the organisation of business related to the use of information and communication technology). The relationship between R&D and innovation is a complex, non-linear one. However, it is recognised also that it is difficult for substantial advances in technology to occur without work undertaken on a systematic basis (even serendipity tends to develop in such a context), and R&D is a good indicator of this broader phenomenon.

R&D performed by business results in new goods and services, in higher quality of output and in new production processes. These are factors of productivity growth at the firm level and at the macroeconomic level. The effect of business R&D on productivity has been investigated in many empirical studies, performed at all aggregation levels – business units, firm, industry and country levels – and for many countries (especially the United States). All these studies reach the conclusion that R&D matters, the estimated output elasticity with respect to business R&D varying from 10 per cent to 30 per cent (see a survey of the literature by Nadiri, 1993). This large variation is mainly due to the fact that studies differ in terms of the econometric specification, data sources, number of economic units, measurement methods for R&D and economic performance, and periods under study. Business performed R&D may be funded by business itself or by government (Guellec and van Pottelsberghe, 1999 and 2001*b*): it might be that business R&D has a different effect on productivity depending on its source of funds (which affects the research agenda and the incentive structure).

Government and university research have a direct effect on scientific knowledge and public missions, they generate basic knowledge. In many cases the effect of government research on productivity is not measured, either because it is indirect² or because its results are not integrated in existing measures of GDP (health-related research allows to improve length and quality of life, which are not taken into account in GDP measures). Basic research performed mainly by universities enhances the stock of knowledge of the society. New knowledge is not considered as an output in the current system of national accounts (contrary to physical investment and software for instance), and as such it is not included in GDP measures: hence the direct outcome of basic research is overlooked. How-

ever, basic research may open new opportunities to business research, which in turn affects productivity.

It is therefore not surprising that there have been very few studies of the effects of public research on productivity. Only some components of public research have been used in empirical frameworks. For instance, Adams (1990) finds that fundamental stocks of knowledge, proxied by accumulated academic scientific papers, significantly contributed to productivity growth in US manufacturing industries. Another example is provided by Poole and Bernard (1992) for military innovations in Canada, who present evidence that a defence-related stock of innovation has a negative and significant effect on the total factor productivity growth of four industries over the period 1961-85.

Foreign knowledge (knowledge generated in other countries) is a third source of new technology for any national economy. There are many ways for technology to cross borders, as knowledge coming out of a given country's research is used by another country's enterprises. Companies can buy patents, licences or know-how from foreign firms, they can observe competition (*e.g.* reverse engineering), they can hire foreign scientists and engineers, they can interact with foreign competitors who invested in their country (foreign direct investment), read the scientific and technological literature, or have direct contacts with foreign engineers in conferences or fairs. The impact of foreign-produced knowledge on a country's productivity may depend on the capacity of the recipient country to digest such knowledge, to make efficient use of it, which requires in turn this country to have sufficient technological activity of its own. This is traditionally labelled as the "absorptive capacity" of an economy.

A few studies, such as by Coe and Helpman (1995), OECD (1997) and van Pottelsberghe and Lichtenberg (2001), have estimated the effect of foreign R&D on productivity. This is done by regressing multi factor productivity (MFP) on a stock of domestic R&D and a stock of foreign R&D. Coe and Helpman find that domestic R&D contributes significantly to productivity growth and that this impact is substantially higher for the G7 than for other developed countries. In addition, foreign R&D has a significant impact on MFP growth. Lichtenberg and van Pottelsberghe (1998) show that foreign R&D can affect domestic performances through both imports and outward foreign direct investment (through technology sourcing and learning practices).

THE MODEL AND DATA

Based on the above framework, we estimate the contribution of technical change to productivity growth. We distinguish the various sources of technical change: domestic, foreign, and public sources. We also take into account business-cycle effects that

strongly influence productivity in the short run. The model on which the estimated equation is based is a simple Cobb-Douglas production function.

$$MFP_{it} = \exp [\phi_i + \phi_t + \mu_{it}] \cdot BRD_{it-1}^{\beta_{rd}} \cdot FRD_{it-1}^{\beta_{frd}} \cdot PRD_{it-2}^{\beta_{prd}} \cdot U_{it}^{\sigma_U} \cdot G^{\sigma_G} \quad (1)$$

The variables (for country i and time t) are defined as follows.

MFP is an index of total factor productivity of industry. MFP has been computed in the usual way (OECD, 2001), as the ratio of the domestic product of industry on the weighted sum of the quantity of labour and fixed capital stock, the weights being the annual labour cost share and the capital cost share, respectively (under assumptions of perfect competition and constant return to scale). The series comes from the OECD National Accounts database.

BRD is the domestic business R&D capital stock. It has been computed using the perpetual inventory method from total intramural business R&D expenditures, in constant 1990 GDP prices and US PPPs. The depreciation rate is 15 per cent (sensitivity analysis shows that the results of the regressions do not change significantly with the chosen depreciation rate). The series come from the OECD Main Science and Technology Indicators.

FRD is the foreign R&D capital stock, which is the weighted sum of the domestic business R&D capital stocks of the 15 other countries of the panel. The weights correspond to the bilateral technological proximity between countries (see Appendix 1). The underlying assumptions are two-fold: first, technology circulates directly, with no need for exchange of goods as a vector (although this may help). This assumption differs from that of Coe and Helpman (1995), who measure foreign capital stock for any country as the sum of other countries' R&D capital stock *weighted by the foreign trade structure of the country*. However, our assumption is consistent with available evidence on the circulation of knowledge across borders.³ The second assumption is that a country will benefit more from foreign knowledge relating to the same technology fields it works on than from knowledge in other fields.

PRD is total public R&D capital stock, which comprises R&D expenditures performed in the higher education sector and in the government sector (public laboratories). The depreciation rate is 15 per cent (again, sensitivity analysis shows that the results of the regressions do not change significantly with the chosen depreciation rate). Since these R&D activities are not performed by the business sectors, we expect a longer delay before they affect business productivity and therefore include them in the model with a two-year lag.

A range of control variables is included in all the regressions. U is intended to capture the business cycle effect: it is equal to 1 minus the unemployment rate. This should be a better proxy than the usually applied rate of utilisation of capital,

which applies to manufacturing industries only (which account for about 20 per cent of GDP in OECD countries). In the context of this study, it is also better than the output gap, as the calculation of the output gap relies on certain assumptions on MFP growth: by using it, we would be faced with simultaneity problems (if MFP is the same on both sides of the equation) or inconsistency (if two different MFPs are used on the two sides of the equation). G is a dummy equal to 1 for Germany in 1991, and 0 otherwise; in order to take into account the exogenous shock of the German unification. ϕ_i are country dummies which allow country-specific framework conditions that might affect long-term growth. φ_t are time dummies which take into account exogenous technical change and exogenous shocks that are common to several countries, such as changes in exchange rates.

The basic equation we estimate, adapted from (1) is an error correction model (ECM), which allows separating short-term from long-term effects. The long-term (stationary) form of the model expressed in logarithmic form is as follows:

$$LMFP_{it} = \beta_{brd} LBRD_{it-1} + \beta_{frd} LFRD_{it-1} + \beta_{prd} LPRD_{it-2} + \sigma_U LU_{it} + \sigma_G G + \phi_i + \varphi_t + \mu_{it} \quad (1')$$

Which translates into the following error correction model (ECM):

$$\begin{aligned} \Delta MFP_{it} = & \lambda \Delta MFP_{it-1} + \alpha_{brd} \Delta BRD_{it-1} + \alpha_{frd} \Delta FRD_{it-1} + \alpha_{prd} \Delta PRD_{it-2} + \eta LMFP_{it-2} \\ & + \beta_{brd} LBRD_{it-2} + \beta_{frd} LFRD_{it-2} + \beta_{prd} LPRD_{it-3} + \sigma_U \Delta U_{it} + \sigma_G \cdot G + \phi_i + \varphi_t + \mu_{it} \end{aligned} \quad (2)$$

Where Δ represents the first logarithmic difference and L the natural logarithm. In this equation, the long-term elasticity of output with respect to, say business R&D (LBRD), is $[-\beta_{RD} / \eta]$.

The parameters that are to be estimated are assumed to be constant across countries and over time; they are defined as follows:

- β_{brd} The elasticity of MFP with respect to domestic business R&D.
- β_{frd} The elasticity of MFP with respect to foreign business R&D.
- β_{prd} The elasticity of MFP with respect to public R&D.
- σ_U The elasticity of MFP with respect to the capacity utilisation rate.
- σ_G The impact of the German unification on MFP in Germany.

Interpretation of these elasticities should take into account the fact that the explained variable is not GDP but MFP. That means that we capture only the spill-over effects of R&D, not the total effect on output growth (which includes also the direct effect on private return). This concerns especially business R&D: part of the private resources devoted to R&D (labour and capital) are already reflected in the calculation of MFP, as they are included in the economy's stock of capital and pool

of labour. Hence, if the social return to R&D is equal to its private return, and if the private return to R&D is equal to its output share (and if the assumptions underlying the calculation of MFP hold, notably perfect competition and constant returns to scale at the aggregate level) then the elasticity of MFP with respect to domestic business R&D should equal zero. A positive elasticity would signal the existence of spillovers.

Foreign business R&D is partly paid for by domestic business users, in the form of international payments for technology transfers (patents, licences and know-how contracts). However, such payments are relatively small in most countries (less than 0.4 per cent of GDP on average in OECD, including payments for software which are not taken into account directly in this analysis) and probably cover only a small fraction of all the benefits that accrue to users: the international market for technology is still very incomplete. Being treated by national accounts as intermediate consumption, payments for technology, be it to domestic or to foreign suppliers, are not accounted for as such in GDP, hence in MFP. For instance, increased payments for foreign technology will not impact *directly* on the level of GDP, as it is not considered as value added. Hence, the effects we will capture are only spillovers, the portion of the benefits for which users do not pay. In general, business users do not fully compensate government for the benefits from public R&D. Hence, most of its effect on business activity is spillovers.

As a consequence, this model captures most of the effect of public and foreign R&D but only the excess private return added to the public effect of business R&D. A further caveat is that the assumptions used for calculating MFP may not hold totally: increasing returns to scale and imperfect competition are often associated with R&D (*e.g.* Romer, 1990). If that is the case, the MFP index that we explain is subject to measurement errors that might be correlated with the right-hand-side variables. In order to mitigate this problem, we conducted estimates with instrumental variables.

We now look at some descriptive statistics. Table 1 reports compound annual growth rates for all the variables and countries, over the 1980-98 period. MFP growth ranges from 0.3 per cent a year in Germany to 3.4 per cent in Ireland. Most countries, however, are very close to 1 per cent a year (ten countries are between 0.9 per cent and 1.4 per cent). MFP growth, as well as R&D growth, is high for Ireland as this country has been catching up over this period. Business R&D (capital stock) growth ranges from 1.9 per cent (United Kingdom) to 8.9 per cent (Finland) and even 10.8 per cent for Ireland, with most countries around 4 per cent to 7 per cent. In most countries the growth of business R&D has been higher in the 1980s than in the 1990s (see Guellec and Ioannidis, 1999, for an analysis of the “levelling off” of R&D). Foreign R&D growth rates fluctuated around 4 per cent for all countries except Ireland where it was about 7 per cent a year, on average. In most coun-

Table 1. **Descriptive statistics: average annual growth rates, 1980-98**

Country	Per cent			
	Business R&D capital stock	Foreign R&D capital stock	Public R&D capital stock	MFP growth
AUT	7.50	3.80	3.69	0.84
BEL	4.07	4.19	2.11	1.34
CAN	6.71	3.84	2.46	0.69
DNK	7.08	3.41	4.23	1.02
FIN	8.86	5.11	5.86	2.60
FRA	3.80	4.10	3.45	1.05
GER	3.62	3.71	2.41	0.30
IRL	10.76	7.15	3.35	3.39
ITA	4.83	3.92	4.18	1.08
JPN	6.31	3.56	3.71	0.94
NLD	2.66	4.27	2.68	1.05
NOR	5.41	4.34	3.32	1.08
ESP	4.40	4.41	1.95	1.38
SWE	5.79	4.27	4.25	1.20
GBR	1.90	4.21	1.83	1.03
USA	3.66	4.47	2.04	0.94

tries, the growth of public performed R&D was much lower than that of business R&D over this period. It ranges from 1.9 per cent (United Kingdom) to 6.6 per cent (Finland), with most countries reporting around 1.8 per cent to 5.9 per cent. The major reasons for that are the end of the cold war (reduced defence spending) and strained budgetary conditions in many countries, as well as intensified efforts of business in this area.

Simple analysis of correlation between the average growth rates (1980-98) of these variables is reported in Table 2. MFP is quite highly correlated with business R&D and with foreign R&D, which are the two variables expected to have the more direct relationships. It is also positively correlated with public R&D, although the relationship is weaker. Business R&D is well correlated with the other two R&D variables. Foreign R&D is not correlated with public R&D and there is no reason to expect such a relationship. The positive correlation between foreign and business R&D can be explained as follows: foreign R&D is a weighted average of

Table 2. **Correlation matrix between average annual growth rates for 16 countries, 1980-98**

	Business R&D	Foreign R&D	Public R&D
MFP	0.675	0.909	0.383
Public R&D	0.622	0.094	
Foreign R&D	0.528		

other countries' R&D, with the weights reflecting technological proximity. As a country expands its R&D expenditures, it is likely to broaden the range of technologies it covers, thus increasing its correlation with other countries' specialisation. Such a mechanism applies especially to countries starting from a relatively low technological level, where the range of technologies covered is quite limited (Ireland is a case in point).

ESTIMATION RESULTS

We run estimates (see Table 3) of the error correction model (2).⁴ Different ranges of estimates were conducted for checking robustness: one constraining the short-term parameters to be the same for all countries; another allowing short-term parameters to vary across countries; SURE and 3SLS procedures have been used; and the model has been estimated over different sub-periods.⁵ There are no significant differences between the parameters estimated with these various techniques, denoting the robustness of our estimates. Lags for the long-term relationships have been set at two years for business and foreign R&D, and three years for public R&D. The choice has been made by testing different lags and choosing those delivering the most significant results. The longer lag for public R&D is consistent with the view that it is more basic than business R&D, and takes more time to affect productivity. Anyway in an ECM, in which long-term coefficients reflect relationships *in level*, the choice of lags is not as important as in other types of models. For detecting possible outliers, and the robustness of the results with respect to the sample of countries, the model was estimated on 16 sub-samples of 15 countries, which means that each country was dropped in turn. The figures reported in Appendix 3 further support the robustness of our estimates. In all cases the coefficients remain significantly different from zero.

The long-term elasticity of MFP with respect to business R&D is 0.13 (Table 3, column 1).⁶ Such an elasticity is quite in line with estimates reported in the literature (Nadiri, 1993), although it is in the low range. As the direct impact of business R&D on output is at least already partly accounted for in MFP, this positive coefficient must capture mainly spillovers and possibly extra return (coming in addition to normal remuneration of capital and labour) arising from R&D. It should be compared with the ratio of business R&D on business GDP (around 2 per cent in the OECD over the 1980s and 1990s). The social return to business R&D therefore is much higher than the "normal private return" (reflected in the income share of R&D).

Crossing the elasticity of business R&D with a time trend (Table 5, column 6) shows that there has been a growing impact of business R&D on MFP over time (an increase of about 0.005 a year). This finding confirms the impression given by

Table 3. Multi-factor productivity estimation results, error correction model

Dependent variable Δ MFP		80-98	80-98	80-98	80-96	84-98
		3SLS	SURE	SURE	3SLS	3SLS
Regressions		1	2	3	4	5
Multi-factor productivity growth (t-1) Δ MFP		-0.396* (-7.81)	-0.088* (-3.41)		-0.419* (-13.17)	-0.370* (-30.20)
Business R&D growth (t-1)	Δ BRD	-0.024 (-1.12)	-0.019 (-1.13)	-0.010 (-0.72)	-0.046* (-2.43)	0.024* (2.78)
Foreign R&D growth (t-1)	Δ FRD	0.055* (2.93)	0.069* (4.14)	0.042* (2.40)	0.044* (2.94)	0.125* (20.34)
Public R&D growth (t-2)	Δ PRD	0.091 (2.66)	0.067* (2.32)	0.041 (1.54)	0.073* (2.23)	0.125* (8.96)
MFP level (t-2)	LMFP	-0.205* (-11.20)	-0.181* (-13.04)	-0.162* (-10.88)	-0.211* (-13.19)	-0.192* (-29.47)
Business R&D (t-2)	LBRD	0.027* (5.28)	0.024* (5.17)	0.024* (5.52)	0.029* (5.82)	0.022* (6.11)
Foreign R&D (t-2)	LFRD	0.094* (7.74)	0.079* (7.83)	0.067* (6.67)	0.090* (8.32)	0.127* (26.85)
Public R&D (t-3)	LPRD	0.035* (5.12)	0.028* (4.20)	0.029* (4.70)	0.025* (4.34)	0.035* (16.28)
Control variables						
Employment rate growth (t)	Δ U	0.380* (8.95)	0.372* (11.05)	0.338* (9.44)	0.376* (10.17)	0.378* (39.41)
German unification dummy (t)	G	-0.100* (-20.78)	-0.096* (-28.63)	-0.097* (-26.94)	-0.099* (-23.30)	-0.094* (-52.81)
Country-specific short-term effects		No	No	Yes	No	No
Number of countries		16	16	16	16	16
Adjusted R-squared		0.501	0.477	0.477	0.525	0.505
Number of observations		302	302	302	272	238

Notes: Panel data, 16 countries, 1980-98. All regressions include country-specific intercepts (within estimates) and time dummies. The SURE estimation method (seemingly unrelated regression equations) corrects for the contemporaneous correlation of the error term across countries and the 3SLS method (three-stage least squares) corrects for the presence of the lagged endogenous variable among the right-hand side variables. * indicates the parameters that are significant at a 5 per cent probability threshold.

The instrumental variables for the 3SLS (three-stage least squares) estimates are all the exogenous variables (including dummies) and the endogenous variables (lagged two years).

The long-term coefficients as mentioned in the main text are obtained as follows (in accordance with the error correction model as developed in the section on estimation results): For variable X (X = BRD, FRD or PRD), divide the estimated coefficient for LX by the opposite of estimated coefficient for LMFP in the same regression. For instance, in the first regression (column 1), the long-term coefficient for business R&D is: $0.024/0.180 = 0.13$.

business reporting that R&D is an increasingly important activity for firms in the knowledge-based economy: when firms in most OECD countries are now at the technological frontier (after several decades of catching up), keeping pace with competition implies not only to build physical capacities, but increasingly to innovate (OECD, 2000).

Table 4. Long-term elasticities of output with respect to R&D variables

	Business R&D	Foreign R&D	Public R&D
Long-term elasticities	0.132	0.459	0.171

Further estimates allow the identification conditions that enhance or reduce this elasticity (Table 5, column 1). A country's business R&D intensity (the ratio of business R&D expenses on business GDP) has a positive effect on the elasticity of business R&D: a further percentage point in a country's R&D intensity increases its elasticity by 0.003 to 0.004. This finding points to some kind of increasing returns from investment in research. By spending more on R&D, businesses in a country are able to reap internal economies of scale, to set up networks, to benefit from each other's discoveries. It also denotes an improved ability to absorb the domestic knowledge generated by other firms and/or industries.

The share of government funding has a negative effect on the elasticity of business R&D, although it is small (Table 5, column 1). However, only the defence-related part of public funding has a significant negative effect on MFP (Table 5, column 2). Only four or five OECD countries have a substantial defence R&D budget and might be concerned by this problem. Actually, public funding with a civilian objective has a (weak) positive effect on the elasticity of business R&D. As this elasticity mainly captures spillovers, this might indicate that government funding is fairly successful in enhancing business R&D with higher social return. This is all the more possible as part of government funding of civilian business R&D is related to health or the environment, with no direct impact on measured MFP.

The long-term elasticity of foreign R&D on productivity is in the range of 0.45 to 0.5. This figure may seem surprisingly high, as this is essentially low-cost technology for the economy (the direct cost of absorbing new technology *when the domestic conditions are right* must be substantially lower than the cost of inventing it, which is the *raison d'être* for technology transfers). Estimates by Coe and Helpman (1995), although lower, are in the same order of magnitude: 0.29. This is high also as compared with the elasticity of domestic R&D reported above, leading to the conclusion that for any one country, other countries' R&D matter more than domestic R&D for the purpose of productivity growth, provided that the country has the capacity to absorb technology from abroad. This result is consistent with the fact that the domestic social return on R&D is higher than the private one: if technology spillovers occur within countries, there is no reason for it to stop at the border, and international spillovers should occur. As any country is small as compared with the whole OECD (or: the share of any country in new knowledge

Table 5. **Multi-factor productivity estimation results: error correction model and interactions**

Dependent variable	Δ MFP 1980-98					
	1	2	3	4	5	6
Regressions						
Multi-factor productivity level (t-2) LMFP	-0.193* (-9.74)	-0.208* (-9.45)	-0.206* (-11.68)	-0.210* (-11.66)	-0.241* (-13.28)	-0.214* (-10.60)
Business R&D (t-2) LBRD	0.020* (3.65)	0.017* (2.94)	0.026* (5.25)	0.019* (3.78)	0.018* (3.27)	0.016* (2.91)
R&D intensity (IRD) * LBRD (t-2)	0.044* (2.96)	0.067* (4.49)				
Share of public funding * LBRD (t-2)	-0.002* (-2.17)					
Defence share of public funding * LBRD (t-2)		-0.011* (-4.83)				
Civilian share of public funding * LBRD (t-2)		0.003* (2.17)				
Trend * LBRD (t-2)						0.001* (2.79)
Foreign R&D (t-2) LFRD	0.088* (7.42)	0.096* (7.56)	0.159* (4.63)	0.092* (7.51)	0.107* (7.94)	0.080* (5.71)
Log (average DPI) * LFRD (t-2)			-0.003* (-2.09)			
IRD * LFRD (t-2)				0.395* (4.34)		
Trend * LFRD (t-2)						0.001 (0.93)
Public R&D (t-3) LPRD	0.033* (4.83)	0.032* (4.45)	0.026* (3.71)	0.024* (3.92)	0.039* (4.30)	0.041* (5.80)
Business R&D intensity * LPRD (t-3)					0.049* (4.30)	
Defence as a percentage of GBOARD * LPRD (t-3)					-0.003* (-3.65)	
Higher education as a percentage of public * LPRD (t-3)					0.004* (4.19)	
Trend * LPRD (t-3)						-0.001* (-3.80)
Adjusted R-squared	0.519	0.532	0.508	0.513	0.538	0.502
Number of observations	302	297	302	302	298	302

Note: Panel data, 16 countries, 1980-98. All regressions include country-specific intercepts (within estimates) and time dummies, the short-term parameters and control variables are not reported for the sake of space. The 3SLS method (three-stage least squares) corrects for the possible simultaneity of the left-hand-side variable and certain of the right-hand-side variables. * indicates the parameters that are significant at a 5 per cent probability threshold.

generated by the 16-country panel is small), the benefits from other countries may dwarf those arising from domestic technology.⁷

A first straightforward deduction from this interpretation is that the impact of foreign spillovers on productivity should be larger for small countries than for large ones (the world abroad is even more important for smaller countries than for larger ones). The number of researchers is lower the smaller a country is. Hence the probability that the colleagues with whom you interact are located abroad is higher when you are from a small country. This is confirmed in Guellec and van Pottelsberghe (2001a), who use patent data for showing that smaller countries have a higher share of their inventions that involve co-operation with other countries (as opposed to inventions made by domestic inventors only). This size effect might be compensated by specialisation, as researchers interact mainly with colleagues *working in a related scientific field*: a small but highly specialised country may be as intensive as larger ones in the fields it covers, but the number of fields it covers may be lower ("specialisation effect"). We tested the "size effect" hypothesis by interacting foreign R&D with an indicator of size for each country: the average over the 1980-98 period of (log) GDP (results are reported in Table 5, column 3). The negative and significant parameter confirms that smaller countries do benefit more from foreign R&D than larger ones, although the effect is quite small. Hence, the "specialisation effect" also is strong, although it does not entirely cancel out the size effect. In addition to previous findings that R&D is more internationalised in smaller countries, the present study shows that smaller countries also benefit more, in terms of productivity, from such internationalisation.

A second straightforward deduction from the interpretation above is that the higher the R&D intensity of a country, the more it should benefit from foreign R&D. We tested this hypothesis by interacting foreign R&D with business R&D intensity for each country.⁸ The results presented in column 4 show that the impact of domestic R&D intensity on the elasticity of foreign R&D is positive and significant: a 0.1 per cent difference in R&D intensity between two countries generates a spread of about 0.002 between their elasticities. If firms from a country want to take full advantage from international spillovers, they have to spend on R&D: the free rider approach clearly does not work. It is clear that any firm intending to adopt or improve the knowledge generated by other firms or public institutions (be they domestic or foreign) will have to invest in "imitative" or "adaptive" research activities. This argument is forcefully stated by Geroski (1995) and has some empirical validation. For instance, the econometric studies of Cohen and Levinthal (1989) and Branstetter and Sakakibara (1998) illustrate that a firm's own R&D activity enhances its absorptive capacity of R&D results generated by other firms. Furthermore, the survey results of Mansfield (1981) show that imitation costs on average are about 65 per cent of the original innovation costs. Finally, the inter-

action of foreign R&D with a time trend (Table 5, column 6) shows that there is no significant increase over time in the elasticity of foreign R&D.

The long-term elasticity of government and university-performed research on productivity is around 0.17. This is much higher than the ratio of public-performed research on business GDP (0.7 per cent to 0.9 per cent in the 1980s and 1990s in OECD), which tends to show that overall public R&D is very valuable to the economy.⁹ The elasticity of public research is higher when the business R&D intensity of the economy is higher: this shows the importance of the business sector being able to seize opportunities raised by public research. Therefore, part of the effect of public research on productivity is indirect, flowing through the use of its discoveries by the business sector. Stronger links between public and private research, which governments in most OECD countries are trying to build, should enhance this effect.

The elasticity of public R&D is positively affected by the share of universities (as opposed to government laboratories) in public research (Table 5, column 5). This may point to the fact that much government performed R&D is aimed at public missions that do not impact directly on productivity (health, environment), whereas universities are providing the basic knowledge that is used in later stages by industry to perform technological innovation. This is confirmed by the negative effect of the share of defence in public R&D budgets, as it is not the main purpose of defence R&D to increase productivity. Another possible explanation for the higher impact of university research has to do with the way funds are allocated: in most OECD countries, at least part of funds for university research are allocated on a project basis, whereas government laboratories have an institutional funding. The former allows more reactivity to changing technological priorities than the latter (dropping technological lines which turn out to offer little opportunities, switching to promising areas), and may have a bigger impact on productivity. More case studies would be necessary to substantiate this assumption. Finally, the trend of the elasticity of public research over time is negative. This is at odds with trends in business and foreign research. This is also surprising, as the share of defence, which has a negative impact, has tended to decrease over time. One explanation may be that in many countries the public research sector has been slow to engage in new technology areas, especially ICT, which have spurred MFP growth in the recent years. This lack of flexibility could have contributed to the decreasing impact of public research on productivity.

The two control variables (for the business cycle and for German unification) are of the expected sign and are significant. The employment rate has a large and positive impact on productivity growth, which confirms previous findings that productivity is essentially pro-cyclical. The German unification dummy takes account of the sharp drop in average productivity in Germany following the 1990 events.

TENTATIVE POLICY CONCLUSIONS

One must be careful in drawing policy conclusions from such an exercise, which is performed at a very aggregated level and shows only OECD-wide averages over almost two decades: any policy lesson should then be confirmed by more detailed and country-level studies. Overall, the study points to the importance of technology for economic growth, be it developed by business, by the public sector or coming from foreign sources. It also shows the strong interactions between the various channels and sources of technology, which underline the necessity for government of having a broad and coherent policy approach:

- *Doing R&D is important for productivity and economic growth.* Business R&D has high spillover effects, it enhances the ability of the business sector to absorb technology coming from abroad or from government and university performed research. The social return on business R&D is then higher than its private return, which is a possible justification for some sort of government support to business R&D.
- *Governments should provide appropriate funding of R&D performed in the public sector, in particular the higher education sector, which has a substantial impact on economic growth in the long run.* The lower impact of research performed in government laboratories compared with research performed in the higher education sector points to the need for reviewing the way research is funded in the government sector (in relation to the way the research agenda is set and performance is monitored). However, as these institutional arrangements differ substantially across countries, country-specific studies would be needed for drawing more robust conclusions.
- *The effect of public performed R&D on productivity depends on the intensity of the business R&D effort.* Actually, business research develops technologies that in many cases have been first explored by the public research. It is therefore important that government provide the right framework for encouraging solid relationships between public and private research, so that knowledge flows more easily between the two sectors.
- *Governments should ensure the openness of their country to foreign technology, through flows of goods, of people or of ideas, and ensure that firms have the absorptive capabilities needed for making the best of foreign technology.* As countries that spend more on R&D also take more advantage of foreign technology, free riding (*i.e.* by waiting for other countries to develop the new technology and just trying to imitate when it is ready) would be ineffective.

NOTES

1. The role of information and communication technology in recent growth has been investigated in Schreyer (2000).
2. The most direct and visible effect of research in defence is to capture resources that could be devoted to more economically productive use, although defence may contribute to supporting the institutional framework that is conducive to technical change, something which escapes from direct measurement.
3. Eaton and Kortum (1999) show that, except for small countries very near the source of information, trade is not the major conduit for the spread of new technology. Their results suggest that benefits from innovation spread primarily through the transmission of ideas themselves, rather than through the export of goods embodying them.
4. We also performed static regressions, in log-level, and in growth rate, in order to check the robustness of the results from the ECM: results are reported in Appendix 2. They support results of the dynamic model presented below.
5. SURE: seemingly unrelated regression, allows to control for contemporaneous shocks affecting the 16 countries. 3SLS (three-stage least squares) allows to control for contemporaneous shocks and for the presence of the lagged dependent variable among the right-hand-side variables. Hence, 3SLS controls for potential simultaneity biases, due to the possible influence of the dependent variable on certain of the right-hand side variables. Instruments for the 3SLS regressions are all the right-hand-side variables (including dummies) and the left-hand-side variable lagged two years.
6. The derivation of long-term elasticities (as reported in the text and in Table 4) from the estimated coefficients (as reported in Tables 3 and 5) is detailed in note 3 of Table 3.
7. A simple thought experiment helps realising the importance of world technology for any particular country: imagine that all countries suddenly disappear, except for one, left untouched; this country would have then to rely exclusively on its own research for advancing its technology. One can easily imagine how dramatic the slowdown in productivity growth would be in this country. In a way, certain countries were in a such a situation (think for instance of Albania until a decade ago), with highly visible effects on their technology span and productivity level.
8. When we introduce simultaneously the average size and R&D intensity (both interacting with foreign R&D), size is not significant any more. We decided to introduce them separately into the model for two reasons. First, it seems that there is a correlation between size and R&D intensity among the countries included in the present analysis. This does not mean that small countries are in general more R&D intensive than large ones (a systematic negative relationship between size and R&D intensity). This is the case with our sample of countries because the small countries for which all the data

were available are generally intensive in R&D (*e.g.* Sweden, Finland, the Netherlands). To state it another way, there is a tendency for small countries with low R&D intensity not to have as much data available as large countries with low R&D intensity. Hence, the latter were included in our sample, whereas the former were excluded. The second reason is more technical. The R&D intensity variable varies across countries and over time, whereas the size variable (average GDP over the 1980-98 period) is fixed over time (this choice was made in order to avoid endogeneity of the right-hand-side variable). Therefore the former variable has a much higher variance, which “secures” its significance as compared with the latter variable.

9. In the absence of spillovers, the elasticity of any factor (which reflects its marginal productivity) should be equal to its income share.

