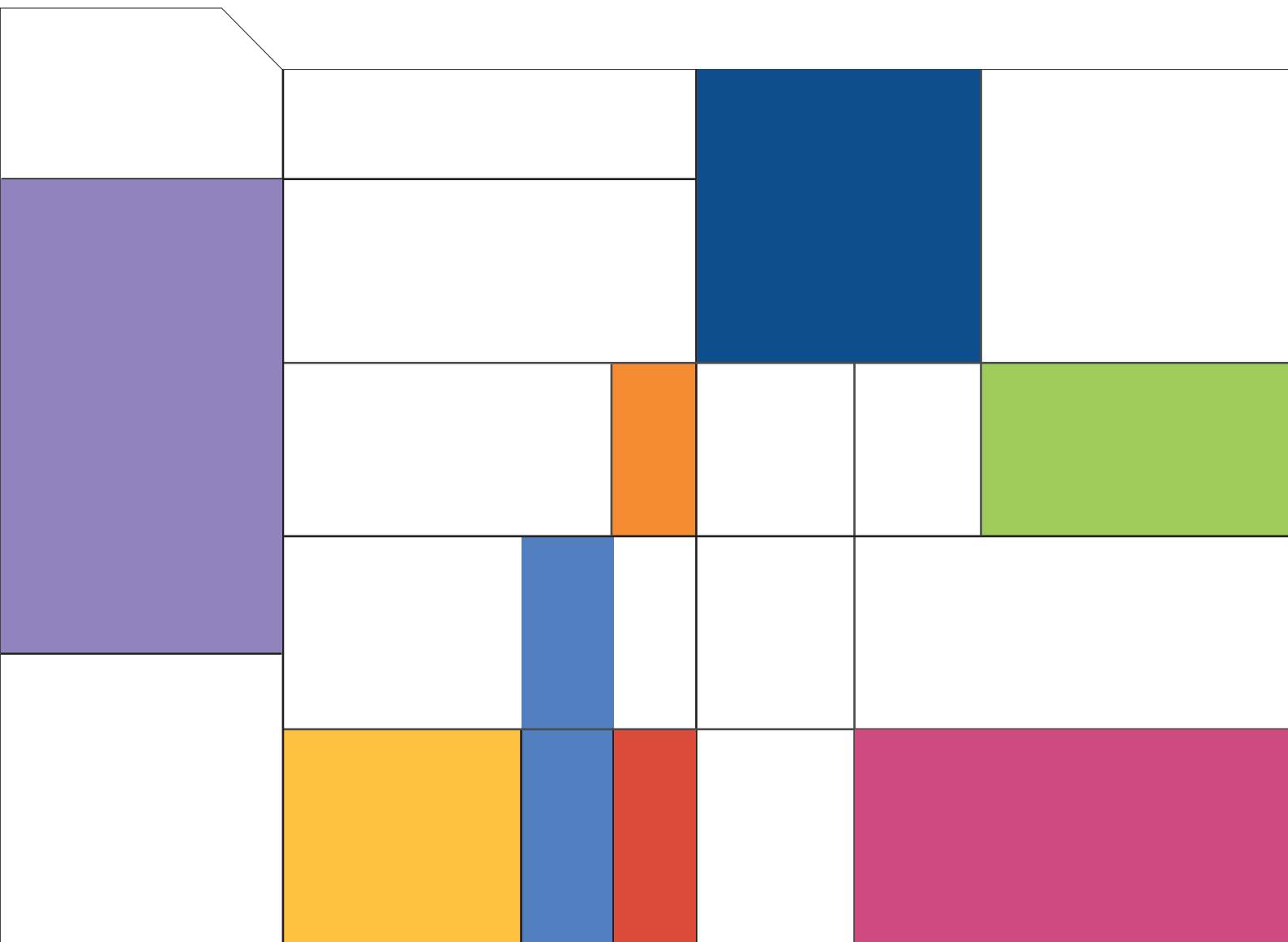




OECD Journal

Economic Studies



OECD Journal

Economic Studies

Volume 2014

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

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

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

OECD Journal: Economic Studies, volume 2014
ISSN 1995-2848 (print)
ISSN 1995-2856 (online)
http://dx.doi.org/10.1787/eco_studies-v2014-1-en

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

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

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

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

© OECD 2015

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

Foreword

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

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

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

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

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

Editorial Board

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

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

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

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

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

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

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

Bo Hansson, OECD Directorate for Education, Paris, France

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

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

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

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

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

Raed Safadi, Deputy Director, OECD Trade and Agriculture Directorate, Paris, France

Vincenzo Spezia, OECD Directorate for Science, Technology and Industry, Paris, France

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

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

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

Lyn Urmston, OECD Economics Department, Paris, France

Acknowledgement

The Editorial Board would like to thank the following external referees who kindly reviewed articles during the preparation of this volume.