Appendix 1

CALCULATION AND DATA SOURCES OF THE TECHNOLOGICAL VARIABLES

R&D capital stocks

R&D capital stocks are calculated following the perpetual inventory method. The stock at time t is equal to the new investment at time t plus the stock at time $t-1$ minus depreciation:

$$R_t = r_t + (1 - \delta)R_{t-1} \quad (\text{A1.1})$$

$$R_t = r_t + (1 - \delta)r_{t-1} + (1 - \delta)^2 r_{t-2} + (1 - \delta)^3 r_{t-3} + \dots \quad (\text{A1.2})$$

To construct the initial stock we assume a constant annual rate of growth of the past investments,

$$R_t = r_t + (1 - \delta)\lambda r_t + (1 - \delta)^2 \lambda^2 r_t + (1 - \delta)^3 \lambda^3 r_t + \dots \quad (\text{A1.3})$$

$$R_t = \frac{r_t}{1 - \lambda(1 - \delta)} \quad (\text{A1.4})$$

where

R_t = R&D capital stock at time t .

r_t = R&D investment at time t .

δ = Depreciation rate (constant over time).

$$\lambda = \frac{1}{1 + \eta} \text{ and } \eta \text{ is the mean annual rate of growth rt.}$$

The same formula has been used to calculate the business R&D capital stock (BRD) and the public R&D capital stock (PRD).

Foreign R&D capital stock

FRD is the foreign R&D capital stock calculated as the weighted sum of the domestic R&D capital stocks of 15 industrialised countries, the weights being the technological prox-

imity between pairs of countries. The technological proximity is computed as in Jaffe (1986, 1988) using patents granted by the USPTO:

$$FRD_i = \sum_{j \neq i} \omega_{ij}^{M3} \cdot BRD_j \quad i, j = 1, \dots, 16 \text{ industrialised countries.} \quad (A1.5)$$

$$\omega_{ij} = \frac{F_i F_j'}{F_{ij} F_{ji}} \quad F_i = \left| \frac{P_i^{TC1}}{\sum_{z=1}^{50} P_i^{PCz}} \quad \frac{P_i^{TC50}}{\sum_{z=1}^{50} P_i^{PCz}} \right|$$

F_i is the frequency distribution across 50 technological classes of patent granted by the USPTO to country i . The weight that are used (ω^{M3}) to compute the foreign R&D capital stock are a three-year moving average of ω .

Appendix 2
REGRESSIONS IN LOG-LEVELS

Robustness of regressions of the ECM can be assessed by estimating regressions in log-levels – according to equation (1'). As such a model misses the dynamics of the linkages between the variables, the purpose is primarily to look for simple, static relationships. Results are reported in Table A1. In columns 1 to 5 we progressively extend the range of variables in the regression (variables of interest and control variables). The estimated coefficients for all variables of interest are of the expected sign and are significant. The coefficient for business R&D is reduced as new variables are introduced into the regression. It drops from 0.2 to 0.1 when all variables are there. The coefficient for foreign R&D is 0.4, which may look high; explanations for that are reported in the main text. Column 6 reports regression results in growth rate (or first logarithmic difference), which can be seen as a test of robustness of the estimated parameters and specification. The coefficient associated with business R&D is not substantially different from the estimation in log-levels. The coefficient for foreign R&D is still significant, but it is much lower than in the regression in levels: this may reflect a dynamic adjustment that is different for this variable. The impact of public R&D is no longer significant. These estimates in growth rate capture short-term variation from a long-term equilibrium relationship. This non-significant estimated parameter may reflect the fact that public research has essentially a long-term impact on MFP growth.

Table A1. Multi-factor productivity estimation results, in log-levels

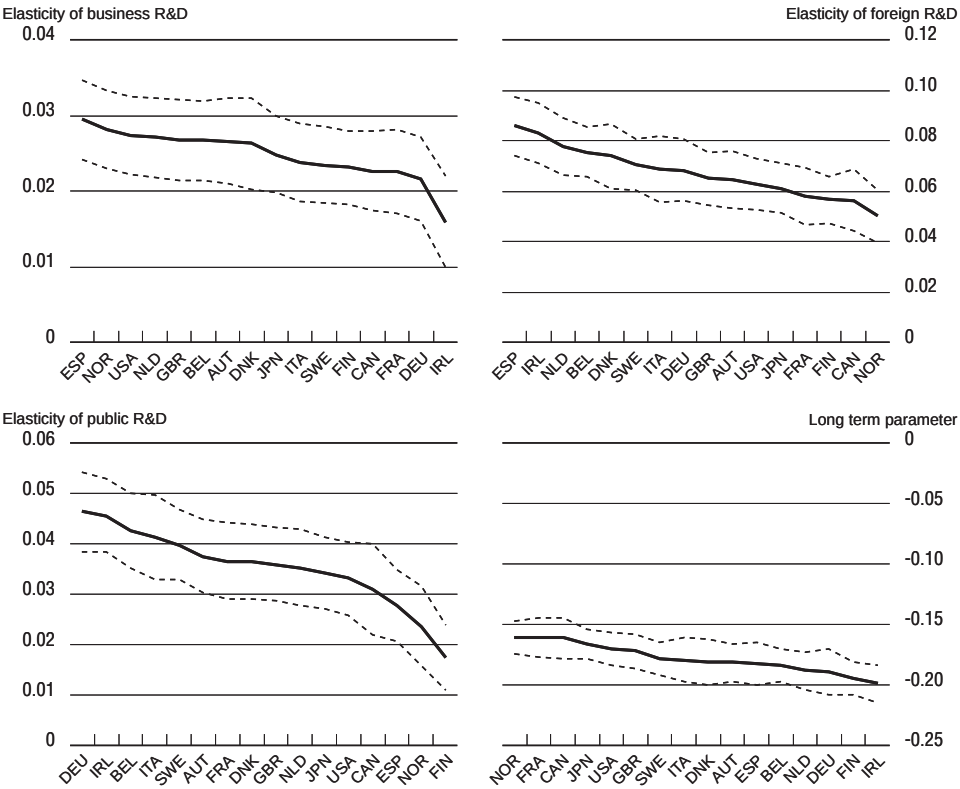
Dependent variable		LMFP	LMFP	LMFP	LMFP	LMFP	ΔMFP
Regressions		1	2	4	5	6	
Business R&D (t-1)	LBRD	0.208*	0.168*	0.127*	0.104*	0.087*	
		(150.8)	(72.2)	(74.27)	(48.11)	(7.11)	
Foreign R&D (t-1)	LFRD			0.385*	0.410*	0.049*	
				(42.39)	(35.64)	(3.01)	
Public R&D (t-2)	LPRD				0.083*	0.015	
					(11.76)	(0.77)	
Control variables							
Employment rate growth (t)	ΔU	1.382*	1.448*	1.156*	1.295*	0.143*	
		(53.21)	(39.03)	(36.96)	(38.98)	(3.76)	
German unification dummy (t)	G	-0.076*	-0.078*	-0.074*	-0.075*	-0.099*	
		(-20.40)	(-20.66)	(-27.29)	(-26.74)	(-26.58)	
Country dummies		Yes	Yes	Yes	Yes	no	
Time dummies		No	Yes	Yes	Yes	Yes	
Adjusted R-squared		0.839	0.835	0.892	0.896	0.274	

Note: Panel data, 16 countries, 1980-98, 302 observations. All regressions include country-specific intercepts (within estimates) and time dummies. The estimation method is SURE (seemingly unrelated regression equations) that corrects for the contemporaneous correlation of the error term across countries. * indicates the parameters that are significant at a 5 per cent probability threshold.

Appendix 3

STABILITY OF THE ESTIMATED PARAMETERS

The following four figures illustrate the stability of the estimated parameters of equation (2) when one out of the 16 countries is withdrawn from the panel.



BIBLIOGRAPHY

- ADAMS, J. (1990),
“Fundamental stocks of knowledge and productivity growth, *Journal of Political Economy* 98(4), pp. 673-702.
- BASSANINI, A., S. SCARPETTA and P. HEMMINGS (2000),
“Economic growth: the role of policies and institutions – panel data evidence from OECD countries”, *OECD Economics Department Working Papers*, No. 283.
- BRANSTETTER, L. and M. SAKAKIBARA (1998),
“Japanese Research Consortia: A microeconomic analysis of industrial policy”, *Journal of Industrial Economics* 46(2), June 1998, pp. 207-33.
- BROOKS, H. (1994),
“The relationship between science and technology”, *Research Policy* 23, pp. 477-486.
- COE, D.T. and E. HELPMAN (1995),
“International R&D spillovers”, *European Economic Review* 39, pp. 859-887.
- COHEN, W. and D. LEVINTHAL (1989),
“Innovation and learning: the two faces of R&D, *Economic Journal* 99, pp. 569-596.
- EATON, J. and S. KORTUM (1999),
“International technology diffusion: theory and measurement”, *International Economic Review* 40(3), pp. 537-570.
- GEROSKI, P.A. (1995),
“Do spillovers undermine the incentive to innovate?”, in S. Dowrick (ed.), *Economic Approaches to Innovation*, Edward Elgar, Aldershot, UK, pp. 76-97.
- GUELLEC, D. and E. IOANNIDIS (1999),
“Causes of fluctuations in R&D expenditures: a quantitative analysis, *OECD Economic Studies* 29, pp. 123-138.
- GUELLEC, D. and B. VAN POTTELSBERGHE DE LA POTTERIE (1999),
“Does government support stimulate private R&D?”, *OECD Economic Studies* 29, pp. 95-122.
- GUELLEC, D. and B. VAN POTTELSBERGHE DE LA POTTERIE (2001a),
“The internationalisation of technology analysed with patent data”, *Research Policy* (forthcoming).
- GUELLEC, D. and B. VAN POTTELSBERGHE DE LA POTTERIE (2001b),
“The effectiveness of public policies in R&D”, *Revue d'Économie Industrielle*, 94(1), pp. 49-68.
- JAFFE, A.B. (1986),
“Technological opportunity and spillovers of R&D: evidence from firm's patent, profits and market value”, *The American Economic Review* 76(5), pp. 984-1001.

- JAFFE, A.B. (1988),
"Demand and supply influences in R&D intensity and productivity growth", *Review of Economics and Statistics* 70(3), pp. 431-437.
- LICHTENBERG, F. and B. VAN POTTELSBERGHE DE LA POTTERIE (1998),
"International R&D spillovers: a comment", *The European Economic Review* 42(8), pp. 1483-1491.
- MANSFIELD, E. (1981),
"Imitation costs and patents: an empirical study", *Economic Journal* 91, pp. 907-918.
- NADIRI, I. (1993),
"Innovations and technological spillovers", *NBER Working Paper Series No. 4423*, Cambridge, Ma.
- OECD (1993),
The Measurement of Scientific and Technological Activities: Standard Practice for Surveys of Research and Experimental Development – Frascati Manual 1993, OECD, Paris.
- OECD (1997),
Technology and Industrial Performance, OECD, Paris.
- OECD (2000),
A New Economy? The Changing Role of Innovation and Information Technology in Growth, OECD, Paris.
- OECD (2001),
The OECD Productivity Manual, OECD, Paris.
- POOLE, E. and J.T. BERNARD (1992),
"Defence innovation stock and total factor productivity growth", *Canadian Journal of Economics* 25(2), pp. 438-52.
- ROMER, P.M. (1990),
"Endogenous technical change", *Journal of Political Economy* 98, pp. 71-102.
- SCHREYER, P. (2000),
"The contribution of ICT to output growth: a study of the G7 countries", *STI Working Paper 2000/2*, OECD, Paris.
- SOLOW, R.M. (1957),
"Technical change and the aggregate production function", *Review of Economics and Statistics* 39, pp. 312-320.
- VAN POTTELSBERGHE DE LA POTTERIE, B. and F. LICHTENBERG (2001),
"Does foreign direct investment transfer technology across borders?", *Review of Economics and Statistics* 83(3) (forthcoming).

MEASURING PRODUCTIVITY

Paul Schreyer and Dirk Pilat

TABLE OF CONTENTS

Introduction	128
Measuring productivity growth.....	129
Gross output and value-added based productivity measures	129
Measuring output	136
Labour input	138
Capital input	142
Index numbers.....	146
Estimating productivity levels	147
Output, labour and capital input.....	148
Purchasing power parities for international comparisons.....	150
Estimates of income and labour productivity levels.....	152
The interpretation of productivity measures	157
The link with technological change	157
Productivity growth and changes in costs	159
The role of the business cycle in productivity growth.....	160
The difference between productivity and efficiency.....	160
How productivity at the industry level is related to that at the firm level	161
Innovation and productivity	162
In conclusion	163
Bibliography	166

National Accounts Division, Statistics Directorate, and Economic Analysis and Statistics Division, Directorate for Science, Technology and Industry, respectively. The authors are grateful for comments from Andrew Dean, Jørgen Elmeskov and Paul Swaim on an earlier draft.

INTRODUCTION

Productivity growth is the basis for improvements in real incomes and welfare. Slow productivity growth limits the rate at which real incomes can improve, and also increases the likelihood of conflicting demands concerning the distribution of income (Englander and Gurney, 1994). Measures of productivity growth and of productivity levels therefore constitute important economic indicators.

In principle, productivity is a rather straightforward indicator. It describes the relationship between output and the inputs that are required to generate that output. Despite its apparent simplicity, several problems arise when measuring productivity. These issues are particularly important for comparing productivity growth across countries, whether for the entire economy or for different industries, and for comparing productivity levels internationally. Some of these measurement difficulties are closely related to technological developments – currently of great interest. For example, assessing the role of information and communication technology (ICT) in productivity growth requires the construction of accurate price and quantity indices of ICT products that are internationally comparable. Other issues, such as the measurement of labour input, have been around for much longer, but remain important. The most important productivity measurement issues have recently been brought together in the *OECD Productivity Manual* (OECD, 2001a) and part of the discussion below draws on this manual.

There are many different measures of productivity growth. The choice between them depends on the purpose of productivity measurement and, in many instances, on the availability of data. Broadly, productivity measures can be classified as single-factor productivity measures (relating a measure of output to a single measure of input) or multi-factor productivity measures (relating a measure of output to a bundle of inputs).¹ Another distinction, of particular relevance at the industry or firm level is between productivity measures that relate gross output to one or several inputs and those which use a value-added concept to capture movements of output.

Table 1 uses these criteria to enumerate the main productivity measures. The list is incomplete insofar as single productivity measures can also be defined over intermediate inputs and labour-capital multi-factor productivity can, in principle, be evaluated on the basis of gross output. However, in the interest of simplicity, the table was restricted to the most frequently used productivity measures. These

Table 1. Overview of the main productivity measures

Type of output measure:	Type of input measure			
	Labour	Capital	Capital and labour	Capital, labour and intermediate inputs (energy, materials, services)
Gross output	Labour productivity (based on gross output)	Capital productivity (based on gross output)	Capital – labour MFP (based on gross output)	KLEMS multi-factor productivity
Value-added	Labour productivity (based on value-added)	Capital productivity (based on value-added)	Capital – labour MFP (based on value-added)	–
Single factor productivity measures			Multi-factor productivity (MFP) measures	

are measures of labour and capital productivity, and multi-factor productivity measures (MFP), either in the form of capital-labour MFP, based on a value-added concept of output, or in the form of capital-labour-energy-materials-services MFP (KLEMS), based on a concept of gross output. Among those measures, value-added-based labour productivity is the single most frequently computed productivity statistic, followed by capital-labour MFP and KLEMS MFP.

A full discussion of the entire set of productivity measures would go beyond the scope of this paper. The following pages will therefore highlight some key conceptual and measurement issues associated with comparisons of productivity growth and productivity levels over time and between countries. The sections on productivity growth and on the interpretation of productivity measures draw heavily on the *OECD Productivity Manual* (OECD, 2001a) to which the reader is referred for a more in-depth discussion. The paper primarily focuses on measurement and comparability issues, leaving the detailed analysis and interpretation of productivity to other papers in this issue of *OECD Economic Studies*. The next section explores measures of productivity growth and the theoretical foundations for these measures. The third section examines estimates of productivity levels, while the fourth section briefly discusses the interpretation of productivity growth and levels.

MEASURING PRODUCTIVITY GROWTH

Gross output and value-added based productivity measures

Every productivity measure, implicitly or explicitly, relates to a specific producer unit: an establishment, a firm, an industry, a sector or an entire economy. The goods or services that are produced within a producer unit and that become

available for use outside the unit are called (gross) output.² Output is produced using primary inputs (labour and capital) and intermediate inputs. This relation is normally presented as a production function H with gross output Q , labour input L , capital input K , intermediate inputs M and a parameter of technical change, A :

$$Q = H(A, K, L, M) \quad (1)$$

Technical change is called “Hicks-neutral” or “output augmenting” when it can be presented as an outward shift of the production function that affects all factors of production proportionately:

$$Q = H(A, K, L, M) = A \cdot F(K, L, M) \quad (2)$$

Differentiating this expression with respect to time and using a logarithmic rate of change, MFP growth (the rate of change of the variable A) is measured as the rate of change of volume output minus the weighted rates of change of inputs. For a cost-minimising producer, the weights attached to the rates of change of factor inputs correspond to the revenue share of each factor in total gross output:

$$\frac{d \ln A}{dt} = \frac{d \ln Q}{dt} - s_L \frac{d \ln L}{dt} - s_K \frac{d \ln K}{dt} - s_M \frac{d \ln M}{dt} \quad (3)$$

Here, MFP growth is positive when the rate of growth of the volume of gross output rises faster than the rate of growth of all combined inputs. When the assumption about factor-augmenting technical change holds, the so computed MFP measure can be interpreted as an index of disembodied technical change. This is the well-known Solow (1959) growth accounting model (Box 1; Box 2).

However, the gross output based approach provides few insights about the relative importance of a firm or an industry for productivity growth of a larger (parent) sector or of the entire economy, due to complications related to intra-industry deliveries. This is best explained by way of an example. Suppose that there are two industries: the leather industry, which only produces intermediate inputs for the shoe industry. By contrast, the shoe industry produces only final output. A productivity measure for the *aggregate shoe and leather industry* would need to address the following problem. Simple addition of the flows of outputs and inputs is not the right procedure for obtaining measures of output and input of the shoe and leather industry *as a whole*, since double counting of the intermediate flows between the leather and the shoe producer would result. These flows have to be netted out, such that the output of the integrated shoe and leather industry would consist only of the shoes produced, and integrated intermediate inputs consist only of the purchases of the leather industry and non-leather purchases of the shoe industry. This has important consequences for productivity measures. Take

Box 1. Growth accounting and links between productivity measures

The economic theory of productivity measurement goes back to the work of Jan Tinbergen (1942) and, independently, to Robert Solow (1957). They formulated productivity measures in a production function context and linked them to the analysis of economic growth. The field has developed considerably since and now offers a consistent approach that integrates the theory of the firm, index number theory and national accounts. Robert Solow's growth accounting approach identifies the contributions of different inputs to output growth. In its simplest form, where output Q is measured as deflated value-added and inputs are confined to primary inputs labour L and capital services K , the growth accounting equation can be stated as:

$$\frac{d \ln Q}{dt} = s_L \frac{d \ln L}{dt} + s_K \frac{d \ln K}{dt} + \frac{d \ln A}{dt}$$

In this expression, labour and capital each contribute to value-added growth and their contribution is measured as the rate of change of each input times its share in total costs. The change in value added that is not explained by these contributions is attributed to multi-factor productivity growth, captured by the variable A . Thus, the rate of change of A is measured residually, *i.e.* by subtracting the contributions of labour and capital from the rate of output growth. Another useful way of presenting the growth accounting equation is in terms of a decomposition of the rate of change of labour productivity. Labour productivity growth is measured as the difference between the rate of change of output growth and the rate of change of labour input growth, or as

$$\frac{d \ln Q}{dt} - \frac{d \ln L}{dt}.$$

Re-arranging the above expression gives a decomposition of the movement in labour productivity into two components. The first depicts the change in labour productivity due to capital deepening (labour productivity rises when more capital is used per worker) and the second shows the effects of MFP growth:

$$\frac{d \ln Q}{dt} - \frac{d \ln L}{dt} = (1 - s_L) \left(\frac{d \ln K}{dt} - \frac{d \ln L}{dt} \right) + \frac{d \ln A}{dt}$$

The following table presents this decomposition for the business sector of several OECD countries. It shows that capital deepening has played an important but not a dominant role in explaining value-added based labour productivity growth in the 1990s. This does not imply that investment has been unimportant in the process of growth. For instance, if output growth is driven both by capital and employment growth, capital intensity remains stable. Much depends also on the measurement of capital input (see discussion below). The present table is based on a measure of capital services, thereby reflecting the theoretically preferred approach towards the measurement of capital input.

Box 1. Growth accounting and links between productivity measures (cont.)**Productivity growth in the business sector**

(annual percentage change)

		Labour productivity	Of which:	
		(Value-added per hour worked)	Capital deepening	MFP growth
Australia	1990-99	2.5	0.6	1.9
	1995-99	2.9	0.9	2.0
Canada	1990-99	1.5	0.2	1.3
	1995-99	1.3	0.2	1.1
France	1990-99	1.8	1.1	0.7
	1995-99	1.5	0.7	0.8
Germany	1990-99	2.4	1.2	1.2
	1995-99	1.7	0.9	0.8
Italy	1990-99	1.9	0.9	0.9
	1995-99	0.7	0.6	0.1
Japan	1990-99	2.0	1.5	0.6
	1995-99	1.7	1.1	0.5
United States	1990-99	1.5	0.7	0.8
	1995-99	2.3	1.0	1.3

Source: Colecchia and Schreyer (2001).

the example where both the shoe and the leather producers' gross output-based MFP growth is 1 per cent. The simple (weighted) average of the shoe and leather producers' MFP growth will be 1 per cent. However, productivity growth of the integrated shoe and leather industry will be more than 1 per cent, because the shoe producer's productivity gains cumulate with those of the leather producer as the former buys inputs from the latter. In sum, it is difficult to compare gross output based MFP growth across different levels of aggregation, as aggregate MFP growth is not a simple weighted average of its component measures.

Double counting does not arise with value-added-based MFP growth.³ Here, productivity is measured as the ratio of value-added to an index of combined primary inputs (*i.e.* labour and capital), with both the numerator and denominator

Box 2. The econometric approach to productivity measurement

The econometric approach to productivity measurement is appealing for being based solely on observations of volume outputs and inputs. For example, it avoids postulating the relationships between production elasticities and income shares that are implied by cost minimisation but may not correspond to reality. Indeed, researchers are able to test these relationships. Further possibilities arise with econometric techniques: *i*) allowance can be made for adjustment cost (the possibility that changes in factor inputs are increasingly costly the faster they are implemented) and variations in capacity utilisation; *ii*) it is possible to investigate forms of technical change other than the Hicks-neutral formulation implied by the index number based approach; *iii*) there is no *a-priori* requirement to assume constant returns to scale of production functions. The literature about the econometric approach is large, and examples of integrated, general models can be found in Morrison (1986) or Nadiri and Prucha (2001).

All these possibilities come at a cost, however. These models raise complex econometric issues that sometimes put a question mark on the robustness of results. Often, researchers are constrained by modest sample sizes to revert to *a priori* restrictions (for example constant returns to scale) to increase the degrees of freedom for estimation. From the point of view of statistical offices responsible for the publication of regular productivity statistics, complex econometric approaches have little attractiveness because: *i*) updating involves full re-estimation of (systems of) equations; *ii*) methodologies are often difficult to communicate to a broad spectrum of users of productivity statistics; and *iii*) significant data requirements tend to reduce the timeliness of results.

Hulten (2001) points out that there is no reason why the econometric and the index number approach should be viewed as competitors and he quotes examples of synergies that proved particularly productive. Synergies arise in particular when econometric methods are used to further explain the productivity residual.

Overall, econometric approaches are a tool that is best suited for academically oriented, single studies of productivity growth. Their potential richness and testable set-up make them a valuable complement to the non-parametric, index number methods that are the normal tool for productivity statistics.

representing deflated (real) volumes. Value-added, which takes the role of the output measure, is gross output corrected for purchases of intermediate inputs. In terms of rates of change, real value-added can be defined⁴ as

$$\frac{d \ln VA}{dt} \equiv \frac{1}{s_{VA}} \left(\frac{d \ln Q}{dt} - s_M \frac{d \ln M}{dt} \right).$$

Here, S_{VA} stands for the current price share of value-added in gross output. Using equation (3) to substitute for the expression in parenthesis yields:

$$\frac{d \ln VA}{dt} = \frac{1}{s_{VA}} \left(s_L \frac{d \ln L}{dt} + s_K \frac{d \ln K}{dt} + \frac{d \ln A}{dt} \right) \quad (4)$$

Value-added-based MFP measures, evaluated residually as the difference between the rate of change of real value-added and the weighted rates of change of primary inputs labour and capital, can then be expressed as in (5), where $s_L^{VA} = s_L/s_{VA}$ is the labour share in value-added and $s_K^{VA} = s_K/s_{VA}$ is the capital share.

$$\frac{d \ln A^{VA}}{dt} = \frac{d \ln VA}{dt} - s_L^{VA} \frac{d \ln L}{dt} - s_K^{VA} \frac{d \ln K}{dt} \quad (5)$$

Value-added-based MFP growth will be positive if volume value-added grows faster than combined primary inputs. The advantage of the value-added measure is that aggregate value-added growth is a simple weighted average of value-added growth in individual industries, and so is value-added-based MFP growth. To stay with the above example, value added (at current prices) of the integrated shoe and leather industry is simply the sum of value-added in the shoe and the leather industry. A 1 per cent growth of value-added-based MFP in both the shoe and leather industry translates into a 1 per cent productivity growth of the shoe and leather industry as a whole. This makes value-added-based productivity measures comparable across different levels of aggregation and turns them into meaningful indicators for an industry's contribution to economy-wide productivity growth. Value-added is, however, not an immediately plausible measure of output: contrary to gross output, there is no physical quantity that corresponds to a volume measure of value added. Also, if the production model (2) is the "true" model of technical change, the value-added-based calculation will overstate⁵ the rate of technical change, as

$$\frac{d \ln A^{VA}}{dt} = \frac{1}{s_{VA}} \frac{d \ln A}{dt}.$$

That is, the value-added-based MFP measure equals the gross output-based measure times a scaling factor that corresponds to the inverted share of value-added in gross output. This share cannot exceed unity and consequently, the value-added-based MFP measure will always be at least as large as the gross output-based term.

Empirically, the choice of concepts matters, as Table 2 on productivity in the machinery and equipment industry in Finland demonstrates. The rate of change of the gross output-based MFP measure is 2.7 per cent over the 1990-98 period, com-

Table 2. **Value-added and gross output-based productivity measures: an example**
 Machinery and equipment industry, Finland
 Averages of annual rates of change (percentages)

	1990-98	1990-94	1994-98
Gross output (deflated)	10.1	4.2	16.0
Value added (deflated)	9.5	3.3	15.8
Labour input (total hours)	1.6	-3.7	6.9
Capital input (gross capital stock)	3.0	1.5	4.5
Intermediate inputs (deflated expenditure)	10.4	4.8	16.1
Share of value-added in gross output (current prices)	37.0	38.9	33.4
Gross output-based productivity (KLEMS MFP)	2.7	2.1	3.3
Value-added based productivity (Capital-labour MFP)	7.8	5.7	9.8

Source: OECD, based on STAN database.

pared with a 7.8 per cent rise in the value-added-based measure. Moreover, the two measures show quite different pictures in terms of the *acceleration* or *deceleration* of productivity growth between two periods, an indicator that is of significant importance to analysts as has been seen in the discussion about the “productivity slowdown” in the years after 1973 or the “productivity acceleration” in the United States in the late 1990s. In the Finnish example, the gross output-based measure rises from 2.1 per cent to 3.3 per cent per year between the first and the second half of the 1990s, or by 1.2 percentage points; meanwhile, the value-added measure rises from 5.7 per cent to 9.8 per cent, or by 4.1 percentage points.

In a closed economy, the difference between the two measures becomes smaller with a rising level of aggregation; at the level of the entire economy, the gross output-based productivity measure will equal the value-added based MFP measure. In an open economy, with imports from abroad, this is not the case and the two measures will continue to produce different results even at the macroeconomic level.

Different interpretations have also to be invoked with respect to gross output and value-added-based measures of *labour productivity*, both widely-used productivity indices. Growth in value-added-based labour productivity depends on shifts in capital intensity (the amount of capital available per unit of labour) and MFP growth. When measured as *gross output* per unit of labour input, labour productivity growth also depends on how the ratio of intermediate inputs to labour changes. A process of outsourcing, for example, implies substitution of primary factors of production, including labour, for intermediate inputs. Everything else equal, gross output-based labour productivity rises as a consequence of outsourcing and falls when in-house production replaces purchases of intermediate inputs,

despite the fact that such changes need not reflect changes in the individual characteristics of the workforce, nor shifts in technology or efficiency. By contrast, the growth rate of value-added productivity is less strongly affected by changes in the ratio between intermediate inputs and labour, or the degree of vertical integration. When outsourcing takes place, labour is replaced by intermediate inputs. In itself, this would raise measured labour productivity. At the same time, however, value-added will fall, and this offsets some or the entire rise in measured productivity.

Overall, it would appear that gross output and value-added based productivity measures are useful complements. When technical progress affects all factors of production proportionally, the former is a better measure of technical change. Value-added-based productivity measures compensate for the extent of outsourcing and provide an indication of the importance of the productivity improvement in an industry for the economy as a whole. They indicate how much extra delivery to final demand per unit of primary inputs an industry generates. When it comes to *labour productivity*, value-added based measures are less sensitive to changes in the degree of vertical integration than gross output-based measures. Practical aspects also come into play. Measures of value-added are often more easily available than measures of gross output although in principle, gross output measures are necessary to derive value-added data in the first place. Intra-industry flows of intermediate products must be accounted for in order to generate consistent sets of gross output measures and that may be difficult empirically.

Measuring output

Differences in the methodologies used to obtain quantity series of output can significantly affect productivity measures. Quantity indices of output are normally obtained by dividing a current-price series or index of output by an appropriate price index (*i.e.* by deflation). Only in a few instances are quantity measures derived by direct observation of volume output series.⁶ Measurement of volume output is therefore often tantamount to constructing price indices – a task whose full description far exceeds the scope of the present paper. Some of the more difficult issues associated with the deflation of output are nevertheless mentioned here.

Independence of measures of output from measures of input. An important precondition for the validity of productivity measures is that price and quantity indices of output should be constructed independently of price and quantity indices of inputs. Such dependence occurs, for example, when quantity indices of outputs are extrapolated using quantity indices for one or several inputs. The quantity indicators used are often inputs to the industry under consideration, in particular observations on employment.

In other instances, output-related measures are used to extrapolate real value-added. Though often imperfect, it is apparent that the implied bias for productivity measurement is less severe than in the case of input-based extrapolation. For example, Eldridge (1999) reports that, in the United States, the quantity indicator for auto insurance expenditure is the deflated value of premiums, where deflation itself is based on a component index of the CPI. In other instances, physical output data are used as the quantity indicator; the United States quantity indicator for brokerage charges is based primarily on BEA estimates of orders derived from volume data from the Security and Exchange Commission and trade sources (Eldridge, 1999).

From the perspective of productivity measurement, the independence of statistics on inputs and outputs is key. Using input-based indicators to construct output series generates an obvious bias in productivity measures; productivity growth will reflect whatever assumption about productivity growth was made by statisticians in constructing the output series (*e.g.* that labour productivity was unchanged). Occurrences of input-based extrapolation are concentrated in activities where market output prices are difficult to observe. For this reason, input-based extrapolation is more frequent and quantitatively more important for service industries than for other parts of the economy (see OECD, 1996*b* for a survey of methods in OECD countries) and can lead to biased productivity measures.