Stefan Ambec, INRA, Toulouse School of Economics

Davide Furceri, International Monetary Fund, Research Department

Prakash Loungani, International Monetary Fund, Research Department

Fabrice Orlandi, European Central Bank

Fabio Pammolli, IMT Lucca Institute for Advanced Studies and CERM

Kozo Ueda, School of Political Science and Economics, Waseda University and TCER

Bruno van Pottelsberghe de la Potterie, Solvay Brussels School of Economics and Management, ULB, iCite

Table of contents

Lessons from OECD forecasts during and after the financial crisis	
By Christine Lewis and Nigel Pain	9
The effect of the global financial crisis on OECD potential output	
By Patrice Ollivaud and David Turner	41
The future of health and long-term care spending	
By Christine de la Maisonneuve and Joaquim Oliveira Martins	61
Japan's challenging debt dynamics	
By Yvan Guillemette and Jan Stráský	97
Rescuing the Phillips curve: Making use of long-term unemployment in the measurement of the NAIRU	
By Elena Rusticelli	109
Foreign direct investment and reverse technology spillovers: The effect on total factor productivity	
By Edmund Amann and Swati Virmani	129
Environmental policies and productivity growth – a critical review of empirical findings	
By Tomasz Koźluk and Vera Zipperer	155

Lessons from OECD forecasts during and after the financial crisis

by

Christine Lewis and Nigel Pain*

This paper assesses the OECD's projections for GDP growth and inflation during the global financial crisis and recovery, focusing on lessons that can be learned. Growth was repeatedly overestimated in the projections, which failed to anticipate the extent of the slowdown and later the weak pace of the recovery. Similar errors were made by many other forecasters. At the same time, inflation was stronger than expected on average. Analysis of the growth errors shows that the OECD projections in the crisis years were larger in countries with more international trade openness and greater presence of foreign banks. In the recovery, there is little evidence that an underestimate of the impact of fiscal consolidation contributed significantly to forecast errors. Instead, the repeated conditioning assumption that the euro area crisis would stabilise or ease played an important role, with growth weaker than projected in European countries where bond spreads were higher than had been assumed. But placing these errors in a historical context illustrates that the errors were not without precedent: similar-sized errors were made in the first oil price shock of the 1970s. In response to the challenges encountered in forecasting in recent years and the lessons learnt, the OECD and other international organisations have sought to improve their forecasting techniques and procedures, to improve their ability to monitor near-term developments and to better account for international linkages and financial market developments.

JEL classification: E17, E27, E32, E37, E62, E66, F47, G01

Keywords: Forecasting, economic outlook, economic fluctuations, fiscal policy

* This article is drawn from a wider project on OECD forecasts undertaken with Thai Thanh Dang, Yosuke Jin and Pete Richardson; we are grateful for their help. The authors also wish to thank Jørgen Elmeskov, Jean-Luc Schneider, Sveinbjörn Blöndal, other colleagues in the OECD Economics Department, Prakash Loungani and Volker Ziemann for comments on earlier drafts and Jérôme Brézillon for statistical assistance.

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

1. Introduction and summary

This paper assesses the performance of OECD projections for GDP growth over the period 2007-12 and places these in the context of the errors made in earlier years and by other international organisations. The focus is on the lessons that can be learned from projection errors and their cross-country differences and the resulting changes to forecasting models and procedures that have occurred since the start of the financial crisis, both inside the OECD and in other international organisations.

Forecasting the timing, depth and ramifications of the global financial crisis proved exceptionally difficult. Particular challenges included the identification of imbalances and unsustainabilities entering the crisis, the timing of their unwinding and the likely impact on real activity. These challenges were compounded by the unusually high speed and depth of cross-country interconnections between real and financial developments, the increased variability of economic growth compared with the pre-crisis period, the lack of timely data on many important financial factors and the limited understanding of macro-financial linkages. All these came on top of the well documented difficulties normally experienced when forecasting around major turning points in activity.