Quality change. The rapid development of information and communication technology products has brought to centre-stage two long-standing questions for the construction of price indices: how to deal with quality changes of existing goods and how to account for new goods.⁷ The distinction between these two issues is blurred because it is unclear where to draw the borderline between a “truly” new good and a new variety of an existing good.⁸

Typically, statistical agencies derive price indices for products by observing price changes of items in a representative sample. New products, quality changes and new variants are common phenomena in the observation of price changes of items and statistical offices have well-established procedures to deal with them.⁹ Unfortunately, these methods are not the same across countries and sometimes yield implausibly large differences. The most widely quoted case is price indices for information and communications technology products such as computers. Their prices decline by between 30 per cent per year in the United States, and about 5 per cent per year in a number of European countries. Given the homogeneity and international tradability of these products, it is likely that some of the differences are due to statistical methods rather than actual price developments. In the present context, the question arises: how much do these differences matter for comparisons of measures of output?

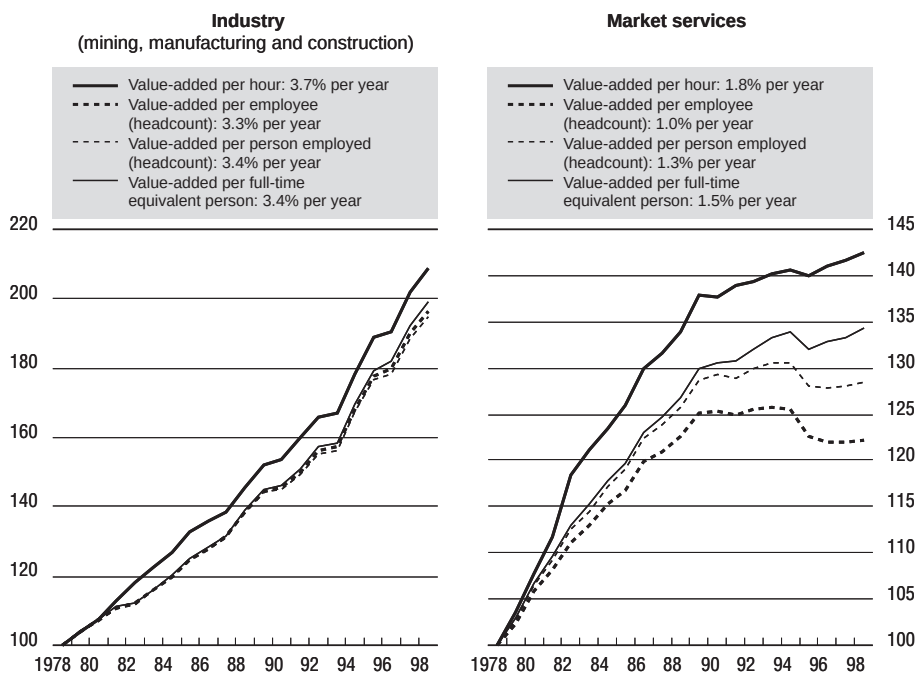
Empirically, the answer to this question depends largely on the level of aggregation at which the analysis is conducted. As shown in Schreyer (2001), the effects of a greater quality adjustment of ICT price measures tend to be comparatively small for the measurement of economy-wide productivity, and certainly not of a size to account for differences in measured productivity growth between countries. This is largely due to the fact that many ICT products are imported, and a different price measure not only affects measures of final consumption (and hence GDP) but also measures of imports, implying that some of the effects on measured GDP are offsetting. On the other hand, the effects on measured output volumes are without doubt significant for individual industries such as the office equipment and computer industry. Similarly, measures of individual demand components, in particular the volume of investment, may suffer from a lack of comparability unless similar methods are used between countries in their efforts to account for quality change in high-tech products. Measures of the volume of investment are of direct importance for productivity analysis as they are important elements in the construction of capital stock series (see section on capital input below).

Labour input

Different measures of employment. In the spirit of production theory, labour input for an industry is most appropriately measured as the quality-adjusted number of hours actually worked. The simplest, though least recommended, measure of labour input is a head count of jobs or employees. Such a measure fails to reflect changes in average work time per employee, multiple job holding (when the number of employees is the measure), self-employment and the quality of labour.

A first refinement to this measure is its extension to total employment, comprising both wage and salary earners, and the self-employed (including contributing family members). A second refinement is the conversion from simple job (or person) counts to estimates of total "hours actually worked". Rates of change of the number of persons employed differ from the rates of change of total hours worked when the number of average hours worked per person shifts over time. Such shifts may be due to a move towards more paid vacations, shorter "normal" hours for full-time workers and greater use of part-time work. Moreover, hours worked will also vary over the business cycle as labour demand rises and falls. These developments have taken place in many OECD countries and underline the importance of choosing "hours actually worked" as the variable for labour input in productivity measurement because it bears a closer relation to the amount of productive services provided by workers than simple head counts.

Figure 1. Labour productivity¹ based on different measures of employment in France



1. Output is measured as a quantity index of value-added.

Source: INSEE.

An example of the impact on labour productivity measures of choosing different measures for employment is given in Figure 1 below. For France, for the period 1987-98, labour productivity indices were calculated using total hours, the number of full-time equivalent persons, the number of employed persons (head counts) and the number of employees (head counts). Results are presented for industry (comprising mining, manufacturing and construction) and for market services. Not surprisingly, the productivity measures based on total hours rise significantly faster than those based on other employment measures. In manufacturing, moving from head counts to full-time equivalent measures hardly changes the productivity series. This is quite different for the service sector where part-time employment has grown rapidly in many countries and now plays an important role in total employment. Even more pronounced are the effects of including or excluding the self-employed in the service sector, as reflected by the differences

in productivity estimates based on total employment and based on the number of employees only.

Full-time equivalent jobs (or persons) are another variable sometimes used for measuring labour input. By definition, full-time equivalent employment is the number of total hours worked divided by average annual hours actually worked in full-time jobs. Conceptually, then, in full-time equivalent measures part-time employed persons are counted with a smaller weight than persons working full time. Consequently, the full-time equivalent measure should avoid the bias arising from a shifting share of part-time employment in the workforce but will not adjust for changes in the number of hours which constitutes a full-time job, *e.g.* as a consequence of changes in legislation or collective agreements. In addition, methodologies underlying the construction of full-time equivalent persons (or jobs) are not always transparent and may vary internationally. For example, in some cases full-time equivalents are based on crude estimates, such as counting each part-time job (often defined as any job with less than normal working hours) as half a full-time job.

Statistical sources. There are two main statistical sources for measures of labour input: household-based labour force surveys (LFS) and establishment or firm-based surveys (ES). LFS are typically conducted from a socio-economic perspective to provide reliable information about personal characteristics of the labour force, such as educational attainment, age, or the occurrence of multiple job holding, as well as information about the jobs (*e.g.* occupation and type of contract). Also, LFS have the advantage of full coverage of the economy. ES are conducted from a production perspective, and describe labour as an input factor. One distinguishing feature of establishment surveys is that they gather information on jobs rather than on persons employed, thus persons who have jobs in more than one establishment will be counted more than once. Another feature is that ES will often only cover a subset of all establishments in an industry, normally those above a certain size limit. If establishments included in the survey have systematically higher productivity levels than those excluded, productivity estimates based on ES will inadequately reflect the effects of the size composition in an industry.

In a few OECD countries (*e.g.* the Netherlands), statistical offices fully consolidate the two sources into a single, final set of labour accounts. In most countries, both sources are used to construct employment data for national accounts (NA). As such, these NA data qualify as the preferred source for productivity analysis. However, NA statistics often stop short of producing all the relevant labour input measures (in particular hours worked) or such variables are not available at the required sectoral detail. In such cases, multiple sources sometimes have to be combined, although this introduces the risk of not comparing like with like. One such example is the application of data on average hours worked per person,

based on LFS, to NA-based statistics on the number of persons employed. This may be acceptable for purposes of constructing measures of productivity *growth* but can create important non-comparabilities when measuring productivity *levels* (see below for a further discussion).

Skill composition of labour. Labour input reflects the time, effort and skills of the workforce. While data on hours worked capture the time dimension, they do not reflect the skill dimension. When total hours worked are the simple sum of all hours of all workers, no account is taken of the heterogeneity of labour. For the estimation of productivity changes, the question is whether, over time, the composition of the labour force changes, *i.e.* whether there is an increase or decrease in the average quality of labour input. By most measures, there has been a steady increase in the quality of labour (OECD, 1998a). An increase in the average quality of labour implies that a quality-adjusted measure of labour input would rise faster than an unadjusted measure of labour input. Successful quality-adjustment is tantamount to measuring labour in constant-quality units. In the context of productivity measurement, Jorgenson *et al.* (1987), Denison (1985) and the US Bureau of Labor Statistics (BLS, 1993) have tackled this issue

Measuring constant-quality labour input is interesting from several perspectives. First, it provides a more accurate indication of the contribution of labour to production. One recalls that MFP measures the residual growth in output that cannot be explained by the rate of change in the services of labour, capital and intermediate inputs. When quality-adjusted measures of labour input are used in growth accounting instead of unadjusted hours worked, a larger share of output growth will be attributed to the factor labour instead of the residual factor productivity growth. In other words, substituting quality-adjusted labour input measures for unadjusted measures can better identify the sources of growth, by distinguishing between externalities or spill-overs – captured by the productivity residual – and the effects of investment in human capital.

Second, a comparison of an adjusted and unadjusted measure of labour input yields a measure of the corresponding compositional or quality change of labour input. This can usefully be interpreted as one aspect of the formation of human capital, and is thus a step towards measuring the effects of intangible investment.

The theory of the firm stipulates that, under certain conditions (*i.e.* the firm is a price-taker on labour markets and minimises its total costs), labour of a certain type will be hired up to the point where the cost of an additional hour of labour is just equal to the additional revenue that using this labour generates. This equality implies that, for a measure of total labour input, the individual labour inputs of different quality can be weighted using their relative wage rates (or, more precisely, the shares of each type of labour in total labour compensation).

Even when labour input is differentiated only by a simple trait, such as occupation, information requirements are severe: data are needed that distribute the number of total hours worked across different occupations, by individual industry and by individual year. In addition, quantity measures of labour input (*i.e.* hours worked) have to be accompanied by price measures (*i.e.* relative average compensation) to construct weights for aggregation. Such rich data sets are normally both difficult and costly to collect and therefore not readily available in practice.¹⁰

Even when such data are not available, “implicit differentiation” can provide a useful, albeit incomplete, adjustment for labour quality. Implicit differentiation arises when labour input (*i.e.* total hours worked) is measured by detailed industry without, however, distinguishing between different types of labour within each industry. If the rate of change in hours worked by industry are aggregated to the economy-wide level using each industry’s share in total labour compensation as its aggregation weight, these weights will be relatively large for industries that pay above-average wages and relatively small for industries with below-average wages. Assuming that above-average wages reflect an above-average skill composition of the workforce, some of the quality change of labour input is implicitly taken into account. Statistics Canada’s industry-level productivity statistics provide an example of implicit differentiation, since the indices of labour input at the sectoral level are built up from hours-worked data for more detailed industries that are weighted by their shares in total sectoral labour compensation.

Capital input

In a production process, labour, capital and intermediate inputs are combined to produce one or several outputs. Conceptually, many facets of capital input measurement are directly analogous to labour input measurement (Table 3). Capital goods, whether purchased or rented by a firm, provide a flow of capital services that constitutes the actual input to the production process. Similarly, employees hired for a certain period can be seen as providing flows of labour services from their stocks of human capital. Differences between labour and capital arise because producers usually own capital goods. When the capital good “delivers” services to its owner, no market transaction is recorded. The measurement of these implicit transactions – whose quantities are the services drawn from the capital stock during a period and whose prices are the user costs or rental prices of capital – is one of the challenges of capital measurement.

Constructing measures of capital services.¹¹ Conceptually, capital services reflect a quantity or physical concept that should not to be confused with the value or price concept of capital. Because flows of the quantity of capital services are not usually directly observable, they have to be approximated. Most often, this is done by assuming that service flows are in proportion to the stock of assets,

after each vintage has been converted into standard “efficiency” units. The capital stock, so computed, is sometimes referred to as the “productive stock” of a given type of asset. Accordingly, the importance of capital stock measures to productivity analysis derives solely from the fact that they offer a practical tool to estimate flows of capital services – were the latter directly observable, there would be no need to measure capital stocks.

Several measures of capital frequently encountered in economic statistics do not provide estimates of capital services suited to use in measuring productivity. These include the net or wealth capital stock, which is the current market valuation of an industry's or a country's productive capital. One of the purposes of the wealth stock is to measure economic depreciation, *i.e.* the loss in value of an asset as it ages. Total depreciation across all vintages of an asset is exactly the amount by which the value of the net capital stock of an asset declines as an effect of ageing. However, the wealth stock is not the appropriate tool to capture the quantity side of capital services.

The “gross capital stock” is a closely related capital measure. It represents the cumulative flow of investments, corrected only for the retirement of capital goods, but based on the assumption that an asset's productive capacity remains fully intact until the end of its service life (sometimes called “one-hoss-shay”). For a single, homogenous asset, the gross capital stock can be considered a special case of the productive stock, where an asset loses nothing of its physical productive capacity until it is retired.

Table 3. Labour and capital inputs

	Labour input	Capital input
	Human capital	Physical capital
Services to production from input factors:		
Quantity	Labour services, measured as total hours worked	Capital services, measured as total machine hours (typically, assumed to be in fixed proportion to capital stock)
Prices	Compensation per hour	User cost of capital per unit of capital service
Differentiation	By industry and by type of labour input	By industry and by type of capital asset
Factor cost or factor income	Compensation per hour x total hours	User costs x productive capital services
Aggregation weights	Industry-specific and labour-quality-specific shares in total labour compensation	Industry-specific and asset-specific shares of total user costs of capital

In empirical applications, the growth rate of capital services typically exceeds that of the wealth stock. Using the wealth stock as a measure of capital input in productivity calculations would thus imply an overstatement of MFP growth compared with the MFP associated with capital services (see below). On the other hand, gross capital stock measures in productivity calculations potentially lead to an understatement of MFP growth, as gross stocks grow more rapidly than capital services.

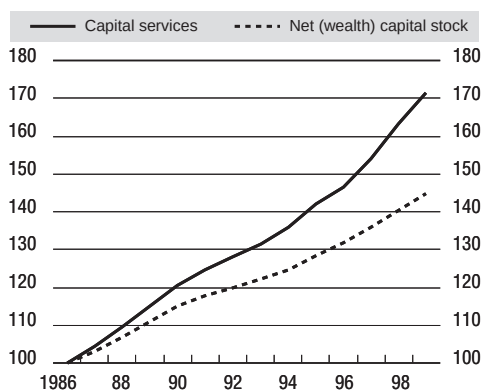
The price of capital services is measured by their rental price. If there were complete markets for capital services, rental prices could be directly observed. In the cases of, for example, office buildings or cars, rental prices do exist and are observable in the market. However, this is not the case for many other capital goods that are owned by producers and for which rental prices have to be imputed. The implicit rent that capital good owners “pay” themselves gives rise to the terminology “user costs of capital”.

Because many different types of capital goods are used in production, an aggregate measure of the capital stock or of capital services must be constructed. For net (wealth) stocks this is a straightforward matter of summing estimates for different types of assets. In so doing, market prices serve as aggregation weights. The situation is different in productivity analysis. Typically, each type of asset is associated with a specific flow of capital services and strict proportionality is assumed between capital services and capital stocks at the level of individual assets. This ratio is not the same, however, for different kinds of assets, so that the aggregate stock and the flows covering different kinds of assets must diverge. A single measure cannot serve both purposes except when there is only one single homogenous capital good (Hill, 1999a).

Under competitive markets and equilibrium conditions, user costs reflect the marginal productivity of the different assets. User cost weights thus provide a means to effectively incorporate differences in the productive contribution of heterogeneous investments as the composition of investment and capital changes. Jorgenson (1963) and Jorgenson and Griliches (1967) were the first to develop aggregate capital service measures that take the heterogeneity of assets into account. They defined the flow of quantities of capital services individually for each type of asset, and then applied asset-specific user costs as weights to aggregate across services from the different types of assets.

Figure 2 shows an example for the differences in capital measures that arise from the two concepts. Over the period under consideration, the capital services measure in Australia grew at a significantly faster pace than the wealth measure in that same country. To explain, note that wealth measures are based on an aggregation across different assets where each asset is weighted by its market price. The weights that are used to construct capital service measures are higher for short-

Figure 2. **Capital services and net capital stock measures**
Australia, 1986-99



Source: Australian Bureau of Statistics.

lived assets than for long-lived ones: one dollar invested in a short-lived asset must yield a higher return per year than a dollar invested in a long-lived asset to make that investment worthwhile. If shorter-lived assets grow more rapidly than longer-lived ones, the measure of capital services will grow faster than the wealth measure.

This feature can also be found in other countries, in particular the United States (Dean *et al.* 1996). It implies that the choice of the capital measure may have non-negligible impacts on measured productivity growth. For example, Australia's multi-factor productivity grew by an annual average rate of 2.0 per cent over the period 1995-99, when based on a capital services measure. The capital services indicator grew by 4.7 per cent per year over the same period, whereas the net (wealth) capital stock measure only showed a 3.1 per cent rise. Assuming a capital share of about 0.3, the resulting 1.6 percentage point difference implies approximately a 0.5 percentage point adjustment to the MFP measure (0.3×1.6 per cent = 0.48 per cent). Thus, based on net stock rather than a measure of capital services, Australia's MFP growth would have been evaluated at 2.5 per cent over the years 1995-99, and hence over-estimated compared with the correct capital services measure; too large a share of output growth would have been attributed to a change in MFP rather than to an increased contribution of physical capital to output.

Capital and capacity utilisation. There are many reasons why the rate of utilisation of capital, or more generally, the rate of utilisation of capacity of a firm varies over time: a change in demand conditions, seasonal variations, interruptions in the supply of intermediate products or a breakdown of machinery are all examples of factors that lead to variations in the flow of capital services drawn from a stock of assets. And yet, it is frequently assumed (for want of better information on utilisation rates) that the flow of services is a *constant* proportion of the capital stock. This is one of the reasons for the pro-cyclical behaviour of productivity series: variations in output are reflected in the data series, but the corresponding variations in the utilisation of capital (and labour) inputs are inadequately captured. If machine hours were measured, adjustments could be made. However, in practice, the required data rarely exist and consequently, swings in demand and output are picked up by the residual productivity measure. There have been several attempts to deal with this issue, but a generally accepted solution has yet to be found.

Index numbers

Productivity is usually measured as the ratio of a quantity index of output to a quantity index of inputs. Indices are required because the heterogeneity of goods and services does not permit simply adding up units of different types of commodities. However, aggregation results are in general sensitive to the choice of a specific index number formula. These formulae should therefore be chosen carefully on both conceptual and practical grounds.

A first choice that must be made for comparisons over several periods is whether to compare two periods directly (say, between period 0 and period 2), or indirectly (in which case the change between period 0 and 2 is derived from the change between period 0 and 1, combined with that from period 1 to 2). The economics literature, as well as the 1993 System of National Accounts, are quite unanimous that inter-temporal comparisons over longer periods should be obtained by chaining, *i.e.* by linking the year-to-year movements. The main reason for chaining is that it allows one to adopt weights that reflect economic behaviour: for example, a relative price fall of a good will typically lead to higher consumption of this good and changes in the expenditure share of this item. Chained indices reflect such changes in expenditure patterns because weights are regularly updated. When indices are based on weights that only reflect conditions in a base period that is several years away from the comparison period, there is a risk of weights being out of date and this may introduce a bias into the price or volume measure.

A second choice pertains to the specific index number formula. The most widely used index number formulae are the Laspeyres and Paasche indices (the

former uses base-period weights, the latter current period weights), the Fisher index (a geometric average of the Laspeyres and Paasche indices) and the Törnqvist index (a weighted geometric average of its components).

To help decide between different index number formulae, a series of intuitively appealing criteria have been developed in the index number literature, starting with the impressive work by Irving Fisher (1922). Examples of such criteria are the *identity test* (if the prices of period 1 are the same as those of period 0, then the price index should take a value of one) or the *commensurability test* (the price index should be independent of the units of measurement). Many other criteria exist (see Balk, 1995 for a survey), and the different index number formulae can be checked against them.

Another approach makes use of economic theory to define price or quantity indices theoretically. A well-known example is the Konüs (1924) cost of living index derived from the micro-economic theory of the consumer. It is represented as the ratio of the minimum expenditure in period 1 over minimum expenditure in period 0, while maintaining utility constant. Empirically, it is not normally possible to measure such a theoretically defined index directly as the specific form and parameters of utility or cost functions are unknown. However, Diewert (1976) showed that there are functional forms (such as the translog functional form) that provide approximations to arbitrary, twice differentiable homogenous functions. He further showed that these functional forms can be exactly represented by certain index number formulae that he called “superlative” index numbers. This provides a strong economic rationale for the use of superlative index numbers, such as the Fisher and the Törnqvist index. Empirically, it turns out the choice between different types of superlative indices matters little and can thus be left to the individual researcher.

ESTIMATING PRODUCTIVITY LEVELS

International comparisons of productivity growth can give useful insights in growth processes, but should ideally be complemented with international comparisons of income and productivity levels. An examination of levels gives insights into the possible scope for further gains, and also places a country's growth experience in the perspective of its current level of income and productivity. OECD has published estimates of productivity levels in various studies (*e.g.* Englander and Gurney, 1994; Pilat, 1997; Scarpetta, *et al.*, 2000). Most of these studies have not looked in detail at measurement issues, or examined how differences in productivity levels should be interpreted.

International comparisons of productivity levels require three components, namely comparable information on output (typically GDP), comparable informa-

tion on factor inputs (labour, capital) and conversion factors (or purchasing power parities, PPPs) to translate output and factor inputs in national currencies to a common currency. This section discusses the available estimates and some of the main measurement issues, in particular the use of PPPs for currency conversions, as well as the correspondence between output and labour input measures for level comparisons. The discussion focuses on productivity at the aggregate level. The estimation of sectoral productivity levels raises additional measurement issues that go beyond the scope of this paper.¹²

Output, labour and capital input

Comparability of output measures

The measurement and definition of economic output is treated systematically across countries in the 1993 System of National Accounts (SNA 93). The revisions to the SNA introduced at that time tend to increase the level of total GDP, although not uniformly over time or across countries. The SNA revisions raised the level of GDP in all of the OECD countries who have implemented these changes, with the increases ranging from 0.3 per cent for Belgium's 1996 GDP, to 7.4 per cent for Korea's 1996 GDP (Scarpetta, *et al.*, 2000). Although most OECD countries have now implemented the new SNA, Switzerland and Turkey are exceptions. Thus, GDP (and consequently productivity) levels in these countries are likely to be underestimated compared with countries that have implemented the new system, although the extent of this bias is unknown. Once the introduction of the new SNA is completed, international comparability is likely to have improved.

A second factor that may influence the comparability of GDP across countries is size of the non-observed economy. In principle, GDP estimates in the national accounts take account of this part of the economy. In practice, questions can be raised about the extent to which official estimates have full coverage of economic activities that are included in GDP according to the SNA, or to which extent mis-reporting is involved. Large differences in coverage could substantially affect comparisons of productivity levels. Little is known about the possible size of the non-observed economy in different OECD member countries, although work is currently underway to address this issue (OECD, 2000).

Comparability of labour-input measures

Equally important for international comparisons of productivity levels are comparable measures of labour input. Labour input is commonly measured along three dimensions: the number of persons engaged; the total number of hours worked of all persons engaged; and the total number of hours worked adjusted for the quality of individual workers. Employment statistics are quite well standard-

used across OECD countries as most countries provide labour force statistics along agreed guidelines. In principle, therefore, they pose few problems for international comparisons. The main difficulty is to ensure that the employment data are consistent in coverage with other data that are required to make comparisons of productivity, notably GDP and hours worked.

There are two issues that arise here. The first question is whether countries integrate estimates of employment in the national accounts. These estimates could be based on different sources, such as labour force surveys and enterprise statistics, and would, in principle, be more consistent with GDP than estimates relying on only one single source. In practice, not all OECD countries integrate employment statistics in the national accounts, implying that labour force statistics remain the most comparable source.

The second, closely related, problem is whether estimates of hours worked are consistent with the employment data. Estimates of hours worked are typically based on two alternative sources, labour force surveys and enterprise statistics. Labour force surveys are based on surveys of households, whereas enterprise statistics survey firms. Both sources seem to have some advantages and disadvantages for comparisons of productivity levels (OECD, 1999a; Van Ark and McGuckin, 1999):

- Labour force surveys may underestimate absences due to illness and holidays.
- Evidence on time use related to labour force surveys suggests that persons who work long hours may overestimate their working time.
- Labour force surveys potentially pick up extra hours worked by managers and professionals that are over and above the conventional hours of work in an establishment and that are clearly not picked up by establishment sources.
- Enterprise statistics are less likely to provide full coverage of all persons engaged in the economy, and may underestimate overtime.

In principle, this suggests that labour force surveys may somewhat overestimate total hours worked, whereas enterprise surveys may underestimate hours worked. Much depends on the quality and coverage of the surveys, however, and several OECD countries provide comprehensive estimates of hours worked based on a mix of sources. The OECD has recently produced estimates of hours worked that draw on such a mix of sources, where hours worked estimates from labour force surveys were adjusted downwards to compensate for known biases (Scarpetta, *et al.*, 2000). Cross-country comparability can thus be improved as compared to the use of original national sources for some countries, but there remains a margin of uncertainty.

The quality of labour input is much more difficult to compare across countries, in particular in terms of levels. Education systems and standards differ con-

siderably across countries, information on firm-level training is not collected or incomplete in most countries, and reliable data are scarce in all these areas. While some efforts have been made – primarily at the level of individual sectors – to account for qualitative differences in labour input across countries, the statistical basis is limited, and the results are therefore not necessarily robust.¹³ The OECD has therefore not yet estimated quality-adjusted levels of labour input for international comparisons of productivity levels. The decomposition of GDP per capita below therefore ignores quality aspects of labour input in the estimation of labour productivity levels.

Comparability of capital-input measures

Estimates of labour productivity levels are quite common. Much more difficult and more controversial is the comparison of levels of capital productivity and of multi-factor productivity.¹⁴ The main controversy concerns the comparison of capital input across countries (Van Ark, 1996). Official estimates of capital stock embody a wide variety of assumed asset lives and retirement patterns.¹⁵ Some of the variation across countries may be accounted for by compositional differences and differences in technological progress, which cause the capital stock to become obsolete more rapidly in some countries. However, this is difficult to verify and in most cases the statistical basis for the variation in asset lives and retirement patterns is weak, since statistical offices collect such information only infrequently.

A second problem concerns the conversion of capital stocks in national currency values into a common currency. This requires PPPs for capital stock. In principle, these can be derived from official PPPs for investment by converting investment series to a common currency and then calculating capital stock in a common currency. This requires reliable PPPs for investment and deflators for investment. PPPs for investment have been criticised in some recent evaluations of the PPP programme (OECD, 1997), however, and it is therefore not clear how reliable they are. Further progress is needed in making capital stocks more comparable across countries so as to improve the basis for comparisons of capital and multi-factor productivity.¹⁶

Purchasing power parities for international comparisons

The comparison of income and productivity across countries also requires purchasing power parity (PPP) data for GDP. Exchange rates are not suitable for the conversion of GDP to a common currency, since they do not reflect international price differences, and since they are heavily influenced by short-term fluctuations. Over the past two decades, OECD has regularly published estimates of PPPs, derived from its joint programme with Eurostat. Benchmark estimates of PPPs are currently available for 1980, 1985, 1990, 1993 and 1996, and work is underway for a new benchmark comparison for 1999.¹⁷ In using these PPP estimates for international compari-

sons of income and productivity, two issues must be addressed, namely the choice of aggregation method and the choice of benchmark year.

Aggregation method

The choice of aggregation method for constructing PPPs has been a source of debate over the past two decades. Initial work on international comparisons, such as the seminal study by Kravis, Heston and Summers (1982), provided a wide range of aggregation methods. The latest benchmark comparisons by OECD and Eurostat offer only two alternatives, namely those based on the Geary-Khamis (GK) method, and those based on the Elteto-Koves-Szulc (EKS) method.¹⁸ Aggregation takes place after price ratios for individual goods and services have been averaged to obtain unweighted parities for small groups of homogeneous commodities. It involves weighting and aggregating the unweighted commodity group parities to arrive at PPPs and real values for each category of expenditure up to the level of total GDP.

The two methods differ substantially. The EKS method treats countries as a set of independent units with each country being assigned equal weight. The EKS prices are obtained by minimising the differences between multilateral binary PPPs and bilateral binary PPPs. The EKS PPPs are thus close to the PPPs that would have been obtained if each pair of countries had been compared individually. The GK method treats countries as members of a group. Each country is weighted according to its share in GDP and the prices that are calculated are characteristic of the group overall. Both methods have advantages and disadvantages:

- For countries with price structures that are very different from the average, the GK approach leads to higher estimates of volumes (and GDP per capita) than if more characteristic prices would have been used. This effect is particularly important when comparing countries with great differences in income levels. The GK approach leads to results that are additively consistent, however, which implies that the real value of aggregates is the sum of the real value of its components. This is an advantage for national accounts and permits comparisons of price and volume structures across countries.
- The EKS method leads to results that are more characteristic of each country's own prices, and thus leads to estimates of GDP per capita that are relatively similar to those resulting from the use of characteristic prices. Its results are not additive, however.

For OECD countries, the differences between the two methods are relatively small, since national price structures are similar. Most comparisons of income and productivity utilise the EKS results, however, since these do not seem to overestimate income levels for low-income economies and are more closely aligned with

index number theory.¹⁹ The EKS method is also the method officially accepted by Eurostat for administrative purposes.

Benchmark year

The second issue to be addressed concerns the choice of benchmark year. For several OECD countries, the OECD/Eurostat estimates of PPPs are currently available for five benchmark years. This raises a problem of which benchmark to choose for international comparisons. In principle, it seems appropriate to use the most recent benchmark, *i.e.* 1996, since this is most likely to reflect current price differences in the OECD area. To indicate the sensitivity of comparisons of income and productivity to the choice of benchmark, Table 4 provides an overview of comparative estimates of GDP per capita for 1999, based on alternative benchmark results. It suggests that there is some variation in results between the different benchmark years, but that there are relatively small differences in results between recent benchmark years, *i.e.* 1990, 1993 and 1996. For most countries, the 1996 benchmark gives slightly higher estimates of GDP per capita relative to the United States than the 1993 benchmark. France, Belgium, Norway and Spain are exceptions and have slightly lower levels of GDP per capita with the 1996 benchmark.

Estimates of income and labour productivity levels

Clearly, data for international comparisons of income and productivity are not perfect and some choices between different sources have to be made. In this paper, GDP is derived from the OECD SNA database, which incorporates the latest comparative information on GDP from OECD member countries. Data on employment are from the OECD *Labour Force Statistics*, as this source is more standardised across countries than employment data from the national accounts and since the estimates of hours worked that are used in this paper are closely linked to labour force surveys. To convert GDP to a common currency, this paper uses the 1996 benchmark PPPs as the most recently available.