On average across the OECD and the BRIICS economies, calendar year GDP growth was overestimated across 2007-12, with the largest errors occurring in the projections for the vulnerable euro area economies. The largest errors were made at the height of the financial crisis in 2009 but there were also growth disappointments during the recovery. The OECD was not alone in finding this period particularly challenging. The profile and magnitude of the errors in the GDP growth projections of other international organisations and consensus forecasts are strikingly similar.

Although recent projection errors were large, they were not unprecedented: the first oil-price shock in the early 1970s also proved to be an equally difficult period for forecasters. Over a longer perspective of forty years, the OECD projections for G7 countries have generally been efficient and informative. Allowing for variations in growth volatility over time, the projection errors over 2007-12 are of a broadly similar magnitude to those in the pre-crisis years.

A number of economic characteristics are associated with the projection errors during this period. Growth was typically weaker than expected and errors higher in countries that are more open to external developments and exposed to shocks from other economies. For example, international trade openness and the presence of foreign banks in the economy are strongly associated with larger errors during the downturn period. This suggests that the projections may have failed to fully reflect the higher exposure of these economies to interconnected negative global shocks. In the recovery, growth was weaker than expected in countries in which banks had low pre-crisis capital ratios and in countries in which non-performing loans had risen strongly. Moreover, growth in countries with more regulated product and labour markets has generally proved more difficult to forecast. In part this

may reflect insufficient attention being paid to the extent to which tighter regulations have delayed the necessary reallocation across sectors in the recovery phase.

Based on the OECD forecasts, stronger projected fiscal consolidation has been associated with growth disappointments, but only in some years and only if Greece is included in the sample used. The repeated assumption that the euro crisis would dissipate over time, and that sovereign bond yield differentials would narrow, appears to have been a more important source of error. This underlines the conditional nature of OECD projections, which are not intended to be forecasts, but also raises questions about the conditioning assumptions chosen.

In response to the crisis, the OECD and other international organisations have been reviewing and changing their forecast procedures and practices. A key change has been to increase the centralisation of the early stages of the forecast round. This ensures that global economic developments and cross-country spillover effects are reflected consistently in the projections for individual economies. Monitoring and statistical modelling of near-term developments has also been enhanced, with the OECD's indicator models for near-term GDP and global trade growth proving to be a useful source of guidance. Anecdotal evidence from contacts with businesses has also become more important. In addition, there is now a stronger focus on financial market developments, with financial market indicators increasingly being integrated into projection processes and macroeconomic models. Communication efforts have also been enhanced to characterise the shape of the risk distribution around the baseline projection, including forecast ranges and the use of fan charts. Greater use is also being made of quantitative scenario analyses to illustrate the implications of key risks.

The structure of this paper is as follows. Section 2 briefly reviews forecasting and risk assessment practices before the onset of the financial crisis and then discusses the properties of the errors made for the 2007-12 period. (The focus is on the period of the crisis and the years that immediately followed because this paper is based on a project covering those years (Pain, et al., 2014); 2013 is not reported in the paper but revisions to the initial projections were generally comparatively small.) Section 3 assesses whether OECD projection errors in the 2007-12 period, or the downturn and recovery sub-periods, are systematically linked to pre-crisis conditions, structural conditions, and policy assumptions. Section 4 places the recent errors in a longer-run context to better understand whether they are unusual errors from an unusual period. A discussion of the nature of the projections follows. The final part of the paper takes stock of recent changes in forecasting procedures and practices in the OECD and other international organisations to address the identified weaknesses with earlier forecasting practices, drawing on a series of interviews undertaken in 2013. Information on the data set used and the statistical procedures undertaken is summarised in Box 1.

2. Forecasting before and during the crisis

2.1. Forecasting and risk assessment practices before the onset of the financial crisis

The pre-crisis period was one of relative economic stability. OECD projections in this period appeared to perform relatively well (Vogel, 2007), with only limited revisions to GDP growth projections typically observed as the forecast horizon narrowed, even for 2007 (Figure 1).¹ However, changes in the global economy and financial system in the early to mid-2000s contributed to the subsequent difficulties of forecasting once the crisis began.