Table 5 presents the resulting income and productivity level estimates for 1999. It shows a considerable diversity in real per-capita GDP levels across the OECD countries. The United States is at the top of the OECD income distribution, followed by Switzerland and Norway that have levels of GDP per capita between 80 and 90 per cent of the US level. Luxembourg also has a very high level of GDP per capita, which is partly due to the large share of frontier workers in total employment (56 000 out of 226 000 workers in 1997). These contribute to GDP and employment, but are not included in the working-age and total population. The bulk of the OECD, including all the other major economies, has income levels that are between 65 and 80 per cent of the US level. Following this group are

Table 4. Estimates of GDP per capita for 1999 and the impact of alternative purchasing power parities

	Alternative PPPs for 1999, national currency/USD					Alternative estimates of GDP per capita, United States = 100				
	Official 1999 PPP based on the 1996 EKS index	Based on the 1993 EKS index	Based on the 1990 EKS index	Based on the 1985 Fisher index	Based on the 1980 Fisher index	Official estimate, based on the 1996 EKS index	Based on the 1993 EKS index	Based on the 1990 EKS index	Based on the 1985 Fisher index	Based on the 1980 Fisher index
Australia	1.30	1.30	1.30	1.41	..	75.6	75.5	75.8	69.4	..
Austria	13.6	13.7	14.2	17.0	14.8	72.8	72.2	69.6	58.2	66.7
Belgium	37.1	36.5	39.2	44.8	37.5	73.4	74.6	69.4	60.8	72.6
Canada	1.17	1.21	1.22	1.19	1.12	79.3	76.5	76.0	77.9	82.6
Czech Republic	13.4	..	..	..	..	39.5	..	..	..	..
Denmark	8.54	8.92	9.41	10.81	9.15	79.1	75.7	71.8	62.5	73.8
Finland	6.15	6.14	6.22	7.05	..	67.2	67.3	66.4	58.6	..
France ¹	6.63	6.34	6.32	7.23	6.22	65.2	68.2	68.5	59.8	69.6
Germany	1.98	2.05	2.12	2.52	2.18	70.4	68.1	65.7	55.3	64.0
Greece	239.2	250.3	277.6	294.2	278.5	44.7	42.8	38.5	36.4	38.4
Hungary	100.7	..	..	..	..	33.3	..	..	..	..
Iceland	85.5	89.8	..	..	..	77.8	74.1	..	..	..
Ireland	0.73	0.72	0.77	0.84	0.72	75.1	75.4	70.6	64.8	75.3
Italy	1 601.1	1 681.2	1 675.2	1 911.2	1 694.7	68.2	64.9	65.1	57.1	64.4
Japan	158.7	161.6	165.9	190.0	168.3	75.3	74.0	72.1	62.9	71.0
Korea	642.9	677.7	..	..	..	47.5	45.0	..	..	..
Luxembourg	40.6	40.7	..	..	..	122.2	121.9	..	..	..
Mexico	5.61	..	..	..	..	25.0	..	..	..	..
Netherlands	1.99	2.13	2.14	2.32	2.11	77.5	72.3	72.0	66.3	72.9
New Zealand	1.46	1.46	1.52	1.84	..	55.1	54.8	52.9	43.6	..
Norway	9.50	9.38	9.76	9.71	8.16	83.1	84.3	80.9	81.4	96.9
Poland	1.84	..	..	..	..	25.6	..	..	..	..
Portugal	128.9	..	..	..	..	48.6	..	..	..	..
Slovak Republic	13.8	..	..	..	..	32.4	..	..	..	..
Spain	130.5	127.5	131.1	146.1	130.7	53.8	55.1	53.6	48.1	53.8
Sweden	9.78	9.77	9.56	10.48	..	68.0	68.1	69.6	63.5	..
Switzerland ²	1.90	1.98	2.11	..	..	84.7	81.1	76.2	..	..
Turkey ²	185 567	183 857	183 799	177 331	..	18.7	18.9	18.9	19.6	..
United Kingdom	0.66	0.67	0.66	0.74	0.69	67.6	66.3	66.9	60.0	63.8
United States	1.00	1.00	1.00	1.00	1.00	100.0	100.0	100.0	100.0	100.0

1. Includes Overseas Department.

2. Countries still using the 1968 SNA, i.e. GDP may be underestimated compared with other OECD countries.

Source: GDP and employment from OECD SNA database; PPP estimates for 1980, 1985 and 1990 from Maddison (1995); 1993, 1996 and 1999 from OECD Statistics Department.

Table 5. Breakdown of GDP per capita in its components, 1999

	GDP per head of population (as % of US)	Effect of % working-age population (15-64 years) to total population	Effect of % labour force to working-age population	Effect of unemployment	Effect of working hours	Total effect of labour force participation	GDP per hour worked (as % of US)	GDP per person employed (as % of US)
	(1)	(2)	(3)	(4)	(5)	(6)(2) + (3) + (4) + (5)	(7)(1) - (6)	(8)(1) - (2) - (3) - (4)
Australia	76	1	-4	-2	-3	-8	84	81
Austria	73	2	-6	0	-17	-22	95	77
Belgium	73	0	-15	-7	-14	-36	110	96
Canada	79	3	-2	-3	-4	-6	86	82
Czech Republic	40	2	-3	-2	3	1	39	42
Denmark	79	1	2	-1	-16	-14	93	77
Finland	67	1	-3	-5	-8	-15	82	74
France ¹	65	-2	-10	-6	-14	-32	97	84
Germany	70	2	-6	-4	-16	-23	94	78
Greece	45	1	-11	-4	2	-12	56	59
Hungary	33	1	-10	-1	-2	-12	45	43
Iceland ²	78	-2	9	1	-2	6	72	70
Ireland	75	1	-11	-1	-9	-21	96	87
Italy	68	1	-19	-7	-13	-38	106	93
Japan	75	3	0	0	-1	1	74	73
Korea	47	4	-9	-1	14	7	40	54
Luxembourg	122	1	13	2	-14	2	120	106
Mexico	25	-2	-5	1	1	-6	31	32
Netherlands	78	2	-4	1	-30	-32	109	79
New Zealand ²	55	0	-2	-2	-3	-7	62	59
Norway	83	-1	3	1	-27	-25	108	81
Poland	26	1	-5	-3	..	-7	..	32
Portugal	49	1	-2	0	-3	-5	53	50
Slovak Republic	32	1	-4	-5	..	-8	..	40
Spain	54	2	-14	-9	-2	-23	76	75

Table 5. Breakdown of GDP per capita in its components, 1999 (cont.)

	GDP per head of population (as % of US)	Effect of % working-age population (15-64 years) to total population	Effect of % labour force to working-age population	Effect of unemployment	Effect of working hours	Total effect of labour force participation	GDP per hour worked (as % of US)	GDP per person employed (as % of US)
	(1)	(2)	(3)	(4)	(5)	(6)(2) + (3) + (4) + (5)	(7)(1) - (6)	(8)(1) - (2) - (3) - (4)
Sweden	68	-2	-1	-2	-10	-15	84	73
Switzerland ²	85	2	5	1	-14	-6	91	77
Turkey ²	19	0	-8	-1	..	-9	..	28
United Kingdom	68	-1	-3	-1	-14	-19	87	72
United States	100	0	0	0	0	0	100	100
Euro-area ³	66	1	-11	-5	-12	-27	92	80
European Union	66	1	-9	-4	-13	-25	91	79
G7 countries	83	1	-4	-2	-5	-10	92	87
OECD ⁴	72	1	-6	-2	-3	-10	82	78

1. Includes Overseas Departments.

2. GDP estimates for Iceland, New Zealand, Switzerland and Turkey are based on the SNA 68.

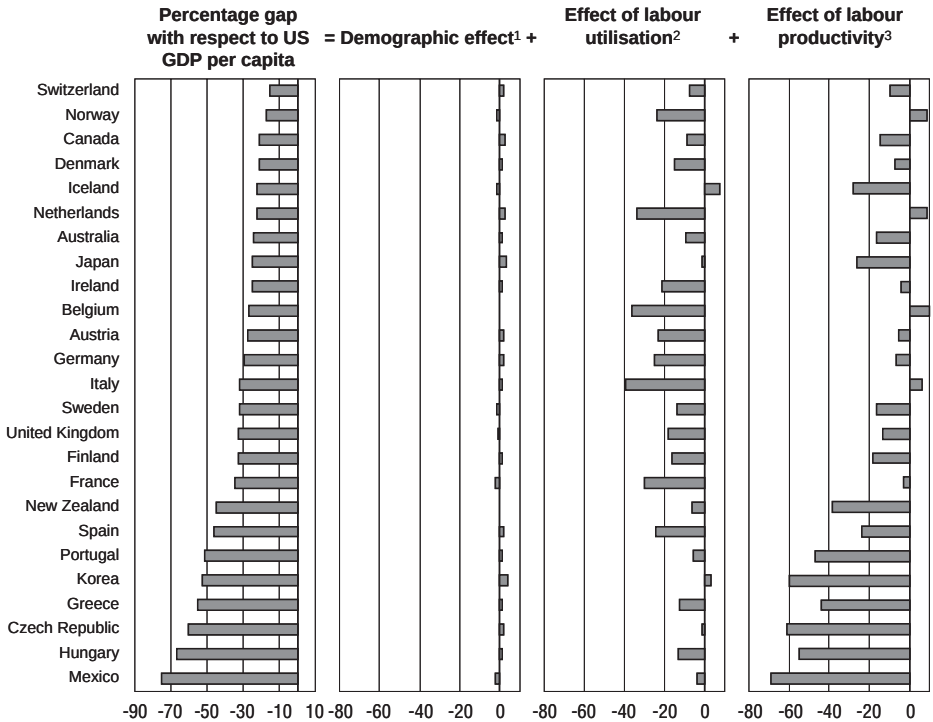
3. Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain.

4. Excluding Poland, Slovak Republic and Turkey.

Source: OECD, GDP and population from OECD National Accounts, working-age population, labour force and employment from OECD Labour Force Statistics; hours worked from OECD calculations, see Scarpetta, *et al.* (2000). GDP converted to common currency by 1999 OECD PPP.

Figure 3. Differentials in GDP per capita and their determinants, 1999

Percentage point differences in PPP-based GDP per capita with respect to the United States



1. Based on the ratio of the working-age population (15-64 years) to the total population.

2. Based on employment rates and average hours worked.

3. GDP per hour worked.

Source: Table 5.

a number of lower-income economies, including Greece, Korea, Portugal, Spain and New Zealand, some of which have experienced very high growth over the recent period. Mexico, Turkey and the former centrally-planned economies (Czech Republic, Hungary and Poland) are at the bottom of the OECD income distribution.

Figure 3 shows clearly that levels of GDP per capita are not the same as differences in labour productivity, *i.e.* GDP per hour worked. The differences is particularly marked for certain European countries, such as France, Italy, Belgium and the Netherlands, that have levels of GDP per hour worked that are higher or comparable

to the United States but much lower levels of GDP per capita. Low labour utilisation, *i.e.* low employment rates and short working hours, explains the bulk of this gap. For most other OECD countries, in particular those at the bottom end of the OECD income distribution, labour productivity is closely linked to levels of GDP per capita.

High labour productivity is often associated with strong economic performance. However, some countries, with high levels of labour productivity have very low levels of labour utilisation, suggesting that high labour productivity may partly be due to increasing the capital-labour ratio and to pushing low-productivity workers into unemployment or out of the labour force. Estimates of GDP per hour worked should therefore ideally be looked at in combination with estimates of GDP per capita.

THE INTERPRETATION OF PRODUCTIVITY MEASURES

Productivity is an important yardstick of economic performance. Different productivity measures are, however, often used without sufficient clarity about the specific measure that is being used and its correct interpretation. This section briefly touches on six of the most common uses made of productivity measures and possible pitfalls: the relationship between productivity growth and technological change; the link between productivity and costs; productivity over the business cycle; the difference between productivity and efficiency; the links between industry and firm-level productivity growth; and innovation and productivity growth.

The link with technological change

Multi-factor productivity growth is often interpreted as an indicator of technological progress. This is not entirely correct for three reasons: *i)* technological change does not necessarily translate into MFP growth; *ii)* MFP growth is not necessarily caused by technological change; and *iii)* MFP may understate the eventual importance of productivity change in stimulating the growth of output. These three factors are discussed below.

Some technological change does not translate into MFP growth because embodied technological change, such as advances in the quality of new vintages of capital or improved human capital, is reflected instead in the measured contributions of capital and labour to output growth. Disembodied technical change, on the other hand, will be reflected in MFP growth as it relates to advances in scientific knowledge, and to the diffusion of knowledge on how things are done, including better management and organisational change. MFP should also include the

spill-over effects from capital and labour, *e.g.* network effects arising from investment in information technology products.

Conceptually, and following Jorgenson (1995*a*), the MFP term reflects all those effects on output growth that are *not* investment, where investment is understood as the commitment of current resources in the expectation of future returns, implying that these returns can be internalised by the investor. The distinction is important because the diffusion of embodied technical change is dependent on market transactions: investment in the improved capital or intermediate good will be undertaken until its marginal contribution to revenue generation just equals its user cost, itself dependent on the market price of the capital good. The diffusion of disembodied technical change is not necessarily associated with market transactions. Information may circulate freely and its use by one person does not normally restrict its use by another one.

When the measure for the quantity of capital services used in the MFP calculation is based on a price index that reflects quality changes in capital goods, embodied technological change is captured by the growth contribution of capital and not by the residual MFP. Conversely, when the capital goods price index is not adjusted for quality change, both embodied and disembodied technological change will be picked up by the MFP residual. As Hulten (1992) showed, the embodiment part of technology can be measured by comparing capital input based on quality-adjusted price indices and capital input based on unadjusted price indices. This idea was also pursued by Greenwood *et al.* (1997) and Bassanini *et al.* (2000).

Data and resource constraints often do not permit a careful differentiation and full coverage of all labour and capital inputs. As a consequence, some of the embodiment effects of technological change and some or all of the changes in the skill composition of labour input are picked up by the MFP residual. Thus, the correct interpretation of the productivity term with respect to technological change requires knowledge about the methodology used to compute time series of capital and labour input.

Just as some technological change does not correspond to MFP growth, some MFP growth is not caused by technological change. Even where the residual reflects part or all of technological change, several other factors will also bear on measured MFP. Such factors include adjustment costs, economies of scale, cyclical effects, inefficiencies and measurement errors. This is confirmed by econometric studies that link MFP growth to technology variables, in particular research and development and patents or those that explicitly control for adjustment costs or allow for non-constant returns to scale. Research and development expenditure, for example, tends to show a statistically significant relation to productivity growth, but only explains a relatively small part of the overall annual movements

in MFP.²⁰ This indicates the presence of other factors. Measures of MFP are thus better interpreted as measures of improvements in overall efficiency than as pure expressions of technical change.

Finally, MFP may understate the eventual importance of productivity change in stimulating the growth of output. This reflects the fact that in growth accounting models, capital is considered an exogenous input to the production process. In a dynamic context, this is no longer the case and a feedback mechanism exists between productivity change and capital. Suppose that technical change allows more output to be produced per person. The static MFP residual measures just this effect of technical change. Typically, however, additional output per person will lead to additional savings and investment, and to a rise in the capital-labour ratio. A traditional growth accounting framework would identify this induced effect as a growth contribution of capital, although it can be traced back to an initial shift in technology. Thus, the MFP residual correctly measures the shift in production possibilities but does not capture the induced effects of technology on growth (Rymes, 1971; Hulten, 2000).

Productivity growth and changes in costs

Productivity measures are typically derived on the basis of production functions and quantity measures of inputs and outputs. Under certain conditions (*e.g.* cost minimisation), an equivalent and intuitively appealing “dual” approach exists that expresses advances in productivity as downward shifts of a cost function. A cost function shows the minimum input cost of producing a certain level of output, given a set of input prices. Thus, the MFP productivity residual can be measured either in the residual growth of output not explained by the growth of inputs, or as the residual decline in average costs not explained by changes in input prices. A slightly different formulation is that productivity growth equals the diminution of total costs that is explained neither by a fall in output nor by substitution of inputs that have become relatively more expensive for those whose relative price has fallen.

The formulation of MFP in terms of average costs demonstrates once more that productivity is not necessarily technological change and that technological advances do not necessarily translate into a change of the MFP residual. It is intuitively plausible that total and average costs can be reduced by factors other than technological change, such as organisational innovations or learning by doing (Harberger, 1998). The cost approach also shows how embodied technological change can reduce the costs of inputs and trigger substitution processes without, however, changing the rate of multi-factor productivity growth.

The role of the business cycle in productivity growth

Most productivity measures are pro-cyclical, accelerating during periods of economic expansion and decelerating during periods of recession. In part, this is due to inadequate measurement. While cyclical variations in volume output tend to be measured quite accurately in economic statistics, variations in the rate of utilisation of inputs are not picked up as fully. In particular, changes in the rate of utilisation of capital equipment (*i.e.* changes in machine hours) are rarely captured in these measures. Labour input, if measured by hours actually worked, better reflects the changing rate of utilisation of manpower, but remains an imperfect measure. Consequently, increases in the rate of capacity utilisation in periods of expansion will cause output measures to show more rapid growth, but input measures may remain stable or grow less rapidly. The result is a rise in measured productivity growth. The converse holds for periods of recession.

Even if capacity utilisation were accurately measured, difficulties would remain in reconciling the standard productivity model with the realities of the business cycle. Much of the economic and index number theory used to guide the construction of these measures relies on long-term, equilibrium relationships. Since little or no account is taken of unforeseen events or disequilibria, the economic model of productivity measurement is more appropriately applied to periods of continued and moderate expansion, than to rapidly changing phases of the business cycle. This limitation means that year-to-year changes in productivity growth should not be interpreted *prima facie* as shifts in disembodied technology. For this purpose it is preferable to examine productivity growth patterns over longer periods of time or adjust productivity estimates for cyclical fluctuations.²¹ Moreover, it is wise not to draw strong conclusions about shifts in productivity from the evidence of just a few years.

The difference between productivity and efficiency

Productivity and efficiency are related, but not identical concepts (Sharpe, 1995). A firm or industry is considered to be inefficient if it could produce more output with existing inputs, *i.e.* the firm is not on the production possibility curve, but within it. Productivity relates the quantity of output produced to one or more inputs used in its production, irrespective of the efficiency of their use. In analysing productivity growth across countries, the difference between these two concepts allows a distinction between three different processes. First, productivity growth can result from innovative activity that results in an outward shift of the global production possibility frontier. Second, firms can improve productivity by adopting production processes and products developed elsewhere (imitation). Diffusion differs conceptually from efficiency gains, as the latter relates to improvements made in using a given technology – even when this technology is

outdated by international standards. Third, productivity growth may also be due to reduced (technical) inefficiency. An inefficient firm or industry uses more resources and factor inputs than required by a particular technology, thus tying resources to low-productivity activities and reducing the overall allocative efficiency of an economy. Understanding the reasons behind productivity growth is therefore necessary before attributing such changes to specific sources.

How productivity at the industry level is related to that at the firm level

The main focus of this paper has been productivity measurement at the industry and aggregate level of the economy. Industries and branches are themselves made up of individual firms and establishments and new micro-level databases have greatly enhanced the possibilities for empirical research to better understand individual units' productivity performance carries over to what is observed at the industry level. Several important conclusions have arisen from a significant body of studies (Haltiwanger, 2000; Bartelsman and Doms, 2000; OECD, 2001*b*). First, there are large differences in productivity performance between individual units. Second, there is a continuous and large-scale re-allocation of outputs and inputs between producers, including within industries. Third, this re-allocation contributes significantly to aggregate productivity growth. For example, Haltiwanger reports that for the United States manufacturing sector, roughly half of multi-factor productivity growth over the course of a recent decade can be accounted for by the re-allocation of outputs and inputs away from less productive to more productive businesses.

These findings do not invalidate the theory of productivity measurement in this paper that essentially treats an industry, a sector or even the entire economy as if it were a single firm. Rather, it adds to the understanding and interpretation of measured productivity growth. For example, it points to one mechanism by which an industry *as a whole* implements technical change: if new technology is mainly adopted by new establishments, productivity growth occurs with entry and exit, and this requires re-allocation. Technical advances at the industry level are then associated with the diffusion of new technology among establishments, rather than with a simultaneous shift in the production frontier of an existing set of businesses. This provides an additional interpretation of changes in the productivity residual at the industry level. Micro-level data studies with their focus on firm dynamics, entry and exit and re-allocation of resources also form a natural link to the question how innovation and "creative destruction" (see below) translate into industry-level productivity growth. Nonetheless, micro-level approaches cannot replace the more aggregate type of productivity measurement. This is due to poor data quality at the firm or establishment level (*e.g.* for capital input) and the timeliness and exhaustiveness of the available data sets. But micro-level studies

enhance our understanding of the underlying drivers and dynamics of productivity growth.

Innovation and productivity

Most approaches to measuring productivity are firmly rooted in a neo-classical equilibrium concept. Equilibrium conditions are very important because they help to guide measurement of parameters that would otherwise be difficult to identify. Although its usefulness is generally recognised, it has been argued that an equilibrium approach sits uneasily with the notion of innovation and productivity growth. Evolutionary economists (*e.g.* Dosi, 1988; Nelson and Winter, 1982; Nelson, 1981) argue that innovation and technical change occur as a consequence of information asymmetries, and market imperfections. In a quite fundamental sense, innovations and information asymmetries are one and the same phenomenon. Indeed, such asymmetries can scarcely be termed market imperfections when they are necessary conditions for any technical change to occur in a market economy (Metcalf, 1996). The point made by evolutionary economists is that equilibrium concepts may be the wrong tools to approach the measurement of productivity change, because if there truly was equilibrium, there would be no incentive to search, research and to innovate, and there would be no productivity growth.

Such criticism has to be taken seriously in the interpretation and use of productivity measures. An important lesson from this debate is that accounting is not explaining the underlying causes of growth. Griliches (1997) makes a related point:

"We can take productivity growth calculation and allocate it in great detail to the various missed components, reducing thereby the role of the "unallocated" residual. But this, while very instructive and valuable, only shifts the problem to a new set of questions: why was there all this investment in human capital? Will it continue? Where did the improvements in capital equipment come from? [...] Real explanations will come from understanding the sources of scientific and technological advances and from identifying the incentives and circumstances that brought them about and that facilitated their implementation and diffusion. Explanation must come from comprehending the historical detail."

This does not invalidate the usefulness of the standard equilibrium approach to productivity measurement, discussed in this paper, but alerts us to some of its limits. What emerges is the complementarity of approaches: growth accounting and productivity measurement allow the systematic and consistent quantification of the proximate sources of growth. It has explanatory power in that it captures the workings of supply of, demand for and substitution between categories of measurable inputs. At the same time, growth accounting has to be complemented by

institutional, historical and case studies if one wants to explore the underlying causes of growth, innovation and productivity change.

IN CONCLUSION

The measurement of productivity continues to generate substantial interest in many countries, since productivity is a key yardstick of economic performance. This paper has discussed the main measurement difficulties that need to be confronted when calculating productivity indicators and has provided guidance to researchers and statisticians in addressing these difficulties. Substantial progress has been made in recent years to improve the comparability of productivity statistics. In many countries, however, basic source data are still the key limitation to the development of comparable indicators of productivity. In addition, statisticians, researchers and policy makers need to be more aware of the way in which productivity statistics are used and their appropriate interpretation.

NOTES

1. Multi-factor productivity is sometimes referred to as total factor productivity. Multi-factor productivity implies that several factors are included as inputs, though not necessarily all. Total factor productivity suggests that all possible factors are included. In practice, this is seldom the case.
2. This definition is consistent with the System of National Accounts 1993.
3. Recently, Nordhaus (2001) discussed the link between industry-level and aggregate productivity measures based on value added. He presented alternative methods for measuring productivity growth that are derived from the link between productivity and welfare. Nordhaus showed that, under a number of restrictive assumptions, a welfare-related productivity measure can be constructed as a weighted average of productivity growth in individual industries. One specific conclusion is that macroeconomic productivity gains that arise from shifts between industries with different productivity levels should be excluded from welfare-oriented productivity measures.
4. This is a general form of double-deflation, expressed in a continuous-time Divisia index.
5. To interpret value-added-based MFP growth as a measure of disembodied technical change, one has to assume that technical change operates only on primary inputs, and not on intermediate ones. There is little empirical evidence to support this assumption.
6. For a discussion regarding the United States, see Eldridge (1999).
7. See the *OECD Handbook on the Quality Adjustment of Price Indices for ICT Products* (forthcoming) and the *Eurostat Handbook on Price and Volume Measures in National Accounts* (forthcoming).
8. For an overview, see Bresnahan and Gordon (1996).
9. For example, Lowe (1996) provides an overview of how quality change is handled in the Canadian National Accounts.
10. For empirical results see United States Bureau of Labor Statistics (1993), Fosgerau *et al.* (2001) and Scarpetta *et al.* (2000).
11. For specific information about the practical and conceptual issues regarding the construction of capital stock measures, see the *OECD Manual on Capital Measurement* (2001).
12. Van Ark (1996) and Pilat (1997) discuss these issues in greater detail. See also O'Mahony (1999).
13. Estimates of quality-adjusted labour input are available for some countries, see for example O'Mahony (1999).
14. Levels of capital productivity and multi-factor productivity for five major OECD economies (United States, Japan, Germany, France, United Kingdom) are estimated in O'Mahony (1999).

15. An additional problem concerns the deflators that are used to measure investment. See Colechia and Schreyer (2001).
16. Ongoing work on the OECD *Capital Stock Manual* and the OECD *Productivity Manual* may help in enhancing the comparability of capital stock estimates across countries.
17. The Internet site of the OECD Statistics Department provides an overview of some of the key issues related to the construction of purchasing power parities, see www.oecd.org/std/ppps.htm. An evaluation of the PPP programme was prepared by the former chief statistician of the Australian Bureau of Statistics, Ian Castles, in 1997 (OECD, 1997), and has led to a range of improvements in the construction of PPPs. The recent benchmark study for 1996 also contains an extensive discussion of many of the issues related to the OECD/Eurostat work on PPPs (OECD, 1999b).
18. See Elteto and Koves (1964); Szulc (1964); Geary (1958); and Khamis (1970). More elaborate descriptions of these methods and the differences between them are available in OECD (1999b). See also Van Ark (1996), Pilat (1997) and OECD (1998) for a discussion of the use of PPPs for international comparisons of productivity.
19. The EKS method is closely related to superlative index numbers, such as the Theil-Tornqvist index.
20. See the paper by Guellec and Van Pottelsberghe in this issue of OECD *Economic Studies*.
21. More detail on the trend-adjustment of productivity growth is available in Scarpetta *et al.* (2000).

BIBLIOGRAPHY

- ARK, B. VAN (1996),
"Issues in Measurement and International Comparison of Productivity – An Overview",
in OECD (1996), *Industry Productivity – International Comparison and Measurement Issues*, Paris,
pp. 19-47.
- ARK, B. VAN and R.H. MCGUCKIN (1999),
"International comparisons of labor productivity and per capita income", *Monthly Labor
Review*, July 1999, pp. 33-41.
- BALK, B. (1995),
"Axiomatic price index theory: A survey", *International Statistical Review* 63, pp. 69-93.
- BARRO, R.J. and X. SALA-i-MARTIN (1995),
Economic Growth, MIT Press.
- BARTELSMAN, E. and M. DOMS (2000),
"Understanding productivity: lessons from longitudinal microdata", *Journal of Economic
Literature* 38(3), September, pp. 569-94.
- BASSANINI, A., S. SCARPETTA and I. VISCO (2000),
"Knowledge, technology and economic growth: recent evidence from OECD countries",
OECD Economics Department Working Papers No. 259.
- BRESNAHAN, T.F. and R.J. GORDON (eds.) (1996),
The Economics of New Goods, Studies in Income and Wealth, University of Chicago Press.
- BUREAU OF LABOR STATISTICS (1993),
Labor Composition and U.S. Productivity Growth, 1948-90, US Government Printing Office.
- COLECCHIA, A. and P. SCHREYER (2001),
"The impact of information and communications technologies on output growth", STI
Working Paper 2001/07, OECD, Paris.
- DEAN, E.R., M.J. HARPER and M.S. SHERWOOD (1996),
"Productivity measurement with changing-weight indices of outputs and inputs", in
OECD (1996a), *Industry Productivity – International Comparison and Measurement Issues*, pp. 183-215.
- DENISON, E.F. (1985),
Trends in American Economic Growth 1929-82, Brookings Institution, Washington, D.C.
- DIEWERT, W.E. (1976),
"Exact and superlative index numbers", *Journal of Econometrics*, pp. 115-45.
- DOSI, G. (1988),
"Sources, procedures and micro-economic effects of innovation", *Journal of Economic
Literature* 36.

- ENGLANDER, S. and A. GURNEY (1994),
 "Medium-term determinants of OECD productivity", *OECD Economic Studies*, No. 22, Spring, pp. 49-109.
- ELDRIDGE, L.P. (1999),
 "How price indexes affect productivity measures", *Monthly Labor Review*, February, pp. 35-46.
- ELTETO, O. and P. KOVES (1964),
 "On an index computation problem in international comparisons" (in Hungarian), *Statistikai Szemle*, Vol. 42, pp. 507-518.
- FISHER, I. (1922),
The Making of Index Numbers, Boston.
- FOSGERAU, M., S.E. HOUGAARD JENSEN and A. SØRENSEN (2001),
 "Measuring educational heterogeneity and labour quality: A note", *CEBR Discussion paper* 2001-6, www.cebr.dk.
- GEARY (1958),
 "A note on comparisons of exchange rates and purchasing power between countries", *Journal of the Royal Statistical Society*, Vol. 121, 97-99, pt. 1.
- GREENWOOD, J., Z. HERCOWITZ, and P. KRUSELL (1997),
 "Long-run implications of investment-specific technological change", *American Economic Review* Vol. 87(3), pp. 342-62.
- GRILICHES, Z. (1997),
 "R&D and the productivity slowdown: is recovery around the corner?", unpublished mimeo.
- HALTIWANGER, J. (2000),
 "Aggregate growth: what have we learned from microeconomic evidence?", *OECD Economics Department Working Papers*, No. 276, OECD, Paris.
- HARBERGER, A.C. (1998),
 "A vision of the growth process", *American Economic Review*, March, pp. 1-32.
- HILL, P. (1999a),
 "The productive capital stock and the quantity index for flows of capital services"; paper presented at the third meeting of the Canberra Group on Capital Stock Statistics, Washington, D.C.
- HULTEN, C. (1992),
 "Growth accounting when technical change is embodied in capital", *American Economic Review*, Vol. 82 (4), pp. 964-80.
- HULTEN, C. (2001),
 "Total factor productivity: A short biography", in: Hulten, Dean and Harper (eds.). *New Developments in Productivity Analysis*, University of Chicago Press for the National Bureau of Economic Research
- JORGENSEN, D. (1963),
 "Capital theory and investment behaviour", *American Economic Review*, Vol. 53, pp. 247-259.
- JORGENSEN, D. (1995a),
Productivity Volume 1: Postwar US Economic Growth, MIT Press.
- JORGENSEN, D. (1995b),
Productivity Volume 2: International Comparisons of Economic Growth, MIT Press.

- JORGENSEN, D., F. GOLLOP and B. FRAUMENI (1987),
Productivity and US Economic Growth, Harvard University Press, MA.
- JORGENSEN, D. and Z. GRILICHES (1967),
 "The explanation of productivity change", *Review of Economic Studies* 34.
- KHAMIS, S.H. (1970),
 "Properties and conditions for the existence of a new type of index number", *Sankhya*, Series B, Vol. 32, Parts 1 and 2.
- KONÜS, A.A. (1924),
 "The problem of the true index of the cost of living", translated in *Econometrica* (1939) 7, pp. 10-29.
- KRAVIS, I.B., A. HESTON and R. SUMMERS (1982),
World Product and Income, Johns Hopkins University Press, Baltimore.
- KURODA, M., K. MOTOHASHI and K. SHIMPO (1996),
 "Issues on the international comparison of productivity: theory and measurement", in OECD (1996a),
Industry Productivity – International Comparison and Measurement Issues, Paris, pp. 49-95.
- LOWE, R. (1996),
 "Handling quality changes in the Canadian National Accounts price deflators", in OECD (1996a), *Industry Productivity – International Comparison and Measurement Issues*, pp. 143-157.
- MADDISON, A. (1991),
Dynamic Forces in Capitalist Development, Oxford University Press, Oxford.
- MADDISON, A. (1995),
Monitoring the World Economy, 1820-1992, OECD, Paris.
- METCALFE, S. (1996),
 "The economic foundations of technology policy: equilibrium and evolutionary perspectives", in P. Stoneman (ed.) (1996), *Handbook of the Economics of Innovation and Technological Change*, pp. 409-512.
- MORRISON, C.J. (1986),
 "Productivity measurement with non-static expectations and varying capacity utilization", *Journal of Econometrics* 33.
- MOULTON, B.R., R.P. PARKER and E.P. SESKIN (1999),
 "A preview of the 1999 comprehensive revision of the national income and product accounts – definitional and classificational changes", *Survey of Current Business*, August, pp. 7-20, Bureau of Economic Analysis.
- NADIRI, I. and I. PRUCHA (2001),
 "Dynamic factor demand models and productivity analysis", Hulten, Dean and Harper (eds.), *New Developments in Productivity Analysis*, University of Chicago Press for the National Bureau of Economic Research.
- NELSON, R. (1981),
 "Research on productivity growth and productivity differences: dead ends and new departures", *Journal of Economic Literature*.
- NELSON, R. and S. WINTER (1982),
An Evolutionary Theory of Economic Change, Cambridge, Mass., Harvard University.

- NORDHAUS, W.D. (2001),
 “Alternative methods for measuring productivity growth”, NBER Working Paper 8095,
 National Bureau of Economic Research, Cambridge, MA.
- OECD (1996a),
Industry Productivity: International Comparison and Measurement Issues, OECD Proceedings,
 Paris.
- OECD (1996b),
Services: Measuring Real Value Added, Paris.
- OECD (1996c),
OECD Input-Output Tables, Paris.
- OECD (1997),
 “Review of the OECD-Eurostat PPP Program”, Paris, available at www.oecd.org/std/ppp/pps.htm.
- OECD (1998a),
Technology, Productivity and Job Creation, Paris.
- OECD (1998b),
 “The dynamics of industrial performance – what drives productivity growth”, OECD Science, Technology and Industry Outlook 1998, Paris.
- OECD (1999a),
Employment Outlook, Paris.
- OECD (1999b),
 “Purchasing power parities and real expenditures – EKS Results, 1996”, Paris.
- OECD (2000),
 “Underground, informal and illegal activities (non-observed economy): creation and content of a measurement handbook”, STD/NA(2000)21, OECD, Paris.
- OECD (2001a),
OECD Productivity Manual: A Guide to the Measurement of Industry-Level and Aggregate Productivity Growth, OECD, Paris, www.oecd.org/subject/growth/prod-manual.pdf.
- OECD (2001b),
 “Productivity and firm dynamics”, OECD Economic Outlook, No. 69, Paris, June.
- O'MAHONY (1999),
Britain's Productivity Performance 1950-1996 – An International Perspective, National Institute of Economic and Social Research, London.
- PILAT, D. (1997),
 “Competition, productivity and efficiency”, OECD Economic Studies, No. 27, 1996/II, pp. 107-145, OECD, Paris.
- RYMES, T.K. (1971),
On Concepts of Capital and Technological Change, Cambridge.
- SCARPETTA, S., A. BASSANINI, D. PILAT and P. SCHREYER (2000),
 “Economic growth in the OECD area: recent trends at the aggregate and sectoral level”, OECD Economics Department Working Papers, No. 248, OECD, Paris.
- SCHREYER, P. (2001),
 “Computer price indices and international growth comparisons”, Growth Project Background Papers, OECD, Paris, www.oecd.org/subject/growth/products/index.htm.

- SHARPE, A. (1995),
"International perspectives on productivity and efficiency", *Review of Income and Wealth*,
Vol. 41, No. 2, June, pp. 221-237.
- SZULC, B. (1964),
"Index numbers of multilateral regional comparisons" (in Polish), *Przegląd Statystyczny*,
Vol. 3, pp. 239-254.

ESTIMATING THE STRUCTURAL RATE OF UNEMPLOYMENT FOR THE OECD COUNTRIES

**Dave Turner, Laurence Boone, Claude Giorno, Mara Meacci,
Dave Rae and Pete Richardson**

TABLE OF CONTENTS

Introduction	172
Conceptual framework and recent studies	173
The NAIRU and the Phillips Curve	173
Estimation methods in recent studies	176
The OECD approach to estimating the NAIRU	181
The estimation framework: the choice of inflation and supply shock variables	181
Specifying the Kalman filter	183
Determining the smoothness of the NAIRU	183
End-point adjustments	186
The estimation procedure	186
Results	187
The estimation results	187
Measures of uncertainty and revisions to the preliminary estimates	193
Recent trends in the NAIRU estimates	198
The relevance of NAIRU estimates for monetary policy and inflation	199
<i>Appendix. The Theoretical Framework</i>	202
Bibliography	211

The work presented in this paper was originally reported in P. Richardson, L. Boone, C. Giorno, M. Meacci, D. Rae and D. Turner (2000), subsequently updated in "Revised OECD Measures of Structural Unemployment", Chapter V, *OECD Economic Outlook* No. 68, December 2000. The authors are grateful to Jean-Philippe Cotis, Jørgen Elmeskov, Michael P. Feiner, Stefano Scarpetta and Ignazio Visco for comments on previous versions. The views expressed are those of the authors and do not necessarily reflect those of the OECD or its Member countries. Special thanks go to Laurence Le Foulher and Isabelle Wanner-Paoletti for their excellent technical support; and to Rosemary Chahed and Jan-Cathryn Davies for document preparation.

INTRODUCTION

An important challenge facing policy makers is to identify the rate of capacity utilisation that is sustainable, in the sense that it is associated with reasonably stable inflation, over the medium to longer term. There are different ways of measuring capacity utilisation. Looking at perhaps the most common measure, unemployment, this notion of sustainable resource utilisation is made operational in the concept of the NAIRU – the non-accelerating inflation rate of unemployment, *i.e.* the unemployment rate consistent with stable inflation.¹

Views are mixed as to the usefulness of the NAIRU concept. Nevertheless, economists analyse future inflation trends, the sustainability of fiscal positions, and the need to undertake structural reforms to permanently reduce unemployment and for these purposes they need a benchmark to identify and distinguish sustainable and unsustainable trends in output and unemployment. The NAIRU concept provides such a benchmark. Estimates of the NAIRU make clear what assumptions lie behind policy analysis and recommendations and therefore increase the transparency of policy advice.

The measurement of the NAIRU is also controversial. By its nature, it is non-observable and depends on a wide range of institutional and economic factors. It follows that even if one accepts the concept, it can only be estimated with uncertainty. Moreover, it may well vary over time – European experience suggests that, in general, inflation would accelerate if unemployment reached the low unemployment rates associated with stable inflation in the 1960s. And at times, such as when there are large fluctuations in oil or raw material prices, it is clear that unemployment would have to rise or fall very steeply to stabilise inflation.

This paper describes the recent work by the OECD to review its procedures for deriving estimates of the unemployment rates consistent with stable inflation. The procedures have been updated and improved in several respects. Most notably, the new procedures allow the distinction between and estimation of a slow-moving NAIRU and a more volatile short-term NAIRU, which is affected by temporary factors, such as oil price fluctuations, impacting inflation in the short term.² They also provide a gauge to the measurement uncertainty surrounding the NAIRU estimates. The paper first develops a consistent conceptual and analytical framework in which the NAIRU can be identified and goes on to review a range of empirical methods used in a number of existing studies. On this basis it then

develops a general empirical framework for estimating the NAIRU across a range of countries. It then discusses the resulting econometric estimates obtained by applying the new procedures to the OECD countries, and the scope for their further refinement given the associated range of uncertainties. And, finally, it reviews recent trends in the NAIRU estimates obtained and illustrates how they can be used to analyse inflation developments and monetary policy.

CONCEPTUAL FRAMEWORK AND RECENT STUDIES

The NAIRU and the Phillips Curve

The dominant view among economic analysts is that there is not a long-term trade-off between inflation and unemployment: in the *long run*, unemployment depends on essentially *structural* variables, whereas inflation is a monetary phenomenon.³ In the short term, however, a trade-off exists such that if unemployment falls below the NAIRU, inflation will rise until unemployment returns to the NAIRU, at which time inflation will stabilise at a permanently higher level. The existence of a NAIRU therefore has immediate implications for the conduct of economic policies, in that: macroeconomic stimulus alone cannot permanently reduce unemployment; and any short-term improvements relative to the NAIRU resulting from stimulative policy actions will be reflected in progressively higher rates of inflation.⁴

The simplest theoretical framework incorporating the NAIRU concept in a transparent fashion is the expectations-augmented Phillips curve, which is also consistent with a variety of alternative structural models.⁵ In particular, as illustrated in the Appendix, it can be derived from structural wage-price setting models of the type described by Layard *et al.* (1991). The Phillips curve also has a long empirical tradition of being used as a means of estimating NAIRU indicators. Refinements of the empirical specification led Gordon (1997) to summarise it in terms of the so-called “triangle model” with inflation being determined by three factors: expectations/inertia, the pressure of demand as proxied by unemployment and supply factors.

Inflation expectations are often slow moving, which means that the effects of demand pressures or supply shocks get built into the inflation process only gradually. With regards to demand pressures, unemployment may be important not just in terms of its level, but also its recent movements. For example rapidly falling unemployment may put upward pressure on inflation even at high levels of unemployment; an effect sometimes referred to as a “speed limit”.

Taking appropriate account of supply shocks is important in order to distinguish between one-off price changes and ongoing inflation. An important distinction to make here is between temporary and long-lasting supply shocks.⁶ Temporary supply shocks (for example, *changes* in real import prices or *changes* in real oil prices) are typically those which are expected to revert to zero over the horizon of one to two years, that is particularly relevant to monetary policy. Such temporary shocks may alter the rate of inflation, at any given rate of unemployment, but the NAIRU will be largely unchanged once they have passed.⁷ By contrast, a long-lasting supply shock (caused by factors such as the level of real interest rates, the tax wedge, demographics, *etc.*) may permanently alter the NAIRU, so that inflation will rise or fall until unemployment adjusts.

Within such a framework, it is useful to identify three distinct concepts (see Box 1 for more formal definitions): the NAIRU (with no qualifying adjective), the short-term NAIRU and the long-term equilibrium rate of unemployment. Each of these relate to the same basic idea of an “unemployment rate consistent with stable inflation”, but differ according to the time horizon to which they refer:

- The NAIRU is defined as the rate towards which unemployment converges in the absence of *temporary* supply influences (in the medium term or when their effects dissipate), once the dynamic adjustment of inflation is completed.
- The *short-term* NAIRU is defined as that rate of unemployment consistent with stabilising the inflation rate at its current level in the next period (where the precise time frame is defined by the specific frequency used in the inflation analysis, for example, the next quarter, the next semester, or the next year). It depends on the NAIRU (as defined above) but is *a priori* more volatile because it is affected by *all* supply influences, including temporary ones, expectations and inertia in the dynamic process of inflation adjustment and possible related speed-limit effects. It follows that the short-term NAIRU concept will be influenced also by the level of actual unemployment.
- The *long-term* equilibrium unemployment rate corresponds to a long-term steady state, once the NAIRU has fully adjusted to *all* supply and policy influences, including those having long-lasting effects.

Of these three concepts, the first two are relatively straightforward to identify empirically and play clearly defined roles in macroeconomic analysis and policy assessments. Because of difficulties in identifying the effects of individual long-lasting supply influences, the long-term equilibrium rate of unemployment is less easy to quantify empirically. However, while important for structural policies, the long-term equilibrium rate may be of limited relevance to macro policy, especially if the complete adjustment of the NAIRU towards the long-run equilibrium is very

Box 1. Three NAIRU concepts

As shown in the Appendix, the expectations-augmented Phillips curve relationship can be derived as a reduced form equation of structural wage and price setting models of the type described in Layard *et al.* (1991), which can be expressed using the following two-equation system. The first equation (1) identifies explicitly only the temporary supply shocks and the second expression (2) includes the long-lasting supply shocks, which fundamentally determine the NAIRU, subject to various long-term adjustment lags.¹

$$\Delta\pi_t = \alpha(L) \Delta\pi_{t-1} - \beta(U_t - U^*_t) - \theta(L) \Delta U_t + v(L) ZT_t + e_t, \quad (1)$$

$$U^*_t = [K_t + \gamma(L) ZL_t] / \beta \quad (2)$$

where Δ is the first difference operator, π_t is inflation, U_t is the observed unemployment rate, ZL_t and ZT_t are vectors of respectively long-lasting and temporary supply shock variables, $\alpha(L)$, $\theta(L)$, $\gamma(L)$ and $v(L)$ are polynomials in the lag operator and e a white noise error term. K_t is a moving parameter capturing all other unspecified influences on the NAIRU.²

On the basis of these equations, three distinct NAIRU concepts can be identified:

- i) The NAIRU, with no qualifying adjective, which is U^*_t in equation (2).
- ii) The short-run NAIRU, US^*_t , is the value of U_t in expression (1) for which the inflation rate is stabilised at that of the previous period, i.e. $\Delta\pi_t = 0$, for a given NAIRU, U^* .³

Equation (1) can hence be rewritten as follows, using the short-term NAIRU concept:

$$\Delta\pi_t = -[\beta + \theta(0)] (U_t - US^*_t) + e_t,$$

$$\text{where } US^*_t = g\{U^*, \Delta U_{t-1}, \alpha(L)\Delta\pi_{t-1}, v(L)ZT_t\} \quad (3)$$

- iii) The long-run equilibrium rate of unemployment, UL^*_t , which is the value of the NAIRU (U^*) associated with a particular realisation of the lasting supply shocks ($ZL_t = zl$) once there has been full adjustment:

$$UL^*_t = f\{K_t + \gamma(1)zl\} / \beta \quad (4)$$

The particular realisation of the supply-shock variables for which the long-run NAIRU is evaluated might for example be based on a projection or represent a view about the long-run steady-state of the supply shocks.

On this basis, the distinction between the NAIRU and short-run NAIRU, is given by equation (3) as a function of the temporary supply shocks and the estimated dynamics of the Phillips curve, including differenced unemployment terms (ΔU_t). The distinction between the NAIRU and long-run equilibrium

Box 1. Three NAIRU concepts (cont.)

unemployment rate concerns the speed of adjustment to long-lasting shocks (captured by the lag polynomials $\gamma(L)$ in equation (2)), and not the specific dynamic terms in the Phillips curve.

1. Equation (2) might possibly be better represented as a non-linear function of supply-shocks. For example, Blanchard and Wolfers (1999) argue that the NAIRU is a function of the interaction of supply shocks and labour market institutions and the latter may change over time.
2. This parameter might for example take into account structural and institutional factors influencing the functioning of the labour and commodity markets, including those related to the cost of gathering information about job vacancies and labour availability, and the costs of mobility.
3. The relevant time period necessarily corresponds to the frequency of equation (1).

protracted. It is particularly important to be sure that when comparing empirical estimates across different studies that they relate to a similar concept of the NAIRU.

Estimation methods in recent studies

Since the NAIRU concept is unobservable it needs to be quantified before it can be useful for policy analysis. Numerous estimation methods exist, which can be divided broadly into three categories: structural, statistical and reduced-form methods. The first group of so-called “structural methods” involves modelling aggregate wage and price setting behaviour in structural form. The NAIRU is then derived from these estimated systems, assuming that markets are in full or sometimes partial equilibrium. The second group of methods attempt to estimate the NAIRU using a variety of purely statistical techniques to directly split the actual unemployment rate into cyclical and trend components, with the latter identified as the NAIRU. The third group constitutes a compromise between the two approaches already outlined. Similarly to structural methods, they allow the NAIRU to be estimated on the basis of a behavioural equation explaining inflation; typically the expectations-augmented Phillips curve. However, they also rely on statistical techniques to impose certain identifying constraints on the path of the estimated NAIRU and/or the gap between it and the actual rate of unemployment. The rest of this section reviews the main features of these three approaches in turn, drawing on a range of recent studies.

Structural methods

Structural methods for quantifying the NAIRU typically involve estimating a system of equations explaining wage- and price-setting behaviour. These can either take the form of wage and price equations specified in levels form (see, for example, Layard *et al.*, 1991, Phelps, 1994, Cotis *et al.*, 1996, Broer *et al.*, 1998, L'Horty and Rault, 1999) or a more *ad hoc* system in which wage determination is represented by an expectations augmented Phillips curve and prices as a mark-up over unit labour costs (for example, Englander and Los, 1983). Given such specifications, an equilibrium level of unemployment can be derived as the set of values such that inflation is stable subject to firms' and workers' decisions regarding profit margins and real wages being compatible. Because such an equilibrium rate of unemployment typically assumes full adjustment of firms and workers behaviour to all shocks, the derived measure of equilibrium unemployment corresponds more closely to a measure of the long-run equilibrium rate of unemployment rather than the NAIRU which commonly appears in reduced-form Phillips curve specifications.

Structural models can provide a strong theoretical framework to explain how various macroeconomic shocks and more importantly policy instruments impact on structural unemployment, but for several reasons they do not allow specific estimates of the NAIRU to be identified with any degree of precision.

Firstly, there is considerable disagreement about the appropriate structural model to be used. For example, Rowthorn (1999) argues that the assumption of a unit elasticity of substitution between capital and labour underlying the widely used model of Layard *et al.* (1991) is implausible and leads to misleading conclusions. More generally there is disagreement from both a theoretical and empirical perspective concerning the long-run effects of changes in real interest rates, taxation and productivity growth on real wages and equilibrium unemployment.

Second, abstracting from the lack of broad agreement on the appropriate theoretical framework, there is little consensus on specification issues. Some of these issues, such as those concerning the modelling of inflation expectations or the functional form (in particular, whether or not the unemployment gap should take a linear form and whether or not it is symmetrical with respect to its effect on inflation) are also common to reduced-form modelling with a Phillips curve. However, a more general specification problem with structural modelling concerns the number and identity of explanatory variables, which is potentially large, and the sensitivity of results to the particular subset of variables chosen for inclusion in the model. This is, itself, an important limitation when the objective is to apply the same specification across many countries.⁸

Third there is a statistical identification problem regarding the estimation of both wage- and price-setting equations, to the extent that all explanatory vari-

ables which enter the former should also enter the latter, as is often suggested by theory (see Bean, 1994 and Manning, 1993). For some countries, notably the United States, it appears difficult to estimate a wage curve based on macroeconomic data because the influence of the (lagged) level of the real wage is often poorly determined, although the reasons for this result are not clear (see Blanchard and Katz, 1997).

Finally, there is considerable difficulty in quantifying many of the relevant institutional variables, such as unemployment benefits, employment protection legislation and the degree of unionisation which theory suggests might be important. Omission of such variables might be particularly problematic given the increasing recognition that the inter-action between institutional factors and macroeconomic shocks plays a key role in determining structural unemployment (Blanchard and Wolfers, 1999).

To overcome the paucity of data relating to the measurement of institutional variables, an increasing number of studies have pooled country information in order to estimate either reduced-form or structural-wage equations or reduced-form unemployment-equations.⁹ This body of work has already provided some very important insights into the causes of structural unemployment. For example, the link between the generosity of benefits and structural unemployment is one of the most robust results in this empirical literature. However, while there is some agreement on the relevant macroeconomic variables (real interest rates, productivity growth, the wage share, the tax wedge, *etc.*) to be used in conjunction with a standard set of institutional variables in empirical studies, there is little or no consensus regarding their relative importance in determining structural unemployment.¹⁰ Nevertheless, structural methods that use pooled country data probably represent the most promising approach for improving understanding of the causes of changes in structural unemployment. However, their usefulness in providing timely estimates of the NAIRU is limited to the extent that it is difficult to obtain reliable and up-to-date time series data on many of the key institutional variables. In such studies it is usually necessary to divide the analysis into sub-periods of several years, where the final period considered is often several years in the past.

Purely statistical methods

Statistical methods focus entirely on the actual unemployment rate and split it into trend (NAIRU) and cyclical components. The assumption behind these approaches is that, since there is no long-term trade-off between inflation and unemployment, on average unemployment should fluctuate around the NAIRU *i.e.* self-equilibrating forces in the economy are strong enough to bring unemployment back to trend.

A wide range of statistical techniques have been developed to decompose time series such as the unemployment rate into cyclical and trend components.¹¹ The basic problem with all these methods is that they depend on arbitrary and sometimes implausible assumptions in order to make this decomposition. Such assumptions typically relate to the way the estimated trend is modelled, its variance and relationship with the cyclical component. For example, in the case of the Hodrick Prescott (HP) filter, trend unemployment is identified as a weighted moving average of actual unemployment, whereas it is assumed to be a random walk by the methods of Watson (1986) and Beveridge and Nelson (1981).¹² More importantly, since all information other than unemployment is ignored (notably the link between the unemployment gap and inflation) the indicators obtained are conceptually not well defined. In practice, trend unemployment estimated with these approaches is usually “centered” around actual unemployment by construction. This is in particular the case of the HP filter, which because of its simplicity is the most frequently used method. Whilst this may be a reasonable approximation when inflation is roughly stable over the estimation period, the estimated NAIRU is likely to be biased when, for example, inflation is falling.

Overall, whilst statistical methods allow indicators of trend unemployment to be estimated in a timely and consistent way across OECD countries, they suffer from a number of practical drawbacks. First, the estimated indicators are often not very well correlated with inflation and are difficult to extrapolate even in the short term.¹³ Second, they tend to be least reliable at the end of the sample, the period of most interest for policy, although this problem can often be mitigated by adding a few years of forecasts to the end of the data sample, which has become standard practice. Third, most of the filters behave like simple moving averages and so perform poorly if there is a large and sudden change in the unemployment rate, for example as occurred for example in Finland and Sweden in the late 1980s and early 1990s. Fourth, there is often no way to judge the degree of precision of the results. Consequently, these methods are seldom used in recent studies to estimate NAIRUs, especially as better alternative methods are now available.

The reduced-form approach

Of the various approaches used to calculate the NAIRU, the most popular technique in recent studies is based on the expectation-augmented Phillips curve. This approach, which follows a relatively long empirical tradition, has the major advantage of being directly related to the definition of the NAIRU, *i.e.* the NAIRU is derived as that rate of unemployment which is consistent with stable inflation, subject to an expectations-augmented Phillips curve relationship. In addition, its relative simplicity and transparency make it consistent with a variety of alternative structural models and hence it is *a priori* likely to be more robust to specification errors than the corresponding structural approach.

Within this framework, as for the purely statistical approach, some identification is required to estimate the NAIRU. The simplest case is to assume the NAIRU to be constant through time (Fortin, 1989; Fuhrer, 1995; Estrella and Mishkin, 1998). For the analysis of periods as long as thirty years, this may be a valid assumption if the observed unemployment rate appears to evolve around a stable mean (as for the United States). However, this clearly is not the case for countries, such as those in Continental Europe, where the unemployment rate has trended upwards since the late 1970s (see Cotis *et al.*, 1996, for France, and the Fabiani and Mestre, 1999, for the Euro area as a whole). In such cases, a constant rate is unlikely to provide a meaningful estimate (Setterfield *et al.*, 1992).

One of the first attempts at estimating time-varying NAIRUs was developed by Elmeskov (1993) and subsequently used by the OECD.¹⁴ This method essentially infers movements in the NAIRU from changes in (wage) inflation based on the notion of an underlying Phillips curve. It is relatively simple and gives plausible and up-to-date indicators for all OECD countries. However there are ways in which this method might be improved. First, the concept could be better defined: *a priori* it is based on a short-term NAIRU concept, but this feature is weakened by smoothing over time (which makes it closer to the “unqualified” NAIRU notion). Second, the Phillips curve relationship could be more sophisticated and the link with inflation strengthened (Holden and Nyomoen, 1998).

More sophisticated estimation techniques help achieve some of these improvements. For example, the Kalman filter, which is used often in the recent literature, allows simultaneous estimation of the NAIRU and the Phillips curve. It also provides some measure of the statistical uncertainty surrounding the NAIRU.¹⁵ In this framework, the estimated NAIRU is time varying, derived from its ability to explain inflationary developments, subject to various constraints on its evolution over time. Such a NAIRU estimate is hence obtained without requiring all factors affecting it to be specified explicitly. In recent years, there has been a proliferation of studies using the Kalman filter in this way. The majority of these were initially applied to the United States, where prominent studies include Gordon (1997 and 1998), King *et al.* (1995), Staiger *et al.* (1997a), but it is now increasingly applied to other countries.¹⁶

As demonstrated in a later section, there is no unique way of using the Kalman filter to estimate the NAIRU. A variety of assumptions may be adopted for the behaviour of the NAIRU or the unemployment gap. In the empirical literature, the most commonly adopted assumption is to specify a random walk for the NAIRU model, although other forms are possible. A closely related case is the HPMV filter, which is an augmented version of the HP filter, and was developed by Laxton and Tetlow (1992).¹⁷ This filter (which, as shown by Boone (2000), belongs to the same class of models) uses a Phillips curve but with a specific restriction on the properties of the unemployment gap.

Overall, reduced-form filtering methods have several important advantages over both the statistical and structural methods. First, by construction, they provide NAIRU estimates directly related to inflation. Second, the fully specified Phillips curve allows the distinction between NAIRU and short-term NAIRU concepts within the same framework. Third, such indicators can be easily produced in a timely and consistent fashion across OECD countries.¹⁸

Despite these attractions, filtering methods also suffer from certain drawbacks. The estimated NAIRU indicators are based on a reduced-form equation, which means that the underlying structural relationships themselves are not identified. This may make it more difficult to extrapolate the NAIRU, especially when the estimated Phillips curve incorporates only temporary supply shocks. The relationship between inflation and unemployment over time also needs to be stable and well specified.¹⁹ The corresponding NAIRU estimates are also likely to be dependent on the specification of the Phillips curve.²⁰

In spite of these limitations, a general conclusion of this review is that filtering methods within a reduced-form Phillips curve framework provide a number of improvements on previous methods for estimating NAIRUs on a timely basis across the range of OECD countries. The following section reports the results of their specific application to these countries and discusses also a range of practical issues arising in their use.²¹

THE OECD APPROACH TO ESTIMATING THE NAIRU

The estimation framework: the choice of inflation and supply shock variables

Following Gordon's "triangle model", the Phillips curve estimation framework can be expressed in the following form:

$$\Delta\pi_t = \alpha(L) \Delta\pi_{t-1} - \beta(U_t - U^*) - \theta(L) \Delta U_t + \gamma(L)z_t + e_t, \quad (1)$$

where Δ is the first difference operator, π is inflation, U is the observed unemployment rate, U^* is the NAIRU, z a vector of temporary supply shock variables, $\alpha(L)$, $\theta(L)$ and $\gamma(L)$ are polynomials in the lag operator and e is a serially uncorrelated error term with zero mean and variance σ^2 . As previously emphasised, only temporary supply shock variables, defined here to be those that might reasonably be expected to revert to zero over a future horizon of 1 to 2 years, are included in the Phillips curve specification. The NAIRU is then estimated with the Kalman filter, to implicitly capture the aggregate effect of all long-lasting shocks, without requiring these shocks to be explicitly identified.

In estimating equation (1), a number of choices have to be made regarding the specification of the dependent and explanatory variables. In principle, theory

suggests that the dependent variable could be either a measure of price inflation or wage inflation, where the latter is adjusted in relation to productivity or trend productivity. In deriving a reduced-form Phillips-type inflation equation from structural wage- and price-setting equations, (as shown in the Appendix) it is possible to substitute out either wages or prices. Hence if there is a stable relationship between wages and prices, then the choice of which to use is not clear-cut. In the empirical work reported here, an inflation measure based on the private consumption deflator is used on the grounds that this is more representative of inflation measures targeted by policy-makers and central banks in most OECD countries, although, for some countries, the robustness of the results to using an alternative measure of wage inflation has also been examined.²² For Canada, a measure of core inflation (excluding food and energy, as used by the Bank of Canada) was found to give more robust results and is used to provide the preferred NAIRU estimates. The unemployment variables used are as defined in the notes to the relevant tables, which for most countries correspond to the national definitions commonly used in the OECD macroeconomic projections.

In practice, the choice of temporary supply shock variables to be included was largely governed by those variables found most often to be statistically significant across the range of country specifications. In particular these include the *change* in real import prices (weighted by the degree of openness of the economy) and the *change* in real oil prices (weighted by a measure of the degree of oil intensity in production).²³ Other possible variables, for example, tax wedge terms and the deviation of productivity growth from trend, were tested in preliminary estimation but found to be much less successful and are not included in the final specifications reported here. Temporary variations in the mark-up of prices over unit labour costs are also a candidate, provided that the mark-up tends to return to trend within the time horizon relevant to monetary policy. For example, Brayton *et al.* (1999) suggest that low inflation in the United States in recent years may partly result from mark-ups returning to their historical norm. A particular concern related to the choice of temporary supply shocks included is that for most OECD countries real import prices have been trending downward over at least the last two decades, so that the expected change in real import prices over the near future (in the absence of other shocks) is likely to be negative rather than zero. For this reason, real import prices were first de-trended by regressing them on split time trends and their own lagged values.²⁴

A further issue is whether the unemployment gap ($U-U^*$) should enter linearly or non-linearly. For simplicity, a linear specification was initially assumed for all countries. However, it became clear that this was not a reasonable approximation for some countries, particularly those in which unemployment had risen considerably over the past three decades. A linear specification assumes, for example, that unemployment at 3 per cent when the NAIRU is 2 per cent has the

same impact on inflation as unemployment at 12 per cent when the NAIRU is 11 per cent. This does not seem economically plausible and, in fact, led to structural breakdowns of some estimates.²⁵ For Belgium, Spain, Finland and Sweden a partial solution was to have the unemployment gap enter in logarithmic terms: $\log(U/U^*)$.²⁶ For Australia, consistent with academic and official studies, a non-linear gap $(U-U^*/U)$ was found to significantly improve the robustness and significance of the estimates.²⁷

Specifying the Kalman filter

There is no unique way of using the Kalman filter technique to estimate the NAIRU, but the approach followed here is similar to that of most other studies, namely augmenting the Phillips curve, as represented by equation (1) (which is referred to as the measurement equation) with one or more additional equations, defining how the NAIRU varies over time – the transition equations (see Box 2 and Boone (2000) for further technical details concerning the specification and use of the Kalman filter). In the empirical literature, the most commonly adopted form for the transition equation is a random walk (2a) below, which is used in the work reported here, as well as an alternative specifying the *change* in the NAIRU as a first order auto-regressive process (2b).²⁸

$$\Delta U^*_t = v^1_t, \text{ where } v^1_t \sim N(0, \sigma_{v1}^2) \quad (2a)$$

or

$$\Delta U^*_t = \phi \Delta U^*_{t-1} + v^2_t, \text{ where } 0 < \phi < 1 \text{ and } v^2_t \sim N(0, \sigma_{v2}^2). \quad (2b)$$

Where possible both the random walk and auto-regressive forms were estimated and the choice between the two was based largely on the statistical significance of the autocorrelation coefficient and the fit of the respective unemployment gaps in the estimated Phillips curve. The assumption of a first order auto-regressive process is of particular interest for some, mainly European countries, because it may provide evidence of slow adjustment of the NAIRU to long lasting supply shocks.

Determining the smoothness of the NAIRU

When using the Kalman filter, the volatility or smoothness of the resulting NAIRU series is determined by the magnitude of the variance of the errors in the transition equation (σ_{v1}^2 in (2a)) relative to those in the inflation equation (σ^2 in (1)). The larger is this ratio (the “signal-to-noise” ratio) the more volatile will be the NAIRU series which, in the limit, soaks up all the residual variation in the Phillips curve equation.

In principle, the Kalman filter technique makes it possible to estimate all the parameters of the model using a maximum likelihood estimation procedure,

Box 2. Using the Kalman filter to estimate a time-varying NAIRU

The Kalman filter is a convenient way of working out the likelihood function for unobserved component models.¹ For that, the system must be written in a state space form, with a **measurement equation** (the Phillips curve):

$$\Delta\pi_t = \alpha_1\Delta\pi_{t-1} + \alpha_2\Delta\pi_{t-2} + \beta(U_t - U^*_t) - \theta\Delta(U_t - U^*_t) + e_t \quad (1)$$

in a matrix format: $y_t = Z_t X_t + R_t D_t + e_t$ (1')

where Z and R are vectors of parameters, X is a vector of unobserved variables (the NAIRU), while D is a vector of observed exogenous variables (lagged inflation, temporary supply shocks)

and a **transition equation**:²

$$U^*_t = U^*_{t-1} + \varepsilon_t \quad (2)$$

in a matrix format: $X_t = T X_{t-1} + \varepsilon_t$ (2')

where e_t and ε_t are iid, normally distributed with mean zero and variances $H_t = \sigma^2$ and $q_t = \sigma^2$. Q respectively. The ratio $q_t/H_t = Q$ is called the signal-to-noise ratio. T is a vector of parameters.