Box 1. Data and definitions

Data

The results in this document make use of data sets for projections made by the OECD for OECD countries and six non-member countries (the BRIICS: Brazil, Russia, India, Indonesia, China and South Africa), with the main focus on OECD countries. Annual calendar year GDP growth data for 2007-12 are used for all countries except India, where annual growth over the fiscal year (to April) is used. Three different sets of projections are considered:

- May *Economic Outlook* projections for calendar year GDP growth and inflation in the same year.
- May *Economic Outlook* projections for calendar year GDP growth and inflation in the following year.
- November *Economic Outlook* projections for calendar year GDP growth and inflation in the following year.

The projection error is defined as the outturn less the projection. The outturn data for GDP growth in any given year are taken from the May *Economic Outlook* in the year immediately after. An issue for all evaluations of forecasting performance is the appropriate vintage of data to use, since the initial outturn estimates may not be especially reliable, particularly at times of rapid change in the economy (Shrestha and Marini, 2013). But use of the latest vintages of data can result in the calculated forecast errors being misleading, since they can also contain changes to national accounting procedures and concepts that were not known about at the time of the projection. This paper follows standard practice in using early realisations of the outcome.

To enable longer-run comparisons, OECD growth projections for the G7 countries from 1971 are also used. For some countries, these include GNP growth projections, rather than GDP growth, over the first part of the sample period.

Use is also made of a short-run and longer-run dataset of other forecasters' projections for GDP growth in the G7 countries. The short dataset, covering 2007-12, includes projections published by the IMF, European Commission and Consensus Economics, where the latter is the average of private-sector economists' forecasts for each country. The longer-run dataset for the G7 countries, covering 1991-2012, also includes consensus forecasts from 1991 onwards.

Key metrics

The paper uses three key measures of the size of the errors:

- The **average error** for each country, which is the average projection error (defined as above) over a given period. The means for various country groups are based on unweighted averages across countries and time.
- The **average absolute error**, defined as the average of the *absolute value* of individual country errors over the time period shown. For country groups, this is calculated as the unweighted average of absolute errors across countries and time. For greater comparability across countries and time, the average absolute error is also scaled by the corresponding average absolute growth rates. The ratio of country groups is calculated as the unweighted average ratio across countries.
- The **root mean squared error (RMSE)**, which is calculated by squaring individual country errors, then averaging these over the time period shown and taking the square root of the result. For country groups, this is the square root of the average of country squared errors, where the average is calculated across countries and time. To improve comparability, the RMSE is also scaled by the average volatility of growth (i.e. standard deviations) for a country or time period. The ratio of country groups is calculated as the unweighted average across countries.

Box 1. **Data and definitions** (cont.)**Forecast evaluation tests**

- **Unbiasedness:** tested by a pooled regression of country projection errors on a constant. Unbiasedness requires that $\alpha = 0$ in:

$$Error_{it} = \alpha + \varepsilon_{it} \quad [1]$$

- **Information content:** tested by a pooled regression. Informative projections have positive β in:

$$Outcome_{it} = \alpha + \beta Projection_{it} + \varepsilon_{it} \quad [2]$$

Information content is also tested relative to two alternative forecasts: a naïve forecast of the previous year's growth rate; and the consensus forecasts.

- **Efficiency:** this can be measured in several ways, with different degrees of strength. A basic requirement is simply that the RMSE is smaller at each forecast horizon. A slightly stronger definition is that the error should not be predictable and the projections should be informative – “weak efficiency” – which requires that both $\alpha = 0$ and $\beta = 1$ in the second regression above. A third, stronger form of efficiency is that projections embody all information available to the forecasters at the time, in which case the projection errors should be uncorrelated with informative data series. Another strong test of efficiency requires that revisions to a specific forecast be small and unpredictable from one forecasting round to the next, and uncorrelated with the previous revision.
- **Directional accuracy:** this measures whether the projections were qualitatively accurate, in the sense of accurately projecting rising or declining growth rates in the forecast period.

This period was one of increasing globalisation and integration of real and financial activity, raising the potential for cross-border and cross-market transmission of economic and financial shocks. Foreign-owned banks became more important in domestic banking markets and in many economies bank funding relied increasingly on international markets. Financial leverage and risk-taking expanded rapidly in a low interest rate environment, reflected in strong asset price and credit growth. External imbalances also built up to what were widely viewed as unsustainable levels.