The Kalman filter is made up of two stages:

1. The **filtering procedure** builds up the estimates as new information about the observed variable becomes available. If a_t is the optimal estimate of the state variable X_t (the NAIRU) and P_t its variance/covariance matrix, then, given a_{t-1} and P_{t-1} , the Kalman filter may be written:³

$$a_{t+1|t} = (T - K_t Z) a_{t|t-1} + K_t (y_t - d_t) \quad (3)$$

with $K_t = T P_{t|t-1} Z' F_t^{-1}$ and $F_t = Z P_{t|t-1} Z + H$ (4)

and $P_{t+1|t} = T (P_{t|t-1} - P_{t|t-1} Z' F_t^{-1} Z P_{t|t-1}) T' + Q$ (5)

These equations permit the computation of the **prediction errors** v_t for period t as:

$$v_t = y_t - Z a_{t|t-1} - R D_t \quad (6)$$

to go into the likelihood function :

$$l_t = -\frac{1}{2} \log 2\pi - \frac{1}{2} \log |F_t| - \frac{1}{2} v_t F_t^{-1} v_t \quad (7)$$

The series $\{a_t\}$ that maximises this function gives an optimal estimate of the one-sided NAIRU.

Box 2. Using the Kalman filter to estimate a time-varying NAIRU (cont.)

2. The smoothing procedure uses the information available from the whole sample of observation. It is a backward recursion which starts at time T and produces the smoothed estimates in the order T,...,1, following the equations:

$$a_{t|T} = a_t + P_t^* (a_{t+1|T} - T_{t+1} a_t) \quad (8a)$$

$$P_{t|T} = P_t + P_t^* (P_{t+1|T} - P_{t+1|t}) P_t^{*'} \quad (8b)$$

$$P_t^* = P_t T_{t+1}' P_{t+1|t}^{-1} \quad (8c)$$

with $a_{T|T} = a_T$ and $P_{T|T} = P_T$.

1. Standard references are Cuthbertson, Hall and Taylor (1992), Harvey (1992) and Hamilton (1994).
2. As explained in the main text, other forms of transition equations may be used. This one is used here for ease of presentation.
3. The initial values for a_0 and P_0 are important for the optimisation process to converge. The starting values may cause real trouble if the user of the Kalman filter has no prior information about it: as with all maximisation procedure, if the starting values are too far away from the true values the system will not converge. There is no standard or theoretical procedure to overcome this problem. When it is possible, a practical solution is to realise an OLS estimation first that will give an idea about the value of the parameter in the vector A. Yet, this does not help with the initial value for the variance/covariance matrix. The usual "trick" is to give this matrix an extremely high value so as to go away from the initial values of the parameters very quickly.

including the signal-to-noise ratio. In common with the findings of most other researchers using the technique, directly estimating the signal-to-noise ratio was found to give disappointing results because it typically leads to very flat NAIRUs.²⁹ The usual response to this problem is to carry out sensitivity analysis and choose these variances by visual inspection of the resulting NAIRU estimates. For example, Gordon (1997) suggests adopting a "smoothness prior", so that "the NAIRU can move around as much as it likes, subject to the qualification that sharp quarter-to-quarter zigzags are ruled out". Such an approach was adopted here, taking into account a number of factors including the combination of goodness-of fit and plausibility of the estimated equations and NAIRU estimates. In practice, the relevant parameterisation was found to vary significantly across countries. This reflects the differing time series properties of their actual rates of unemployment, in par-

ticular whether or not they have been stationary, as well as the differing goodness-of-fit of the estimated Phillips curves.³⁰

End-point adjustments

An issue for concern when using filter procedures is the sensitivity of the NAIRU estimates for the most recent observations, which are typically of greatest interest from a policy perspective. A variety of studies (see, for example, Giorno *et al.* (1995) show that without further adjustments, the Hodrick Prescott filter may be “drawn towards” values at the end-point of the sample, thereby reducing the estimated “gap”, whether or not this appropriately reflects the cyclical position of the economy in question. Boone *et al.* (2001) demonstrates that by making use of additional information about inflation, in a Phillips curve framework, Kalman filter estimates of the NAIRU are much less subject to end-point revisions than estimates from an HP filter.

To examine the degree of end-point sensitivity for both Kalman filter and HPMV estimation methods, NAIRU estimates for two countries where the cycle in unemployment has been pronounced, the United States and the United Kingdom, were obtained using truncated and full samples. On this basis, the estimated revisions to the Kalman filter NAIRUs over the period 1990-95 were found to be about one-quarter of a percentage point for the United States, with corresponding revisions for the United Kingdom found to be somewhat larger immediately after a turning point in actual unemployment but otherwise averaged about 0.4 percentage points. These revisions were about half the size of those obtained for a comparable HPMV filter and were judged sufficiently small not to warrant specific treatment. Nonetheless, this analysis suggests that particular caution needs to be attached to NAIRU estimates when the end-point is close to a cyclical turning point or where there are reasons for suspecting that there might be strong movements in the NAIRU, perhaps reflecting the effects of recent policy actions.

The estimation procedure

For most countries, Kalman filter estimation was carried out using a maximum likelihood method with the Phillips curve equation estimated jointly with the transition equation(s). However, for five of the 21 countries, direct estimation failed to produce plausible results because of difficulties in jointly identifying the NAIRU series and the coefficient on the unemployment gap.³¹ For these countries an alternative iterative procedure was used, similar to that used by Fabiani and Mestre (1999), in which the Phillips curve coefficients were first imposed on the basis of preliminary estimates based on the HPMV filter, and an initial NAIRU series then estimated using the Kalman filter.³² The resulting NAIRU series was

then substituted into the Phillips curve equation and the parameters re-estimated using OLS. This process was repeated until the NAIRU series converged, usually within a few iterations.

RESULTS

This section describes the preliminary NAIRU estimates that are obtained from a Phillips curve relationship using a Kalman filter, following the framework previously outlined. However, as discussed, these estimates are subsequently adjusted for possible biases, particularly to allow for the effects of recent policy reforms given the uncertainty surrounding the empirical estimates.

The estimation results

Following the procedures outlined in the previous section, it was possible to estimate Phillips curves and corresponding NAIRU estimates using the Kalman filter method for all 21 OECD countries for which the OECD currently publishes NAIRU estimates (see Table 1). Similar Phillips curve specifications were used across countries to ensure comparability of results.³³ Speed limit effects (ΔU terms) were tested for all countries, but found to be insignificant for most of them. Occasional outlier dummies have also been used in places, such as to account for price controls in the United Kingdom in the 1970s. For the United States, special adjustments were made to the unemployment rate variable to take account of specific demographic composition effects.³⁴

For approximately half of the countries, an auto-regressive process was preferred to a random walk when using the Kalman filter. In nearly all of these cases, the auto-regressive coefficient is statistically significant and typically takes a value in the range 0.6 to 0.8. Examining the four major European countries for which both specifications could be stably estimated, the differences between the two NAIRU series are generally small.³⁵ The average absolute difference over the entire sample estimation period is 0.4 percentage points for France and Italy, one-quarter of a percentage point in the case of the United Kingdom, and 0.1 percentage points in the case of Germany. The maximum difference between the two series over the entire sample period for all four countries is about 0.6 to 0.8 percentage points. These relatively small differences lend some support to the predominant use of the random walk assumption in the empirical literature. Nevertheless, the auto-regressive form is intuitively more appealing because it is consistent with the NAIRU adjusting only slowly to long-lasting supply shocks. Moreover, the auto-regressive form is also of interest in a short-term forecasting context insofar as changes in the estimated

Table 1. Estimated Phillips curves and diagnostic tests using the Kalman filter

Estimation method: Kalman filter

Dependent Variable is $\Delta\pi_t$.

Sample	United States	Japan ¹	Germany	France	Italy	United Kingdom	Canada
	63:2 to 99:2	63:2 to 99:1	62:2 to 99:1	70:2 to 99:2	62:2 to 99:1	63:1 to 99:1	65:2 to 99:1
<i>Dynamics</i>							
$\Delta\pi_{t-1}$	-0.33 (3.1)	-0.49 (5.7)	-0.39 (6.0)	-0.43 (3.3)	-0.11 (1.2)	-0.33 (5.2)	-0.46 (4.6)
$\Delta\pi_{t-2}$	-0.24 (2.5)	-0.44 (7.4)		0.00 (0.0)	-0.33 (4.7)	-0.30 (4.3)	-0.51 (5.5)
$\Delta\pi_{t-3}$		-0.34 (5.0)		-0.21 (2.1)		-0.27 (4.3)	
$\Delta\pi_{t-5}$		-0.22 (4.0)					
<i>Unemployment</i>							
$U-U^*$	-0.13 (4.5)	-1.85 (7.9)	-0.19 (6.0)	-0.17 (3.8)	-0.27 (3.7)	-0.20 (5.6)	-0.50 (8.6)
ΔU					-0.65 (2.6)	-0.26 (1.8)	
$\Delta (U^{North}-U)$					-1.18 (2.7)		
<i>Import prices</i>							
$\omega_{-1}(\pi^m - \pi_{-1})_{-1}$	1.53 (4.6)	1.40 (5.6)	0.52 (3.8)	0.89 (3.6)	0.76 (3.5)	0.45 (2.5)	0.87 (5.6)
$\omega_{-1}\Delta\pi^m$	0.83 (2.8)	0.40 (1.8)	0.34 (2.5)	0.55 (2.5)	0.80 (5.1)	0.16 (1.1)	0.24 (2.1)
$\omega_{-1}\Delta\pi^m_{-1}$		-0.81 (3.3)					-0.43 (3.2)
$\omega_{-1}\Delta\pi^m_{-2}$							
<i>Oil prices</i>							
$v_{-1}(\pi^o - \pi_{-1})_{-1}$						0.21 (1.8)	
$v_{-1}\Delta\pi^o$	0.07 (5.9)	0.13 (2.7)	0.11 (4.0)	0.11 (2.4)	0.11 (3.0)	0.10 (1.3)	
$v_{-1}\Delta\pi^o_{-1}$	0.06 (4.6)	0.16 (4.1)			0.18 (4.8)	-0.21 (2.4)	
$v_{-1}\Delta\pi^o_{-2}$							
NAIRU in 99:1	5.2	3.9	7.8	10.1	10.4	6.7	8.5
Sacrifice Ratio	3.1	0.3	1.8	2.4	1.3	2.4	1.0
<i>Standard error</i>							
R^2	0.32	0.50	0.33	0.55	0.59	0.58	0.44
<i>adjusted-R²</i>	0.67	0.83	0.53	0.57	0.77	0.84	0.59
	0.64	0.80	0.50	0.52	0.74	0.80	0.55
<i>Diagnostic tests³ (p-values reported)</i>							
Chow forecast test ⁴	0.76	0.99	0.87	1.00	0.97	0.99	0.80
RESET test ⁵	0.23	0.14	0.07	0.91	0.05	0.11	0.21
Serial correlation ⁶	0.21	0.02	0.10	0.10	0.19	0.88	0.53
Normality ⁷	0.97	0.71	0.20	0.01	0.65	0.23	0.27
Chow breakpoint ⁸	0.02	0.81	0.00	0.14	0.26	0.16	0.13

Table 1. Estimated Phillips curves and diagnostic tests using the Kalman filter (cont.)
 Estimation method: Kalman filter
 Dependent Variable is $\Delta\pi$.

Sample	Australia ² 66:2 to 99:1	Austria ² 66:2 to 99:1	Belgium ² 71:2 to 99:1	Denmark 66:2 to 99:1	Finland ² 70:1 to 99:1	Greece 75:1 to 99:1	Ireland 78:2 to 99:1
Dynamics							
$\Delta\pi_{-1}$	-0.58 (6.6)	-1.07 (9.8)	0.05 (0.4)	-0.57 (5.1)	-0.76 (8.1)		-0.23 (1.8)
$\Delta\pi_{-2}$		-0.55 (5.0)	-0.52 5.30	-0.43 (4.5)	-0.31 (3.4)	-0.44 (-5.7)	
$\Delta\pi_{-3}$		-0.31 (3.4)					
Unemployment							
$U-U^*$	-0.93 (3.8)	-1.60 (5.6)	-0.66 (3.5)	-0.23 (4.0)	-1.00 (3.8)	-0.34 (-6.4)	-0.22 (4.9)
ΔU							
Import prices							
$\omega_{-1}(\pi^0 - \pi_{-1})_{-1}$	0.77 (3.8)	0.89 (3.1)	0.37 (3.8)	0.46 (1.9)	1.40 (4.7)	1.21 (6.2)	0.28 (3.2)
$\omega_{-1}\Delta\pi^m$	0.54 (3.6)	0.66 (3.3)	0.25 (2.8)	0.71 (3.5)	0.24 (1.3)	0.70 (4.8)	
$\omega_{-1}\Delta\pi^m_{-1}$				0.40 (2.0)			
Oil prices							
$v_{-1}(\pi^0 - \pi_{-1})_{-1}$		0.25 (3.3)	0.11 (2.3)				
$v_{-1}(\pi^0 - \pi_{-1})_{-2}$							
$v_{-1}\Delta\pi^0$		0.13 (2.2)	0.08 (3.7)	0.11 (2.4)		0.09 (2.9)	
$v_{-1}\Delta\pi^0_{-1}$	0.11 (4.0)				0.21 (4.7)		0.16 (2.6)
NAIRU in 99:1	7.0	4.9	8.4	7.8	10.2	7.9	9.0
Sacrifice Ratio	0.4	0.5	0.6	2.2	0.5	1.1	1.4
Standard error	0.60	0.50	0.45	0.64	0.80	0.59	0.66
R²	0.67	0.69	0.73	0.59	0.84	0.69	0.63
adjusted-R²	0.63	0.65	0.69	0.55	0.82	0.66	0.57
Diagnostic tests³ (p-values reported)							
Chow forecast test ⁴	0.84	0.00	0.76	0.94	0.78	0.67	0.91
RESET test ⁵	0.03	0.45	0.35	0.91	0.30	0.13	0.42
Serial correlation ⁶	0.81	0.16	0.51	0.10	0.84	0.48	0.90
Normality ⁷	0.88	0.01	0.77	0.32	0.58	0.80	0.87
Chow breakpoint ⁸	0.25	0.23	0.07	0.78	0.14	0.42	0.51

Table 1. Estimated Phillips curves and diagnostic tests using the Kalman filter (cont.)

Estimation method: Kalman filter
Dependent Variable is $\Delta\pi$.

Sample	Netherlands	New Zealand	Norway	Portugal	Spain ²	Sweden	Switzerland
	72:1 to 99:1	80:1 to 99:1	66:2 to 99:1	70:2 to 99:1	66:2 to 99:1	66:2 to 99:1	78:1 to 99:1
Dynamics							
$\Delta\pi_1$	-0.67 (7.1)	-0.60 (6.0)	-1.02 (11.7)	0.01 (0.1)	-0.63 (5.8)	-0.85 (7.1)	-0.34 (4.0)
$\Delta\pi_2$			-0.43 5.20	-0.35 (4.2)	-0.25 (2.5)	-0.26 (2.0)	-0.40 (5.0)
$\Delta\pi_3$					-0.51 (5.4)	-0.22 (2.0)	
Unemployment							
U-U*	-0.20 (4.7)	-0.62 (7.0)	-1.16 (5.1)	-0.21 (3.9)	-2.62 (6.2)	-0.43 (3.5)	-0.23 (6.2)
ΔU					-0.42 (3.5)		
Import prices							
$\omega_{-1}(\pi^m - \pi_{-1})_{-1}$	0.32 (2.9)	1.72 (7.8)	0.39 (2.2)	0.50 (2.8)	1.58 (5.0)	1.20 (3.7)	
$\omega_{-1}\Delta\pi^m$	0.12 (1.2)	0.64 (5.3)		0.80 (5.5)	0.80 (3.3)	0.95 (3.8)	
$\omega_{-1}\Delta\pi^m_{-1}$		0.20 (1.4)					
Oil prices							
$V_{-1}(\pi^o - \pi_{-1})_{-1}$				0.12 (2.2)			0.57 (6.9)
$V_{-1}(\pi^o - \pi_{-1})_{-2}$			0.49 (4.7)				
$V_{-1}\Delta\pi^o$				0.14 (3.7)			0.35 (5.6)
$V_{-1}\Delta\pi^o_{-1}$							
NAIRU in 99:1	4.8	5.4	3.7	4.7	15.4	5.6	4.1
Sacrifice Ratio	2.1	0.6	0.5	1.6	0.2	1.4	1.9
Standard error	0.49	0.60	1.01	0.80	0.80	1.24	0.36
R²	0.70	0.88	0.75	0.74	0.60	0.57	0.86
adjusted-R²	0.68	0.86	0.73	0.70	0.56	0.53	0.83
Diagnostic tests³ (p-values reported)							
Chow forecast test ⁴	0.96	0.97	0.98	0.74	0.97	0.85	0.17
RESET test ⁵	0.24	0.72	0.71	0.29	0.08	0.28	0.27
Serial correlation ⁶	0.19	0.41	0.10	0.19	0.88	0.58	0.74
Normality ⁷	0.38	0.29	0.72	0.88	0.39	0.58	0.99
Chow breakpoint ⁸	0.73	0.57	0.49	0.01	0.14	0.39	0.17

Table 1. Estimated Phillips curves and diagnostic tests using the Kalman filter (cont.)

Definition of variables:	
All data is semi-annual and is taken from the OECD's Analytical Database (ADB), except where otherwise noted. All inflation rates are expressed as the change in the relevant price index on the previous semester, with the rate not being annualised.	
π	= inflation rate based on private consumption deflator. For Canada a measure of the core CPI, excluding food and energy, is used (source: Statistics Canada).
U	= unemployment rate.
U^{NORTH}	= unemployment rate in the Centre-North region of Italy (source: Bank of Italy with OECD interpolations).
U^*	= the NAIRU which is estimated using the Kalman filter.
π^m	= inflation rate of the de-trended non-oil import price of goods and services. Import prices were de-trended by regressing real import prices on split time trends and lagged real import prices. The time trends included were one covering the entire sample estimation period or one beginning in 1980.
ω	= weight of non-oil import prices in total demand, measured as the share by value of imports of goods and services (excluding oil) in total demand.
π^o	= inflation rate of the unit value of energy imports.
v	= measure of oil supply in relation to GDP (source: Energy Balances of OECD Countries, International Energy Agency). Semi-annual values were interpolated from annual figures and most recent values were derived by extrapolation.
Δ denotes the first difference operator, subscripts denote lags.	
1. For Japan the recent forecast performance of the estimated Phillips curve was substantially improved by introducing a reduced effect from the unemployment gap on inflation when inflation is already low. Specifically, when the level of inflation is below 2 per cent per annum and unemployment is above the NAIRU the coefficient on the unemployment gap is reduced in magnitude from that shown in the table: to -1.04 (from -1.85) when estimated with the Kalman filter. 2. For some countries the unemployment gap term does not take a linear form: for Austria, Belgium Finland and Spain in the equation estimated with the Kalman filter, the unemployment rate and NAIRU are specified in logarithmic form, $\log(U/U^*)$; for Australia the (linear) unemployment gap is normalised on the actual rate, $(U-U^*)/U$. 3. The p-values of the diagnostic tests are reported. Failures at the 5 per cent significance level are highlighted in bold. 4. Chow forecast test from 1995:1. 5. Ramsey reset test of functional form based on the inclusion of squared and cubed fitted values. 6. Breusch-Godfrey Lagrange-Multiplier test for up to second order serial correlation of the residuals. 7. Jarque-Bera test for normality of residuals. 8. Chow breakpoint test for break in 1985:1, except for Ireland and New Zealand where a break of 1990:1 was chosen because of the shorter sample estimation period.	
List of dummy variables used in the estimation:	
France	+1 in 1982:2; -1 in 1983:1
United Kingdom	-1 in 1977:2 and +1 in 1979:2; 1 in 1974:1; 1 in 1974:2 and 1975:1
Germany	-1 in 1992:2 and +1 in 1993:1; +1 in 1991:2
Italy	+1 in 1970:1 and -1 in 1970:2; -1 in 1971:2 and +1 in 1972:2; -1 in 1984:2
Japan	+1 in 1974:1 and -1 in 1974:2
Australia	+1 in 1973:2; +1 in 1976:2
Belgium	+1 in 1985:2
Finland	+1 in 1973:2
Ireland	+1 in 1982:1, 1982:2 and 1993:2
Netherlands	+1 in 1985:1, -1 in 1985:2
New Zealand	+1 in 1983:1
Norway	+1 in 1970:1
Portugal	+1 in 1976:1 and -1 in 1976:2
Switzerland	+1 in 1980:1, 1983:1 and 1985:2

NAIRU over the recent past may provide information relevant to its likely future profile.

The temporary supply shock (non-oil import and oil-price inflation) and unemployment gap terms are correctly signed and statistically significant for nearly all countries. In order to test the robustness of the Phillips curve, the estimated unemployment gaps were included in the preferred Phillips curve specification, which was then estimated by OLS and subject to a battery of standard diagnostic tests, as reported in Table 1.³⁶ Among the G7 countries the most serious diagnostic test failure relates to the structural stability (using a Chow break-point test) for Germany which may be related to the effects of reunification. For Italy the inclusion of a country-specific variable, namely the change in the difference between the unemployment rate in the Centre-North region of the country

Table 2. **NAIRU estimates and standard errors**

	1980	1985	1990	1995	1999	Standard errors ¹	
						Average	Final year
Australia	5.1	6.0	6.5	7.1	6.8	1.0	1.6
Austria	1.9	3.2	4.6	5.0	4.9	0.2	0.3
Belgium	5.5	6.8	8.4	8.0	8.2	1.3	1.3
Canada	8.9	10.1	9.0	8.8	7.7	0.6	0.9
Denmark	5.8	5.9	6.9	7.1	6.3	1.0	1.3
Finland	4.3	3.9	5.6	10.6	9.0	1.4	1.8
France	5.8	6.5	9.3	10.3	9.5	1.1	1.7
Germany	3.3	4.4	5.3	6.7	6.9	0.9	1.2
Greece	4.6	6.5	8.4	8.8	9.5	0.8	1.1
Ireland	12.8	13.2	14.1	10.8	7.1	1.2	2.0
Italy	6.8	7.8	9.1	10.0	10.4	0.8	1.1
Japan	1.9	2.7	2.2	2.9	4.0	0.2	0.3
Netherlands	4.7	7.5	7.5	6.1	4.7	1.0	1.3
New Zealand	1.6	5.1	7.0	7.5	6.1	0.6	0.8
Norway	2.2	2.6	4.6	4.9	3.7	0.5	0.6
Portugal	6.1	5.4	4.8	4.2	3.9	1.0	1.4
Spain	7.8	14.4	17.4	16.5	15.1	1.2	1.2
Sweden	2.4	2.1	3.8	5.8	5.8	0.8	1.0
Switzerland	2.3	2.9	3.0	3.3	2.4	0.8	1.0
United Kingdom	4.4	8.1	8.6	6.9	7.0	1.1	1.5
United States	6.1	5.6	5.4	5.3	5.2	0.9	1.2
<i>Memorandum items:</i>							
Euro area	5.5	7.1	8.8	9.2	8.8		
Weighted average of above countries ²	5.0	5.9	6.3	6.5	6.5		

1. Estimated standard errors around initial econometric estimates.

2. Weighted by size of labour force.

Source: OECD Secretariat calculations.

and the aggregate unemployment rate, is needed to pass the test for structural stability.³⁷

Measures of uncertainty and revisions to the preliminary estimates

A particular advantage of the Kalman filter is that when a direct maximum likelihood estimation method is used it is also possible to generate standard errors for the NAIRU estimates.³⁸

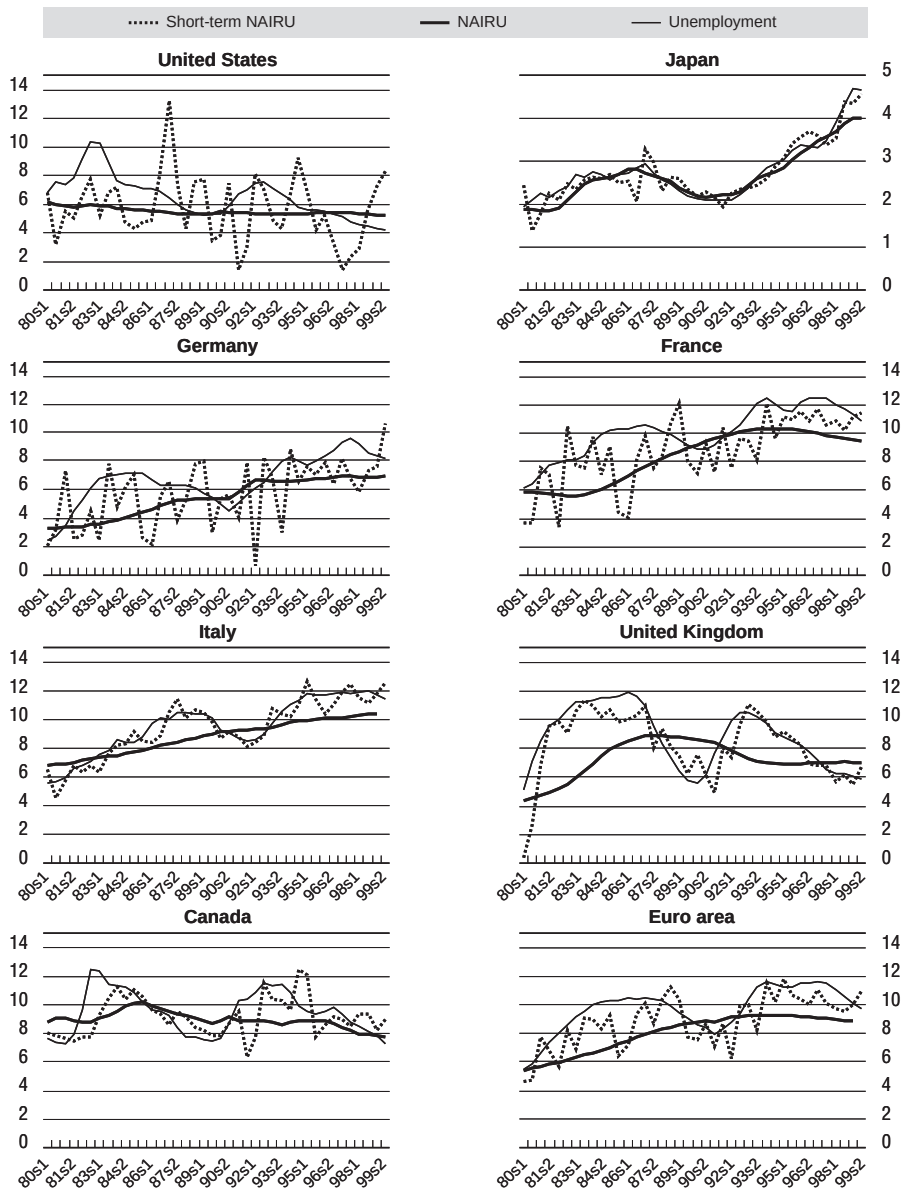
There are three sources of uncertainty surrounding the Kalman filter estimates; those because: *i*) the NAIRU is unobserved and has to be inferred; *ii*) the parameters of the model are unknown and must be estimated; and *iii*) the model specification may be wrong. In the empirical literature, the third source of uncertainty is typically ignored. The estimation techniques used normally provide a means for dealing with the first source of uncertainty in terms of the estimated prediction error variance for the state variable, at each point of time, whilst in some studies, the second source on uncertainty is dealt with either by means of Monte Carlo methods (Laubach (1999), Irac (1999)), or by using the Ansley and Kohn delta' method (Staiger *et al.* 1997).

In the present study the standard errors associated with the first two sources of uncertainty were derived by use of Monte Carlo methods, following Hamilton (1986, 1994).³⁹ For the full sample (see Table 2), these are found to range between 0.2 for Japan and 1 for France, the United Kingdom and Ireland, and are well in line with those reported elsewhere in the recent literature.⁴⁰ Figure 2, which follows, illustrates the corresponding error bands and their evolution over time for the major economies.

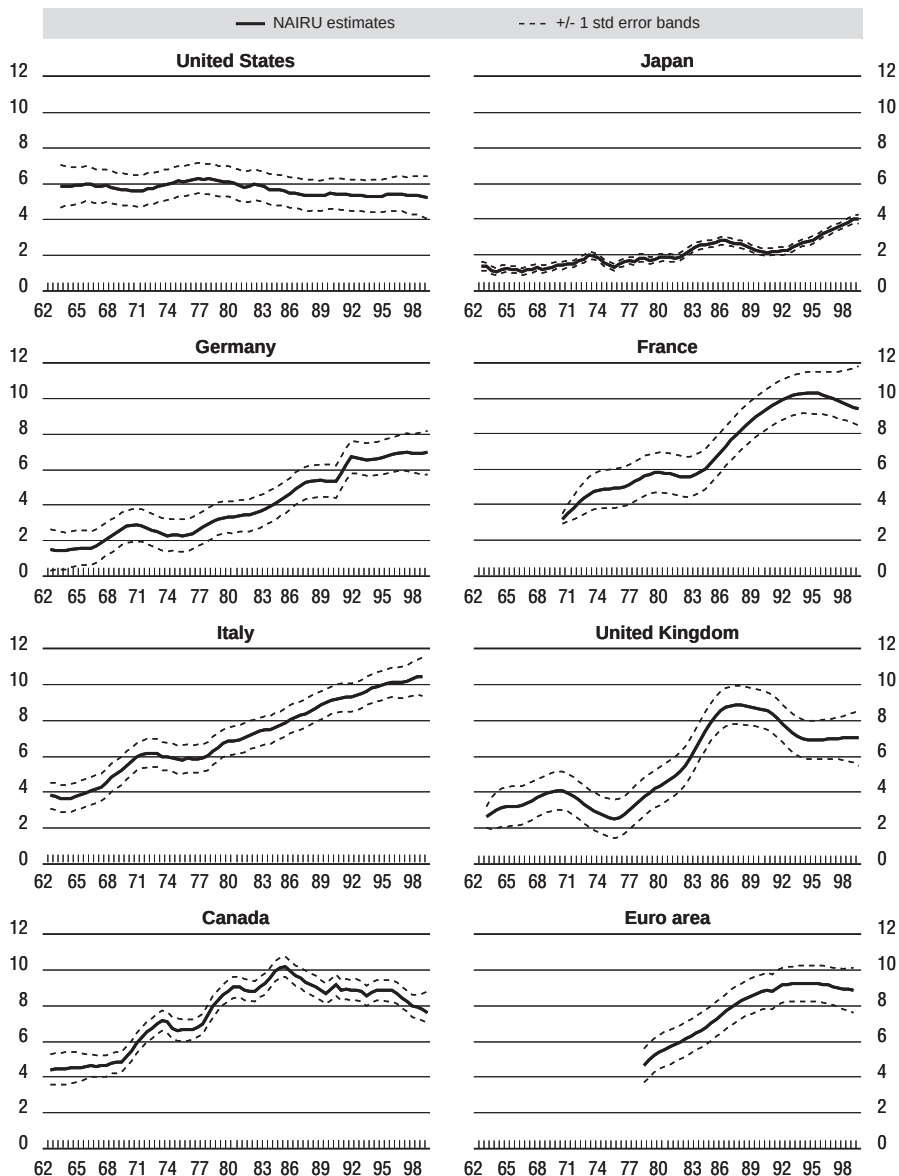
Given the range of uncertainties and the estimated error bands, the estimates generated by the econometric procedure described above were subject to scrutiny by country experts and sometimes revision for specific biases particularly to allow for the effect of recent reforms. In some cases these revisions simply involved using a more appropriate definition of inflation or unemployment in the Phillips curve estimation, which led to a better fitting Phillips curve and a profile for the NAIRU that was judged to be more plausible.⁴¹

For two countries (Canada and Greece) a more fundamental change of specification to the Phillips curve involved more explicit modelling of inflation expectations. For a further three countries (Australia, France and Switzerland) the preliminary estimates appeared to contradict other information, particularly relating to the likely effect of recent labour market reforms, and so were judgementsally adjusted. These latter revisions occur at the end of the estimation period where uncertainty surrounding any filter-based estimates of the NAIRU is greatest.⁴² Two countries (Finland and Ireland) were considered as special cases in so far as the

Figure 1. NAIUR and short-term NAIUR¹



1. Japan is shown on a different scale.
Source: OECD.

Figure 2. NAIRU estimates and standard error bands¹


1. Estimated standard errors are centred around the initial econometric estimates. For France and Canada, where these initial estimates are judgementally revised (see appendix) the NAIRU is not in the centre of the band.

Source: OECD.

basic estimation framework was considered inadequate for explaining recent episodes.⁴³ These revisions are discussed in further detail below.

More explicit modelling of inflation expectations (Canada and Greece)

In the original estimation, inflation expectations in the Phillips curve for most countries are proxied by a distributed lag of past inflation rates. However, this assumption may be particularly inappropriate and lead to biased estimates of the NAIRU following a change in policy regime. Canada and Greece are two countries where allowing for such a regime change seemed appropriate and leads to significant changes in the estimated NAIRU.

Canada was one of the first countries to introduce explicit inflation targeting in 1991. Empirical evidence from the Bank of Canada suggests that this has significantly influenced inflation expectations and following this evidence, inflation expectations from 1991 onwards are modelled as a weighted average of the (mid-point of the) inflation target and a distributed lag of past inflation rates (with weights of about half on each component).⁴⁴ The inflation variable used in the Phillips curve is the core measure of CPI inflation (excluding the effects of food, energy and indirect taxes) that the Bank focuses on for the purposes of monetary policy (although formally the inflation target is formulated in terms of the headline CPI). Given that inflation has consistently undershot the (mid-point of the) inflation target, the new policy regime may have provided an anchor for inflation expectations that has prevented further disinflation. Thus, not taking into account the effect of the change in policy regime on expectations is likely to have led to the NAIRU being over-estimated over recent years. Indeed, allowing for the change in policy regime lowers the NAIRU estimate on average by 0.3 percentage points over the period since the target has been in operation and by slightly more at the end of the estimation period.⁴⁵

Over the course of the 1990s, consumer price inflation in Greece has fallen from 20 to 2½ per cent per annum. One factor underlying this fall, at least over the past several years, may have been the effect that prospective membership of the EMU has had on lowering inflation expectations. To allow for this effect in the estimation of the NAIRU, inflation expectations from 1991 onwards are specified as a weighted average of past inflation and average euro area inflation, where the weight is estimated but allowed to increase at a linear rate over time.⁴⁶ Allowing for this regime shift implies a systematically higher NAIRU (because some of the disinflation is attributed to an expectations effect rather than the unemployment gap), that is on average nearly a percentage point higher than implied by the standard Phillips curve specification.

Allowing for the impact of recent reforms (Australia France and Switzerland)

As mentioned previously, a practical limitation of the estimation method concerns the greater uncertainty at the end of the sample period and, in particular, with respect to the effects of recent and on-going reforms. For those countries where such reforms took place in the late-1980s to mid-1990s (for example: the Netherlands, New Zealand, Spain and the United Kingdom), their impact on the NAIRU are typically found to be substantial but relatively slow to emerge.⁴⁷ To the extent that a number of other OECD countries are currently undergoing similar reforms, it may be too soon to see any appreciable reduction in the NAIRU reflected in current econometric estimates. In such cases, further adjustments are, therefore, made on the basis of the scale and nature of these recent reforms.⁴⁸

In Australia there have been significant reforms to both product and labour market institutions since 1996, including changes to the coverage of industrial awards, a move towards more decentralised bargaining and ongoing deregulation and privatisation of utilities. To incorporate the effect of these changes, the NAIRU was progressively revised downwards from 1998 to 6¾ per cent in 1999 (compared with a preliminary estimate of 7¼ per cent).

For France the preliminary econometric estimates suggested that the NAIRU had been broadly stable over the 1990s (at just over 10 per cent), although the standard error surrounding the estimate is among the largest of any country. Such a profile is not easily reconciled with the structural reforms that have been implemented since 1995, in particular large cuts in social security contributions, as well as evidence that the labour market has become more flexible with a growing share of temporary and part-time employment. To reflect these reforms the NAIRU is progressively revised downward from 1995, so that by 1999 it has fallen to 9½ per cent.

Switzerland has recently undergone a major reform of the unemployment insurance system that involved a tightening of unemployment benefit eligibility criteria in 1996 and 1997, with more intensive use of active labour market policies in 1998 and with participation becoming a condition of unemployment benefit eligibility.⁴⁹ The preliminary econometric estimates of the NAIRU were adjusted to reflect these reforms; a fall of ¾ per cent is imposed from 1997 to give an estimate of the NAIRU of 2½ per cent in 1999.

Special cases (Finland and Ireland)

In two special cases (Finland and Ireland) the specific estimation framework is considered inadequate for explaining past and recent experiences.

Finland has been affected by a number of major shocks in the early 1990s: the bursting of an asset price bubble, a sharp terms-of-trade fall and the collapse

of trade with the former Soviet Union. To reflect the impact of these shocks the profile of the estimated NAIRU has been judgementally adjusted in order to give a profile with a more pronounced rise in the early 1990s, that falls in the second half of the 1990s (consistent with supply side improvements in taxes, replacement rates and employment protection legislation) to a level of about 9 per cent in 1999.

The case of Ireland is unusual given the importance of immigration flows, which may mean that the NAIRU is more volatile than for most other countries with a greater tendency to follow the actual unemployment rate. Attempts to allow for this in the estimation process were, however, unsuccessful. Instead the econometric estimate was progressively revised downwards from 1995 to be more in line with the sharp fall in actual unemployment, so that by 1999 it had fallen to 7 per cent (compared with an econometric estimate of 9 per cent).

Recent trends in the NAIRU estimates

Combining the above judgmental adjustments with the econometric estimates gives a final set of NAIRU estimates for OECD countries reported in Table 2 and Figure 1.

Overall these estimates suggest that the extent and direction of changes in the NAIRU over the 1990s is distinctly mixed across OECD countries, although this might be favourably contrasted with the 1980s during which the NAIRU rose across virtually all of them (the United States and Portugal being exceptions). Countries where the NAIRU has risen by about 2 percentage points or more during the 1990s include Finland, Germany, Japan and Sweden, while Italy and Greece experienced a rise of just over 1 percentage point. Conversely, countries where the NAIRU has fallen by about a percentage point or more – Canada, Netherlands, New Zealand, the United Kingdom, Spain, Portugal, Ireland and Norway – include those where labour market reforms have been most extensive.⁵⁰ Nevertheless, the experience of these countries suggests that even following major reforms the NAIRU may only fall gradually (typically less than ½ percentage point per year) and with considerable lags. A striking exception is Ireland for which the NAIRU appears to have fallen by a remarkable 7 percentage points over the past decade.

There does appear to be a more uniform improvement in labour market performance across many countries in the second half of the 1990s compared with the first half, with two-thirds of the countries examined having experienced some fall in the NAIRU over the past five years. For example, Denmark, Finland, France, New Zealand and Norway have all had substantial falls in the NAIRU (of at least a percentage point) over the second half despite it rising earlier in the decade. Moreover, there are other countries (Canada, Ireland and Spain) for which the NAIRU has fallen more steeply in the second half of the 1990s. A major exception

is Japan where the NAIRU has risen more steeply, by over a percentage point, in the second half of the decade. Overall, while there do seem to be signs of recent progress, there remains considerable scope for further improvement: a weighted average of the NAIRUs across all the countries examined (which cover about 82 per cent of the total OECD labour force) suggests structural unemployment in the OECD is significantly higher now than in 1980 (let alone in earlier decades). Moreover, while disparities have narrowed marginally, large differences across countries remain.

At the same time, the revised estimates imply that for most OECD countries actual unemployment has been well in excess of the NAIRU for much of the 1990s, consistent with the substantial reduction in area-wide inflation. This is particularly the case for the euro area; the average gap between unemployment and the NAIRU since 1993 is about 1¾ percentage points (Figure 1). Much of this gap is accounted for by the three largest euro area economies, for which unemployment was still between 1 and 1½ percentage points higher than the estimated NAIRU in the second half of 1999, although the gap was narrowing. Conversely, for some of the smaller euro area countries the unemployment gap has just closed (Austria and Spain) or unemployment has been below the NAIRU (Ireland and Netherlands) for a year or more. On this basis recovery is even more advanced in both the United Kingdom and United States, where unemployment has been below the estimated NAIRU for 3 and 4 years, respectively. In order to reconcile inflation outcomes with these differing profiles of the gap between unemployment and the NAIRU, it is necessary to consider the short-run NAIRU.

THE RELEVANCE OF NAIRU ESTIMATES FOR MONETARY POLICY AND INFLATION

Indicators of structural unemployment provide a useful input to the setting of monetary policy if they help policymakers assess inflationary developments in the short term.⁵¹ In this respect, the short-term NAIRU concept may be a useful indicative synthesis of information concerning current inflationary pressures – see Estrella and Mishkin (1998) and King (1999) – even though its inherent volatility means that it is unsuitable as a target. Indeed, fluctuations in the short-run NAIRU provide an indication of which inflationary shocks policy-makers can ignore. For example, the effect of adverse temporary supply shocks that may dissipate in the near future should not be seen as necessitating a permanent rise in unemployment. In this situation, policy-makers need to assess, before taking action, whether or not inflation is likely to be consistent with policy objectives when the shock wears off.⁵²

The importance of the distinction between the NAIRU and short-run NAIRU is illustrated in Figure 1, which shows estimates for the G7 and euro-area economies: periods when unemployment is higher (lower) than the short-run NAIRU generally signal periods of falling (rising) inflation, even though the short-run NAIRU gap is sometimes of the opposite sign to that of the NAIRU gap. For the United States, the top left-hand panel of Figure 1 shows the unemployment rate over the period 1996 to 1998 was consistently above the short-run NAIRU, a period during which inflation fell, even though the unemployment rate was below the NAIRU.

Since 1996 unemployment has tended to exceed both the NAIRU and the short run NAIRU for the three largest euro-area economies, implying that demand pressures have been an important influence behind the fall in inflation, at least until the end of 1998. Over the same period, favourable movements in the short-run NAIRU in the United Kingdom and United States relative to euro-area economies are explained by the relative strength of exchange rates and their effects on imported inflation. However, since 1999 the rise in oil prices has become a major factor explaining the upturn in inflation and the corresponding increases in the short-run NAIRU across most OECD countries.

For Japan the rise in inflation during 1996 and 1997 can be related to unemployment falling below the NAIRU combined with pressure from import prices following depreciation of the yen. However, since 1997 the relatively rapid rise in unemployment, to levels in excess of the rising NAIRU has played an important role in driving inflation down to negative rates. Indeed, the relatively large unemployment gap coupled with the strengthening of the yen led to a further fall in inflation in 1999, despite the sharp rise in oil prices.

If speed limit effects are strong then the short-run NAIRU will show a tendency to track the actual unemployment rate because pronounced changes in unemployment will generate considerable changes in inflation in the short-run. In these circumstances, a rapid closing of a positive gap between actual unemployment and the NAIRU may generate unacceptable short-term inflationary effects. Among the G7 economies, such effects are found to be particularly important for Italy and the United Kingdom as reflected in the path of the short-run NAIRU estimates, which for these countries tend to fluctuate around the actual unemployment rate rather than around the NAIRU (Figure 1). Thus, for both countries there have been prolonged periods during the 1980s and 1990s when the actual unemployment rate has exceeded the NAIRU, but the profile of the short-run NAIRU suggests that the scope for reducing unemployment without (temporarily) increasing inflation was limited. Recently, speed limit effects may have been particularly important for Italy in 1999 and the United Kingdom during 1996-97; in both cases the inflationary effect of a relatively rapid fall in unemployment may have outweighed the deflationary effect of unemployment remaining in excess of the NAIRU. Such speed limits may be less pronounced in other countries, but never-

theless have represented a constraint in reducing unemployment quickly, even while it has remained well in excess of the NAIRU during most of the 1990s.

Finally, the limitations of any analysis based on the NAIRU and short-run NAIRU should be emphasised, particularly that they depend on estimated econometric relationships that explain inflation developments imperfectly, and are sometimes subject to large margins of error. As illustrated in Table 2 and Figure 2, standard errors surrounding the NAIRU estimates are on average about $\frac{3}{4}$ of a percentage point across all countries, but rise above 1 percentage point at the end of the estimation period.⁵³ Moreover, different specification choices may lead to different policy conclusions. For example, the choice of temporary supply shocks, oil and import price inflation, in the current analysis is based on what variables explain inflation consistently well across most OECD countries, but other choices are possible.⁵⁴ These factors all suggest that the NAIRU and short-term NAIRU can only serve as one of a range of possible indicators that are useful for assessing inflationary pressures.

Appendix

THE THEORETICAL FRAMEWORK

In previous OECD work on labour market issues (in particular, see OECD, 1996 and OECD, 1999) a framework based on the system of wage and price setting equations popularised by Layard *et al.* (1991), has been used extensively to illustrate how institutional characteristics and macroeconomic shocks interact and affect labour market performance, in particular the unemployment rate. Using this framework, this appendix reviews the theoretical underpinnings of the NAIRU concepts, showing the Phillips curve to be generally consistent with this theoretical model; one that can be interpreted as a reduced form relationship derived from the interaction of wage and price setting.

The structural model

The model used assumes an economy where wages are bargained between workers and firms – the latter deciding on the level of employment, output and prices once a wage agreement has been reached (the so-called “right-to-manage” model, see Layard *et al.*, 1991; Bean, 1994). Firms are assumed to operate in markets with imperfect competition, facing exogenously determined product market conditions, capital stocks and technology. Ignoring, for simplicity, labour force growth, this simple model can be summarised using three equations: 1) price-setting; 2) wage-setting; and 3) labour supply.

Price-setting

The price equation summarises the aggregate demand for labour by firms as a function of the (decreasing) marginal product of labour. If the product market is characterised by imperfections, the equation establishes a relationship between the optimal choice of employment and real wages for the firm, where prices are fixed as a margin over labour costs:

$$p - w = a_0 + a_1 n + a_2 \Delta n - a_3 (p - p^e) - q + ZL_p + ZT_p \quad a_1, a_2, a_3 > 0 \quad (1)$$

where Δ is the first difference operator,⁵⁵ n , w and p are respectively the logarithms of employment, wages (including payroll taxes) and prices, q is the logarithm of trend labour efficiency, ZL_p is a vector of variables having a “long-lasting” influence on price formation of firms, such as factors affecting the competitive structure of the market or the cost of capital. The ZT_p vector represents temporary factors affecting the price setting process (*i.e.* ZT_p represents supply shocks with zero ex-ante expectation) such as import or oil price shocks, p^e is the logarithm of expected prices.

Wage-setting

The wage equation can be obtained from different microeconomic models. Real wages are assumed to be a decreasing function of the unemployment rate (level and changes)⁵⁶ and

an increasing function of wage push factors (ZT_w and ZL_w) and labour efficiency, allowing for unanticipated wage changes ($w - w^e$).⁵⁷ Thus:

$$w - p = b_0 + b_1 U - b_2 \Delta U - b_3 (w - w^e) + q + ZT_w + ZL_w \quad b_1, b_2, b_3 > 0 \quad (2')$$

The ZL_w vector includes variables having long-lasting or “permanent” effects on the wage bargaining. This includes unemployment income support measures, indicators representative of the relative bargaining strength of unions and other relevant characteristics of the wage bargaining process as well as the degree of mismatch between skills and geographical location of job seekers and unfilled job vacancies. It might also take into account other supply factors such as changes in trend productivity growth or taxes as employees might be able to resist to downward adjustment in their after-tax real wage compensation. The ZT_w vector represents temporary factors affecting the wage bargaining process (*i.e.* ZT_w represents supply shocks with zero ex-ante expectation) like terms of trade effects. Thus, the specification of the wage setting equation encompasses various theoretical models, including those focusing on the matching process, efficiency wages and wage bargaining.

Labour-supply

Labour supply is assumed, for simplicity, to be inelastic with respect to real wages. It is a function of the unemployment rate (discouragement effect) and other factors affecting participation decisions (ZL_l), including some of the elements of the wage push (ZL_w).⁵⁸

$$l = c_0 - c_1 U + ZL_l \quad c_1 > 0 \quad (3)$$

where l is the logarithm of the labour force.

The different concepts of NAIRUs and the Philips curve equation

The *long-term equilibrium unemployment rate*, UL^* , is the solution to equations (1), (2), and (3), when price and wage expectations are met (*i.e.* $(w - w^e) = (p - p^e) = 0$), the unemployment rate is stabilised ($\Delta U = 0$), there are no temporary supply shock ($ZT_w = 0$ and $ZT_p = 0$) and long-lasting supply factors have adjusted fully to their long-term equilibria ($ZL_w = z_l w$, $ZL_p = z_l p$ and $ZL_l = z_l l$):

$$UL^* = \frac{d_0 + a_1 z_l l + z_l p + z_l w}{d_1} \quad (4)$$

where $d_0, d_1 > 0$ are functions of a_s, b_s and c_s parameters. This long-term equilibrium unemployment rate, which is fundamentally of the “natural rate” type (as stressed by Layard *et al.*, 1991), corresponds to the *long-term equilibrium concept* discussed in the main text. Its dependence on $z_l l, z_l p$ and $z_l w$ as well as the d_0, d_1 and a_1 parameters implies it is affected by the main institutional characteristics of the labour and product markets.

When the long-lasting supply factors are at their current values rather than their long-term equilibrium value following the response of the economy to macroeconomic shocks, one can define *the NAIRU concept (with no qualifying adjective)*, U^* , mentioned in the main text:

$$U^* = \frac{d_0 + a_1 ZL_l + ZL_p + ZL_w}{d_1} \quad (5)$$

The difference between the long-run NAIRU, UL^* , and the NAIRU, U^* , is that the former is associated with a particular realisation of the long-lasting supply shock variables ($ZL = z_l$) which corresponds to the long-run steady-state of the supply shocks.

The Phillips curve, related to this NAIRU concept U^* , can be obtained from equations (1), (2), (3) and (5) as a reduced form relationship, under the assumption of equal wage and price surprise (i.e. $(w - w^e) = (p - p^e)$).

$$\Delta p - (\Delta p)^e = \Delta w - (\Delta w)^e = -\frac{d_1}{d_3}(U - U^*) - \frac{d_2}{d_3}\Delta U + \frac{1}{d_3}(ZT_w + ZT_p + a_2\Delta ZL_l) \quad (6)$$

Defining $\pi = \Delta p$ inflation and $\pi^e = \Delta p^e$ inflation expectation and assuming expectations are adaptive and dependent on past inflation performance:

$$\Delta p - (\Delta p)^e = \pi - \pi^e = \Delta\pi - \Delta\pi^e = \Delta\pi - \alpha(L)\Delta\pi_{-1} \quad (7)$$

where $\alpha(L)$ is a polynomial of the lag operator. Using equation [6] and [7] we obtain then:

$$\Delta\pi = \alpha(L)\Delta\pi_{-1} - \frac{d_1}{d_3}(U - U^*) - \frac{d_2}{d_3}\Delta U + \frac{1}{d_3}(ZT_w + ZT_p + a_2\Delta ZL_l) \quad (8)$$

This *Phillips curve* is the equation referred to in the conceptual section of the main text. It is also the equation used to estimate the NAIRU (U^*) in the empirical analysis.⁵⁹

Equation (8) can also be used to define the concept of *short-term* NAIRU, US^* , corresponding to the value of unemployment which stabilises inflation over two consecutive period. Solving for $\Delta\pi = 0$:

$$US^* = \theta U^* + (1 - \theta)U - 1 + \frac{d_3}{(d_1 + d_2)}\alpha(L)\Delta\pi_{-1} + \frac{1}{(d_1 + d_2)}(ZT_w + ZT_p + a_2\Delta ZL_l), \quad (9)$$

where $\theta = d_1 / (d_1 + d_2)$

So, the short-term NAIRU can be expressed as a weighted average of actual (lagged) unemployment, the NAIRU, temporary supply shocks and lags of inflation. Similarly to Estrella and Mishkin (1998), equation (8) can hence be rewritten to relate inflation changes directly and only to the unemployment gap measured relative to the short-term NAIRU concept:

$$\Delta\pi = -\frac{(d_1 + d_2)}{d_3}(U - US^*) \quad (10)$$

NOTES

1. As is well known, the acronym is a misnomer, the concept is correctly defined as a “non-increasing” inflation rate of unemployment.
2. Previous Secretariat estimates also relate to wage inflation and the NAWRU, as described in Elmeskov (1993) and elaborated in OECD (1999), as opposed to price inflation and the NAIRU.
3. Friedman (1968) and Phelps (1968) are jointly credited with introducing the concept of the structural or natural rate, whilst the term NAIRU was first introduced by Modigliani and Papademos (1975).
4. This “orthodox” view contrasts with the alternative of “full hysteresis”, whereby the level of unemployment exerts no influence on inflation, although inflation is affected by the rate of change in unemployment. In this extreme case, unemployment is not anchored by structural variables, but will instead reflect the cumulative effect of all past shocks to the economy, including those to demand. A further implication is that unemployment can be maintained indefinitely at any level with stable inflation; which undermines the NAIRU concept. However, there is considerable empirical evidence against the hysteresis model in this extreme form; in particular, a substantial number of empirical studies suggest that the level of unemployment does have an effect on inflation, see for example the recent survey by Nickell (1998).
5. Friedman and Phelps explain the natural-rate model in terms of nominal wage rigidities in the labour market. For unemployment to remain below the natural rate, workers must be surprised by higher-than-expected price inflation into working for lower real wages *ex post* that they anticipated *ex ante*, when nominal wages were set. A similar Phillips curve can be derived from models in which nominal rigidities originate from the product market rather than the labour market if, for example, firms face costs in adjusting prices, as in the models of Calvo (1983) and Rotemberg (1982). More generally, a Phillips curve type relationship emerges as the reduced form of a variety of structural models (Roberts, 1997).
6. The latter may include, potentially, a fairly wide range of influences affecting pricing policies (changes in mark-ups, input, prices, etc.), the transformation and distribution process (competition, regulation, price controls, etc.), and wage determination (tax wedges, unionisation, income policies, etc.).
7. It is also possible that factors which permanently change the level of the wedge between the real product wage and the real consumption wage may also affect the NAIRU.
8. For a discussion of the sensitivity of NAIRU estimates coming from the structural approach to the precise way in which such models are formulated and estimated see Croub (1993).

9. Relevant studies include Nickell and Layard (1998), Elmeskov *et al.* (1998), OECD (1999), Nicoletti *et al.* (1999), Di Tella and MacCulloch (1998), and Daveri and Tabellini (1997).
10. Reviewing the empirical literature, Blanchard and Katz (1997) conclude that, "Economists are a long way from having a good quantitative understanding of the determinants of the natural rate, either across time or across countries."; while Nickell (1998) asserts, "What we lack is a satisfactory empirical explanation of the time series pattern of OECD unemployment".
11. These methods have most commonly been developed to measure potential output. See, for example, the methods developed by Watson (1986); Beveridge and Nelson (1981); Hodrick and Prescott (1997). Other approaches include the band pass filter, which gives results that are similar to an HP filter, Baxter and King, (1995); Christiano and Fitzgerald, (1999); the running median filter and the wavelet filter, Scacciavillani and Swagel, (1999).
12. In the case of Watson (1986), it is assumed that the trend and cyclical components are uncorrelated, while they are supposed to be perfectly correlated with the Beveridge Nelson filter. This latter assumption is economically not plausible.
13. If the arbitrary parameters are "tuned" to ensure that the resulting trend unemployment is sufficiently smooth and gives a reasonable proxy to the NAIRU, then the results, however, might be useful for forecasting inflation. The Bank of England (1999), for example, finds that a simple HP filter of unemployment works reasonably well in this respect.
14. Early attempts to take possible changes in the NAIRU into account involve allowing for different means of the unemployment rate across the sample, or different growth rates (Staiger *et al.*, 1997a, Gordon, 1997, Fabiani and Mestre, 1999). However, these appear unsatisfactory since it is difficult, in this way, to predict the next break in the NAIRU.
15. Confidence intervals for the NAIRU can be derived, although only a few papers do so. Staiger *et al.* (1997a) compute standard errors for the United States that range between 0.7 and 1.2. Irac (1999) uses a Monte Carlo approach to provide standard errors for the French NAIRU that lie between 0.8 and 1.7 depending on the sample period.
16. For example, in Bank of England (1999) it is applied to the United Kingdom, Gruen *et al.* (1999) to Australia, Irac (1999) to France, Meyler (1999) to Ireland, Apel and Jansson (1998, 1999) to Sweden, Rasi and Viikari (1998) to Finland, Orlani and Pichelman (2000) for the European Union and Fabiani and Mestre (1999) to the Euro area. There are fewer studies where the approach is applied consistently across a number of countries, although Laxton *et al.* (1998b) and Laubach (1999) both apply it to all the G7 countries.
17. The HPMV filter has been used to estimate the NAIRU by the Bank of Canada in the QPM model, and by OECD (1999) for a number of Member countries. Côté and Hostland (1994) also report use of a hybrid method combining an HP filter and the HPMV approach to provide estimates for Canada.
18. The approach can also be extended to encompass more complex models. Examples include adding equations that explain other price variables (CPI, PPI, wages) but using the same unemployment gap in each equation. To our knowledge, this has not been done yet, but examples of similar work to estimate core inflation can be found in Cecchetti (1997) and Le Bihan and Sedillot (1999). Preliminary OECD work along these lines is reported in Richardson *et al.* (2000).

19. This might require expectations of inflation to be formally introduced in the model (Roberts, 1996, 1997). The Bank of England (1999) shows that the speed of adjustment of unemployment for the United Kingdom appears quicker when one takes expectations into accounts. However, Meyer (1999) gets a worsening of the Phillips curve for Ireland when explicitly accounting for inflation expectations. For the most part, the empirical work reported here does not attempt to explicitly allow for forward-looking expectations because of the difficulties of doing so consistently across all countries. Exceptions are the estimated Phillips curves for Canada and Greece, as discussed later, which attempt to capture the effects of changes in policy regimes on expectations.
20. For example, Debelle and Laxton (1996) find it necessary to incorporate the idea that bottlenecks may start to develop as the unemployment rate falls below the NAIRU. This, in turn, means that further increases in demand will have even larger inflationary consequences. They show that for the United Kingdom, United States and Canada, such an asymmetric Phillips curve fits the data better and gives more sensible estimates of the unemployment gap.
21. In practice, two filtering methods – the Kalman filter and the Hodrick-Prescott Multi-Variate filter – were used in preliminary estimation but, as reported by Richardson *et al.* (2000), the Kalman filter procedure was preferred in being less restrictive and subject to fewer biases with respect to end-point sensitivity and cyclicity of the NAIRU estimates.
22. For small open economies, the GDP deflator might be a better indicator of inflationary pressures because it excludes the direct effects of terms of trade variations.
23. See, for example, Gordon (1997 and 1998), Hogan (1998), Irac (1999) and Meyer (1999).
24. For most countries, separate time trends were used for the full sample and from the beginning in 1980.
25. This was the case for Belgium, Spain, Finland, Australia, and Sweden.
26. For Belgium, as for the OECD *Jobs Strategy*, OECD (1999), a standardised unemployment role was used.
27. See Gruen *et al.* 1995, and comparable studies by the Australian Treasury and Reserve Bank.
28. Formally, equation (2b) is specified in state space form as two transition equations:
29. $(2b^I) U^*_t = (1 + \phi) U^*_{t-1} - \phi g_{t-1} + v^2_t$, and $(2b^{II}) g_t = U^*_{t-1}$
 Alternative specifications of the transition equation that were investigated, but without success across more than a few countries, were a random walk with stochastic drift and a random walk with deterministic drift.
29. This result is common to many such applications. Stock (1999) and Stock and Watson (1999) showed that when the true variances of (non-stationary) unobserved variables are small, the maximum likelihood estimates of the variances generally tend towards zero. Effectively, the estimation procedure gets trapped at a corner solution involving no fluctuations in the unobservable variable. This is why most of the literature tends to fix the value of the variance of the unobserved variable, or alternatively the signal-to-noise ratio. Another solution (Apel and Jansson, 1999) is to extend the model to provide more information on the evolution of the state variables (for example, adding an extra measurement equation specified in terms of an Okun rule), although such attempts sometimes give unsatisfactory results.

30. The estimation strategy was also guided by the use of comparable estimates based on an alternative HPMV filter. To achieve this, the procedure followed was to first fix the relative variances for the HPMV filter according to certain "rules of thumb". Then having obtained satisfactory HPMV estimates, the variance of the error term in the transition equations for the Kalman filter was chosen to give approximately the same degree of smoothness, where smoothness was judged according to the variance of the change of the NAIRU. The first difference, rather than the level of the NAIRU, was chosen because in many countries there is a clear upwards trend in the NAIRU.
31. The countries for which the direct maximum likelihood procedure did not work are Japan, Italy, Denmark, Norway and Sweden.
32. Experimentation suggests that the best choice for the initial Kalman filter iteration is to impose a coefficient on the unemployment gap which is half that obtained from the HPMV estimation (this follows the observation that the coefficient on the unemployment gap in the Phillips curve from using the Kalman filter in maximum likelihood estimation is typically much lower than that obtained from the HPMV filter). Using this choice the iterative procedure typically converges quickly on the full maximum likelihood procedure for those countries where the latter estimates were available.
33. There are a few counties where specification differences have arisen (mainly concerning whether unemployment and the NAIRU are specified in linear or log terms) because they lead to a significant improvement in the diagnostic tests.
34. Following the seminal study by Perry (1970) it has become common place for empirical studies of the NAIRU in the United States to use a demographically adjusted unemployment rate. These alternative unemployment rates are constructed as a fixed-weighted average of unemployment rates for various demographic categories, where the weights are the labour force shares of each group in some reference year (see Katz and Krueger, 1999 for a recent example). A demographic adjustment to the unemployment rate can be calculated as the difference between the actual unemployment rate and a demographically adjusted unemployment rate. For the present study an initial estimate of the NAIRU was derived by using the demographically adjusted unemployment rate in the Phillips curve. However, the NAIRU shown in the tables and charts is directly comparable to the published aggregate unemployment rate data; it is the sum of the NAIRU from the Phillips curve and the demographic adjustment, described above.
35. It was possible to estimate an auto-regressive NAIRU for Italy, but the random walk specification was preferred because of its superior performance in explaining inflation in the Phillips curve.
36. For simplicity as well as comparability across countries, this procedure was followed even in those cases where the Kalman filter NAIRU was estimated by maximum likelihood methods.
37. The inclusion of this additional term follows Fabiani *et al.* (1997) and is intended to capture the differential inflationary effect of changes in unemployment in the Centre-North region, compared to changes in unemployment in other regions.
38. The direct (rather than iterative) maximum likelihood estimation method was used for 16 of the 21 countries for which a NAIRU were estimated when using the Kalman filter.
39. See Boone (2000) and Richardson *et al.* (2000) for further details of the derivation of estimated standard errors for the NAIRU estimates.

40. Irac (1999) reports standard errors of 0.7 to 1.2 for France, Laubach (1999) reports values between 0.6 and 2.0 for the G7 countries.
41. In the case of Spain this involved using an inflation rate based on core CPI rather than the consumers' expenditure deflator. For Denmark a standardised rate of unemployment was used in place of a register-based definition, because the latter might not be a consistent basis for estimating the NAIRU given recent policy reforms which have eliminated a number of those on the rolls who would not fit within the standardised unemployment definition. In the case of Germany a distinct break in the NAIRU series was introduced to allow for the effect of re-unification (although this change had virtually no effect on the estimated NAIRU at the end of the estimation period).
42. See Table 1 and Figure 2.
43. Finland and Ireland are also the two countries with the largest standard errors surrounding the Kalman filter NAIRU estimates.
44. See for example Perrier (1998) and J-F. Fillion (1997).
45. For Canada the econometric NAIRU estimate was also revised down 0.1 percentage points in 1999 to reflect the effect of recent reforms to the unemployment insurance system.
46. By the end of the sample the weights on lagged and euro area inflation are around 85 and 15 per cent, respectively.
47. The fall in NAIRU estimates for these countries since implementing labour market reforms has, on average, been up to ½ per cent per annum, typically over a period of four to five years.
48. For further details of the reforms, see the most recent OECD Survey relating to the country concerned.
49. For Switzerland the tighter eligibility criteria has implied a significant drop in register-based unemployment – an effect which the Kalman filter can pick up only gradually.
50. Previous analysis has found that there is a correlation between falling structural unemployment and the extent to which *Job Study* recommendations have been implemented, see OECD (1999).
51. Here and in the following paragraphs discussing the policy usefulness of the NAIRU, it should be noted that to avoid confusion the terms “NAIRU” (*i.e.* without qualifying adjective) and “short-run NAIRU” are used strictly according to the definitions of Box 1.
52. See King (1999) for a discussion of how the appreciation of sterling in 1996 and 1997 was assessed by the Bank of England's Monetary Policy Committee in broadly these terms.
53. See Richardson *et al.* (2000) for a description of the Monte Carlo methods used to calculate these standard errors.
54. For example, Brayton. *et al.* (1999) suggest that variations in the mark-up of prices over unit labour costs explain low inflation in the United States in recent years.
55. The first difference operator appears here as a result of lagged response in employment, often caused by the presence of adjustment costs on labour inputs (see, for example, Lindbeck and Snower, 1988).
56. The presence of ΔU in the wage-setting schedule can be justified by the behaviour of both firms and workers. On the basis of the “insider-outsider” hypothesis, it could be argued that real wages may be more responsive to the threat of large-scale redun-

dancy and rising unemployment than to the level of unemployment *per se*. Likewise, in the context of rising unemployment, the proportion of short-term unemployed (*i.e.* those most likely to compete directly with the employed) generally increases and this could put more downward pressure on wages than a stable level of unemployment (see, for example, Blanchard and Summers, 1987 and Layard *et al.* 1991).

57. Wage and price surprises appear in equations (1) and (2') in this form as a result of aggregation. They are derived from the absence of knowledge of aggregate values of those variables that are contemporaneously set at the microeconomic level by workers and firms (see, for instance, Layard *et al.* 1991). Other inertial effects (such as the staggering of wage contracts) can be allowed for in the same way without changing the qualitative properties of the model.
58. For the simplicity of exposition ZL_1 vector is supposed to incorporate only factors having a long-lasting or permanent influence on labour supply.
59. Note that prices and wages play a similar role in the derivation of the reduced-form Phillips curve equation, so that *a priori* this equation may be based on either variable.

BIBLIOGRAPHY

- APEL, M. and P. JANSSON (1998),
"System estimates of potential output and the NAIRU", *Empirical Economics*, Vol. 24,
pp. 373-388.
- APEL, M. and P. JANSSON (1999),
"A theory-consistent system approach for estimating potential output and the NAIRU",
Economics Letters, Vol. 64, pp. 271-275.
- BANK OF ENGLAND (1999),
"Economic models at the Bank of England", pp. 77-156.
- BAXTER, M. and R.G. KING (1995),
"Measuring business cycles approximate band-pass filters for economic time series",
National Bureau of Economic Research Working Papers, No. 5022.
- BEAN, C. (1994),
"European unemployment: a survey", *Journal of Economic Literature*, Vol. XXXII, pp. 573-619.
- BEVERIDGE, S. and C.R. NELSON (1981),
"A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'", *Journal of Monetary Economics*, Vol. 7, No. 2, pp. 151-174.
- BLANCHARD, O. and L.H. SUMMERS (1987),
"Fiscal increasing returns, hysteresis, real wages, and unemployment", *European Economic Review*, Vol. 31, No. 3, pp. 543-560.
- BLANCHARD, O. and L.F. KATZ (1997),
"What we know and do not know about the natural rate of unemployment", *NBER Working Papers*, No. 5822.
- BLANCHARD, O. and J. WOLFERS (1999),
"The role of shocks and institutions in the rise of European unemployment: the aggregate evidence", *NBER Working Papers*, No. 7282.
- BOONE, L. (2000),
"Comparing semi-structural methods to estimate unobserved variables", *OECD Economics Department Working Papers*, No. 240.
- BOONE, L., M. JUILLARD, D. LAXTON and P. N'DIAYE (2001),
"How well do alternative time-varying parameter models of the NAIRU help policymakers forecast unemployment and inflation in the OECD countries?", draft *IMF Working Paper*, February 2001.
- BROER, P., N. DRAPER and F. HUIZINGA (1998),
"The equilibrium rate of unemployment, an econometric analysis", *Centraal Planbureau, Interne Notitie*.

- BRAYTON, F., J.M. ROBERTS and J.C. WILLIAMS (1999),
 "What's happened to the Phillips curve?", *Federal Reserve Board Working Papers*, No. 1999-49 (September).
- CALVO, G. (1983),
 "Staggered prices in a utility maximizing framework", *Journal of Monetary Economics*, Vol. 12, No. 3, pp. 383-98.
- CECCHETTI, S.G. (1997),
 "Central Bank policy rules: conceptual issues and practical considerations", *NBER Working Papers*, No. 6306.
- CHRISTIANO, L. and T. FITZGERALD (1999),
 "The band pass filter", *NBER Working Papers*, No. 7257.
- COTE, D. and D. HOSTLAND, (1994),
 "Measuring potential output and the NAIRU as unobserved variables in a systems framework", Bank of Canada, ed. *Economic Behaviour and Policy Choice Under Price Stability*, pp. 357-411.
- COTIS, J.P., R. MEARY and N. SOBCZAK (1996),
 "Le chômage d'équilibre en France : une évaluation", Contribution française à la session du Groupe de travail No. 1 du Comité de politique économique de l'OCDE. Also published as *Direction de la Prévision Working Papers*, No. 96-14.
- CROMB, R. (1993),
 "A survey of recent econometric work on the NAIRU", *Journal of Economic Studies*, Vol. 20, No. 1/2 pp. 27-51.
- CUTHBERSON, K., S.G. HALL and M.P. TAYLOR (1992),
Applied Econometric Techniques, Harvester Wheatsheaf, U.K.
- DAVERI, F. AND G. TABELLINI (1997),
 "Unemployment, growth and taxation in industrial countries", *Innocenzo Gasparini Institute for Economic Research Working Papers*, No. 122.
- DEBELLE, G. and D. LAXTON (1996),
 "Is the Phillips curve really a curve? Some evidence for Canada, the United Kingdom and the United States", *IMF Working Papers*, No. 96/111.
- DI TELLA, R. and R. MACCULLOCH (1998),
 "The consequences of labour market flexibility: panel evidence based on survey data", Harvard Business School, mimeo.
- ELMESKOV, J. (1993),
 "High and persistent unemployment: assessment of the problem and its causes", *OECD Economics Department Working Papers*, No. 132.
- ENGLANDER, A. and C.A. LOS (1983),
 "Recovery without accelerating inflation?", *Federal Reserve Bank, Quarterly Review*, Vol. 8, No. 2, pp. 19-28.
- ESTRELLA, A. and F.S. MISHKIN (1998),
 "Rethinking the role of NAIRU in monetary policy: implications of model formulation and uncertainty", *NBER Working Papers*, No. 6518.
- FABIANI, S. and R. MESTRE (2000),
 "Alternative measures of the NAIRU in the euro area: estimates and assessment", *European Central Bank Working Papers*, No. 17.

- FILLION, J.-F. et A. LÉONARD (1997),
 "La courbe de Phillips au Canada : un examen de quelques hypothèses", *Bank of Canada Working Papers*, No. 97-3.
- FORTIN, P. (1989),
 "How natural is Canada's high unemployment rate?", *European Economic Review*, Vol. 58, pp. 1-17.
- FRIEDMAN, M. (1968),
 "The role of monetary policy", *American Economic Review Papers and Proceedings*, Vol. 58, No. 1, pp. 1-17.
- FUHRER, J.C. (1995),
 "The Phillips curve is alive and well", *New England Economic Review*, The Federal Reserve of Boston, March/April, pp. 41-56.
- GIORNO, C., P. RICHARDSON, D. ROSEVEARE and P. VAN DEN NOORD (1995),
 "Estimating potential output, output gaps and structural budget balances", *OECD Economics Department Working Papers*, No. 152, pp. 167-209.
- GORDON, R.J. (1997),
 "The time-varying NAIRU and its implications for economic policy", *Journal of Economic Perspectives*, 11(1), Winter 1997, pp. 11-32.
- GORDON, R.J. (1998),
 "Foundations of the Goldilocks economy: supply shocks and the time-varying NAIRU", *Brookings Papers on Economic Activity*, Vol. 2, pp. 297-346.
- GRUEN, D., A. PAGAN and C. THOMPSON (1999),
 "The Phillips Curve in Australia", *Journal of Monetary Economics*, Vol. 44, No. 2, pp. 223-258.
- GRUBB, D. (1986),
 "Topics in the OECD Phillips curve", *Economic Journal*, Vol. 96, pp. 55-79.
- GRUEN, D., A. PAGAN and C. THOMPSON (1999),
 "The Phillips Curve in Australia", *Journal of Monetary Economics*, Vol. 44, No. 2, pp. 223-258.
- HAMILTON, J.D. (1986),
 "A standard error for the estimated state vector of a state space model", *Journal of Econometrics*, Vol. 33, pp. 387-397.
- HAMILTON, J.D. (1994),
Time Series Analysis, Princeton, Princeton University Press.
- HARVEY A.C. (1985),
 "Trends and cycles in macroeconomic time series". *Journal of Business and Economic Statistics* 3: pp 216-27.
- HODRICK, R.J. and E.C. PRESCOTT (1997),
 "Post-war US business cycles: an empirical investigation", *Journal of Money, Credit and Banking*, Vol. 29, No. 1, pp. 1-16.
- HOGAN, S. (1997),
 "What does downward nominal-wage rigidity imply for monetary policy?", *Bank of Canada Working Paper*, No. 97-13.
- HOLDEN, S. and R. NYMOEN (1998),
 "Measuring structural unemployment: is there a rough and ready answer? ", OECD.

- IRAC, D. (1999),
 "Estimation of a time varying NAIRU for France", Banque de France, *ECB Workshop on Structural Unemployment*.
- KALMAN, R.E. (1960),
 "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering, Transactions ASME, Series 82*, pp. 35-45.
- KATZ, L.F. and A.B. KREUGER (1999),
 "The High Pressure US Labour Market of the 1990s", *Brookings Papers in Economic Activity*, 1:1999, pp. 1-87.
- KING, M. (1999),
 "Monetary policy and the labour market", Bank of England, *Quarterly Bulletin*, February.
- KING, R.J., J.H. STOCK and M.W. WATSON (1995),
 "Temporary instability of the unemployment relationship", *Economic Perspectives of the Federal Reserve Bank of Chicago*, May/June, pp. 2-12.
- LAXTON, D., D. ROSE and D. TONBAKIS (1998),
 "The US Phillips Curve: the case of asymmetry", *Journal of Economic Dynamics and Control*, No. 23, 1459-1486.
- LAXTON, D. and R. TETLOW (1992),
 "A simple multivariate filter for the measurement of potential output", *Technical Report*, No. 59, Bank of Canada.
- LAYARD, R., S. NICKELL and R. JACKMAN (1991),
Unemployment, Macroeconomic Performance and the Labour Market, Oxford Press.
- LAUBACH, T. (1999),
 "Measuring the NAIRU: evidence from seven economies", *Federal Reserve Bank of Kansas City Working Papers*, No. 39-13.
- LE BIHAN, H. and F. SEDILLOT (1999),
 "Une comparaison des différents indicateurs d'inflation sous-jacente", Banque de France.
- L'HORTY, Y. and C. RAULT (1999),
 "Les causes du chômage en France: une ré-estimation du modèle WS-PS", Conseil Supérieur de l'emploi, des revenus et des coûts, *Document de Travail*, No. 99-01.
- LINDBECK, A. and D.J. SNOWER (1998),
 "Price dynamics and production lags", Institute for International Economic Studies, University of Stockholm.
- MANNING, A. (1993),
 "Wage bargaining and the Phillips curve: the identification and specification of aggregate wage equations", *The Economic Journal*, Vol. 103, pp. 98-118.
- MEYER, A., (1999),
 "The NAIRU in a small economy: the Irish context", Bank of Ireland, *ECB Workshop on Structural Unemployment*.
- MODIGLIANI, F. and L. PAPADEMOS (1975),
 "Targets for monetary policy in the coming year", *Brookings Paper on Economic Activity*, Vol. 1, pp. 141-163.
- NICKELL, S. (1998),
 "Unemployment: questions and some answers", *The Economic Journal*, Vol. 108, pp. 802-816.

- NICKELL, S. and R. LAYARD (1998),
“Labour market institutions and economic performance”, The Centre of Economic Performance, London School of Economics.
- NICOLETTI, G., S. SCARPETTA and O. BAYLAUD (1999),
“Summary indicators of product market regulation with an extension to employment protection legislation”, *OECD Economics Department Working Papers*, No. 226.
- OECD (1994),
The OECD Jobs Study: Evidence and Explanations, Paris.
- OECD (1999),
Implementing the OECD Job Strategy – Assessing Performance and Policy, Paris.
- ORLAND I. and K. PICHELMAN (2000),
“Disentangling trend and cycle in EUR-11 unemployment series – an unobserved component modelling approach”, European Commission Directorate-General for Economic and Financial Affairs, *Economic Papers*, No. 140.
- PERRIER, P. (1998),
“Un examen de la crédibilité de la politique monétaire au Canada”, *Bank of Canada Working Papers*, No 98-12.
- PERRY, G. (1970),
“Changing labour markets and inflation”, *Brookings Papers in Economic Activity*, 3:1970, 411-48.
- PHELPS, E.S. (1994),
Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest and Assets, Cambridge, Harvard University Press.
- PHELPS, E.S. (1968),
“Money, wage dynamics and labour market equilibrium”, *Journal of Political Economy*, Vol. 76, No. 4 (part 2), pp. 678-711.
- RASI, C.M. and J.M. VIIKARI (1998),
“The time-varying NAIRU and potential output in Finland”, *Bank of Finland Working Papers*, No. XX/98.
- RICHARDSON, P., L. BOONE, C. GIORNO, M. MEACCI, D. RAE and D. TURNER (2000),
“The concept, policy use and measurement of structural unemployment: estimating a time varying NAIRU across 21 OECD countries”, *OECD Economics Department Working Papers*, No. 250.
- ROBERTS, J. (1997),
“The wage curve and the Phillips curve”, Board of Governors of the Federal Reserve System.
- ROTEMBERG, J.J. (1982),
“Sticky prices in the United States”, *Journal of Political Economy*, Vol. 90, No. 6, pp. 1187-1211.
- ROWTHORN, R. (1999),
“Unemployment, wage bargaining and capital-labour substitution”, *Cambridge Journal of Economics*, Vol. 23, pp. 413-425.
- SCACCIAVILLANI, F. and P. SWAGEL (1999),
“Measures of potential output: an application to Israel”, *IMF Working Papers*, No. 99/66.

- SETTERFIELD, M.A., D.V. GORDON and L. OSBERG (1992),
"Searching for a will o'the wisp: an empirical study of the NAIRU in Canada", *European Economic Review*, No. 36, pp. 119-136.
- STAIGER, D., J.H. STOCK and M.W. WATSON (1997a),
"How precise are estimates of the natural rate of unemployment?", in *Reducing Inflation: Motivation and Strategy*, C.D. Romer and D.H. Romer, eds., Chicago: University of Chicago Press.
- STOCK, J.H. (1999a),
"Monetary policy in a changing economy: indicators, rules , and the shift towards intangible output", *IMES Discussion Paper Series*, No. 99-E-13.
- STOCK, J.H. and M.W. WATSON (1999b),
"Forecasting inflation", *Journal of Monetary Economics*, Vol. 44, pp. 293-335.
- WATSON, M.W. (1986),
"Univariate detrending methods with stochastic trends", *Journal of Monetary Economics*, Vol. 18, No. 1, pp. 49-75.

OECD ECONOMIC STUDIES

Recent issues

No. 29, 1997/II

Early retirement in OECD countries:
the role of social security systems

Sveinbjörn Blöndal
and Stefano Scarpetta

Income distribution and poverty
in 13 OECD countries

Howard Oxley, Jean-Marc Burniaux,
Thai-Thanh Dang
and Marco Mira d'Ercole

Does government support stimulate
private R&D?

Dominique Guellec and
Bruno van Pottelsberghe de la Potterie

Causes of fluctuations in R&D
expenditures – a quantitative analysis

Dominique Guellec
and Evangelos Ioannidis

International comparisons of part-time
work

Georges Lemaitre, Pascal Marianna
and Alois van Bastelaer

(13 97 29 1 P) ISBN 92-64-15379-9, 152 pages

No. 30, 2000/I

Poverty dynamics in six OECD
countries

Howard Oxley, Thai-Thanh Dang
and Pablo Antolín

Determinants of health outcomes
in industrialised countries: a pooled,
cross-country, time-series analysis

Zeynep Or

What works among active labour
market policies: evidence from OECD
countries' experiences

John P. Martin

Estimating prudent budgetary margins
for EU countries:
a simulated SVAR model approach

Thomas Dalsgaard and Alain de Serres

Is the health of older persons in OECD
countries improving fast enough to
compensate for population ageing?

S. Jacobzone, Emmanuelle Cambois
and Jean-Marie Robine

Revisiting real social spending
across countries: a brief note

Willem Adema

(13 2000 30 1 P) ISBN 92-64-17556-3, 197 pages

No. 31, 2000/II

An overview: what do we know about policies to make work pay?

Mark Pearson and Stefano Scarpetta

Not perfect, but still pretty good: the EITC and other policies to support the US low-wage labour market

V. Joseph Hotz and John Karl Scholz

Experimental evidence on the use of earnings supplements as a strategy to “make work pay”

John Greenwood and Jean-Pierre Voyer

The family credit system and the working families tax credit in the United Kingdom

Andrew Dilnot and Julian McCrae

The importance of inclusion and the power of job subsidies to increase it

Edmund S. Phelps

Payroll tax reductions for the low paid

Jean-Paul Fitoussi

The use of wage floors as policy tools

Paul Gregg

Eligibility criteria for unemployment benefits

David Grubb

Policy choices and interactions with existing instruments

Giuseppe Bertola

(13 2000 31 1 P) ISBN 92-64-17557-1, 198 pages

No. 32, 2001/I

The implementation and the effects of regulatory reform: past experience and current issues

Rauf Gönenç, Maria Maher and Giuseppe Nicoletti

Regulation, market structure and performance in telecommunications

Olivier Boylaud and Giuseppe Nicoletti

Regulation, industry structure and performance in the electricity supply industry

Faye Steiner

Regulation, market structure and performance in air passenger transportation

Rauf Gönenç and Giuseppe Nicoletti

Regulatory reform in road freight

Olivier Boylaud and Giuseppe Nicoletti

Regulatory reform in retail distribution

Olivier Boylaud and Giuseppe Nicoletti

(13 2001 32 1 P) ISBN 92-64-29004-4, 274 pages

Individual issues are available via the OECD's online bookshop:

www.oecd.org/bookshop

OTHER ECONOMIC PERIODICALS

Economics Department, OECD

OECD Economic Outlook (ISSN 0474-5574)

Each June and December, the OECD *Economic Outlook* presents projections of output, employment, prices and current balances over the next two years, based on a review of each Member country and of the feedback effect of each of them on international developments.

OECD Economic Outlook Flashfile

A data file containing a summary of the Economic Outlook forecasts is now available on INTERNET at the time of its preliminary publication (six to eight weeks before the final publication date) at the following address: www.oecd.org/eco/data/dp.htm. This includes key macroeconomic variables for all OECD countries and regions in text file form, which can be input directly into most statistical and analytical software. The Economic Outlook Flashfile is available free of charge to existing subscribers to the OECD Economic Outlook data diskettes, but can be purchased separately for FF 500 per edition. Subscription can be registered at the above address or via our regular e-mail address: eco.contact@oecd.org.

Statistics and projections on microcomputer diskette

The full set of historical time series data and projections underlying the OECD *Economic Outlook* is regularly available on microcomputer diskettes at the same time as its publication. A basic version of the OECD *Economic Outlook* diskettes contains approximately 3 500 macroeconomic time series for OECD countries and non-OECD zones, beginning in 1980 and extending to the end of the published forecast horizon. In addition, the OECD *Economic Outlook Reference Supplement* contains data for most OECD countries and the majority of National Accounts and other country-specific statistics, starting as early as 1960. This supplement provides data in a form that is fully compatible with other OECD *Economic Outlook* files, permitting a wide variety of analyses and statistical applications, *e.g.* for modelling and longer-term comparative and graphical analyses.

The general subject and country coverage for both versions are as follows:

Subject coverage

- Gross national product and its components
- Government and households appropriation accounts
- Fiscal and monetary indicators
- Labour market and supply indicators
- Wages, prices and profitability
- International trade and payments

Country coverage

- 30 OECD countries
- OECD area aggregations
- Non-OECD zones
- The euro area

Subscription information

Annual subscription (including *Reference Supplement*) is for two issues per year – June and December. Subscriptions, which also include the printed version of the OECD *Economic Outlook*, may be made at any time of the year.

Diskette format

Data are provided on 3½-inch diskettes suitable for IBM PC-compatible microcomputers. These include a simple programme to view and to translate these series into a variety of formats that can be read and used by most PC software packages.

OECD ECONOMIC SURVEYS by the Economic and Development Review Committee.
Detailed surveys of trends and prospects for OECD Member countries (**ISSN 0376-6438**)

Australia	(August 2001)	Korea	(July 2001)
Austria	(May 1999)	Luxembourg	(February 2001)
Belgium	(March 2001)	Mexico	(July 2000)
Canada	(September 2001)	Netherlands	(March 2000)
Canada	(August 2000)	Poland	(January 2000)
Czech Republic	(July 2001)	New Zealand	(December 2000)
Denmark	(July 2000)	Norway	(February 2001)
Finland	(July 2000)	Poland	(May 2001)
France	(July 2000)	Portugal	(April 2001)
Germany	(May 2001)	Slovak Republic	(February 1999)
Greece	(February 2001)	Spain	(June 2001)
Hungary	(November 2000)	Sweden	(March 2001)
Iceland	(April 2001)	Switzerland	(December 2000)
Ireland	(June 2001)	Turkey	(February 2001)
Italy	(May 2000)	United Kingdom	(June 2000)
Japan	(December 2000)	United States	(May 2000)
Euro Area	(May 2001)	Romania	(February 1998)
Brazil	(June 2001)	Russian Federation	(March 2000)
Bulgaria	(April 1999)	Slovenia	(May 1997)
Baltic States	(February 2000)		

For more information, visit the OECD website at the following address:

<http://webnet1.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-home-nothème-8-no-no-no-no,00.html> and also the OECD's online bookshop: www.oecd.org/bookshop.

Send your order to your local distributor or to:

In the United States

OECD WASHINGTON CENTER
2001 I Street NW, Suite 650
Washington, DC 20036-4922
Telephone: (202) 785-6323
Toll-free number for orders: (800) 456 6323
Fax: (202) 785-0350
E-mail: washington.contact@oecd.org
Internet: www.oecdwash.org

In Latin America

OECD MEXICO CENTRE
Edificio INFOTEC
Av. San Fernando No. 37 Col. Toriello
Guerra Tlalpan C.P. 14050, Mexico D.F.
Telephone: (525) 528 10 38
Fax: (525) 606 13 07
E-mail: mexico.contact@oecd.org
Internet: www.rtn.net.mx/ocde

In Asia

OECD TOKYO CENTRE
Landic Akasaka Building
2-3-4 Akasaka, Minato-ku
Tokyo 107-0052, Japan
Tel: (81-3) 3586-2016
Fax: (81-3) 3584-7929
E-mail: center@oecdtokyo.org
Internet: www.oecdtokyo.org

In the rest of the world

OECD PARIS CENTRE
2, rue André-Pascal
75775 Paris Cedex 16
France
Tel: (33) 1 45 24 81 67
Fax: (33) 1 45 24 19 50
E-mail: sales@oecd.org
Internet: www.oecd.org
On-line ordering: www.oecd.org/publications
(secure payment with credit card)

WORKING PAPERS

Economics Department, OECD

This series of *Working Papers* is designed to make available to a wider readership selected studies that the Department has prepared for use within OECD. Authorship is generally collective, but main individual authors are named from *Working Paper* No. 26 onwards, papers are generally available on request in their original language, English or French, with a summary in the other. Earlier papers are available in both English and French. *Working Paper* No. 144 and subsequent papers are available through Internet at www.oecd.org/eco/eco.

Enquiries should be addressed to:

OECD
Economics Department
2, rue André-Pascal
75775 PARIS CEDEX 16
FRANCE
Fax No. (33) 1 44 30 63 80
E-Mail: eco.contact@oecd.org

- 311. *Modelling Import Responsiveness for OECD Manufactures Trade*
(October 2001) Mara Meacci and David Turner
- 310. *Trade Linkages and the Trade Matrices in the OECD Interlink Model*
(October 2001) Laurence Le Foulher, Wim Suyker and Dave Turner
- 309. *Encouraging Environmentally Sustainable Growth in Australia*
(October 2001) Ann Vourc'h and Robert Price
- 308. *Financial Market Liberalisation, Wealth and Consumption*
(September 2001) Laurence Boone, Nathalie Girouard and Isabelle Wanner
- 307. *The Economic Integration of Germany's New Länder*
(September 2001) Eckhard Wurzel
- 306. *Standard Shocks in the OECD Interlink Model*
(September 2001) Thomas Dalsgaard, Christophe André and Pete Richardson
- 305. *Fiscal Implications of Ageing: Projections of Age-related Spending*
(September 2001) Thai Thanh Dang, Pablo Antolin and Howard Oxley
- 304. *The Width of the Intra-European Economic Borders*
(August 2001) Alain de Serres, Peter Hoeller and Christine de la Maisonneuve

- 303. *Surveillance of Tax Policies: A Synthesis of Findings in Economic Surveys*
(July 2001) Paul van den Noord and Christopher Heady
- 302. *Reforming the Tax System in Portugal*
(July 2001) Chiara Bronchi, José C. Gomes-Santos
- 301. *Tax Systems in European Union Countries*
(June 2001) Isabelle Joumard
- 300. *Encouraging Environmentally Sustainable Growth in Belgium*
(June 2001) Paul O'Brien, David Carey, Jens Høj, Andreas Woergoetter
- 299. *Encouraging Environmentally Sustainable Growth in Poland*
(June 2001) Grzegorz Peszko, Patrick Lenain
- 298. *Tracking the Euro*
(June 2001) Vincent Koen, Laurence Boone, Alain de Serres, Nicola Fuchs
- 297. *Firm Dynamics and Productivity Growth: A Review of Micro-evidence from OECD Countries*
(June 2001) Sanghoon Ahn
- 296. *How should Norway Respond to Ageing?*
(May 2001) Pablo Antolín and Wim Suyker
- 295. *How will Ageing Affect Finland?*
(May 2001) Pablo Antolín, Howard Oxley and Wim Suyker
- 294. *Sectoral Regulatory Reforms in Italy: Framework and Implications*
(May 2001) Alessandro Goglio
- 293. *Encouraging Environmentally Sustainable Growth: Experience in OECD Countries*
(May 2001) Paul O'Brien and Ann Vourc'h
- 292. *Increasing Simplicity, Neutrality and Sustainability: A Basis for Tax Reform in Iceland*
(May 2001) Richard Herd and Thorsteinn Thorgeirsson
- 291. *Options for Reforming the Tax System in Greece*
(April 2001) Chiara Bronchi
- 290. *Encouraging Environmentally Sustainable Growth in Canada*
(March 2001) Ann Vourc'h
- 289. *Encouraging Environmentally Sustainable Growth in Sweden*
(March 2001) Deborah Roseveare
- 288. *Public Spending in Mexico: How to Enhance its Effectiveness*
(March 2001) Bénédicte Larre and Marcos Bonturi
- 287. *Regulation in Services: OECD Patterns and Economic Implications*
(February 2001) Giuseppe Nicoletti
- 286. *A Small Global Forecasting Model*
(February 2001) David Rae and David Turner
- 285. *Managing Public Expenditure: Some Emerging Policy Issues and a Framework for Analysis*
(February 2001) Paul Atkinson and Paul van den Noord

OECD STATISTICAL SERIES ON DISKETTE

International Trade and Competitiveness Indicators (ITCI)

The *International Trade and Competitiveness Indicators* statistical series are designed for use in a wide range of empirical analyses related to the international trade performance of OECD Member countries. The set of diskettes includes principal series for external trade classified by four broad SITC commodity groups (basic materials, food, fuels and manufactures) in nominal and constant-price value terms, along with corresponding price deflators, market shares and competitiveness indicators. These statistics are available in a consistent quarterly form for OECD countries. The length of individual series is determined by the availability of individual country source information, but in general all series are available from 1975 onwards.

The set of diskettes is updated to take account of new information and data revisions on a twice-yearly basis.

Business Sector Data Base (BSDB)

The *Business Sector Data Base* contains data for 26 OECD Member countries related to business-sector value added, employment, investment, factor prices and capital stocks, from 1960 to the present day, and are stored in quarterly form. The underlying database was developed for use in the analysis of production and supply issues in the context of the OECD's regular surveys and assessments, and related empirical studies.

The BSDB is available in two issues per year.

Fiscal Positions and Business Cycles (FPBC)

The *Fiscal Positions and Business Cycles* (FPBC) diskettes contain detailed quantitative information on potential output and output gaps, on actual and cyclically-adjusted government revenues, outlays and balances, and on government financial liabilities (including debt series consistent with the Maastricht criterion). Historical annual data from 1970 onwards and the projections of the OECD *Economic Outlook* are given for 20 OECD countries. The diskettes are a valuable data source for those interested in fiscal policy and in potential output and output gaps.

The set of diskettes is updated on a twice-yearly basis to take into account data revisions and to make available new projections.

For more information, visit the OECD website at the following address: <http://webnet1.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-home-notheme-8-no-no-no-no-no,00.html>. A full listing of OECD statistics is available at the following address: www.oecd.org/statlist.htm.

OECD ELECTRONIC BOOKS

Economics Department Publications in Electronic Form

Recognising the growing importance of multimedia electronic publications, the OECD is in the process of producing a new range of electronic books corresponding to many of its key economic publications. A special version of the OECD *Economic Outlook* (excluding the statistical annex) is now available and electronic book versions of the OECD and Centre for co-operation with the Economies in Transition (CCET) *Economic Surveys* will shortly be released based on the 1997 series. These typically provide the text, tables and figures of the relevant publication on screen and in an identical format to that of the printed version, including the use of colour in graphs.

The electronic book, which retains the quality and readability of the printed version throughout, will enable readers to take advantage of the new tools that the ACROBAT software (included on the diskette) provides by offering the following benefits:

- User-friendly and intuitive interface.
- Comprehensive index for rapid text retrieval, including a table of contents, as well as a list of numbered tables and figures.
- Rapid browse and search facilities.
- Zoom facility for magnifying graphics or for increasing page size for easy readability.
- Cut and paste capabilities.
- Printing facility.
- Reduced volume for easy filing/portability.

Working environment: DOS, Windows or Macintosh.

OECD *Economic Outlook* (June and December)

OECD (and CCET) *Economic Surveys*

For further information contact the OECD online bookshop: www.oecd.org/bookshop

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(13 2001 33 1 P) ISBN 92-64-19005-8 – No. 52269 2001