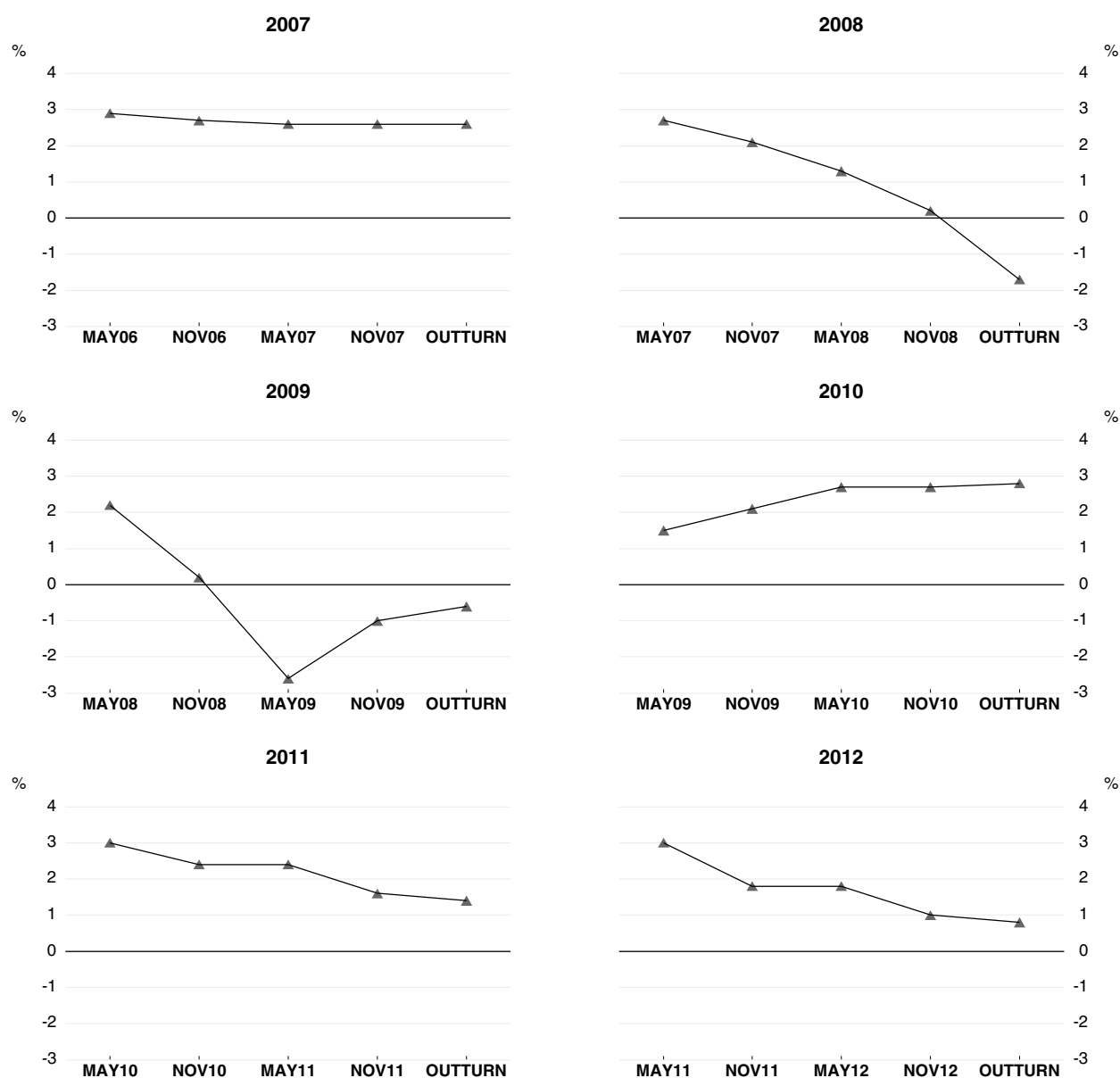
Limited weight was given to the possible impact of excessive risk-taking in the projections made before the crisis, not least because the extent of risk-taking was often hidden in off-balance-sheet activities or masked by derivative positions. That said, frequent risk analysis was done of the extent of over-valuation of housing markets and the activity implications stemming from developments such as the run-up and subsequent correction in US house prices.² However, only a handful of financial variables were integrated fully into the forecast process and the background models used; typically, account was taken only of policy interest rate and asset price effects on activity.³ Certain structural policy settings, notably less stringent product and labour market regulations, were believed to enhance the resilience of economies by improving the flexibility to bounce back relatively quickly from downturns, conditional on the assumption that financial markets and the monetary transmission mechanism functioned in a normal manner (Duval et al., 2007; Duval and Vogel, 2008). In this context, increases in financial depth were seen as being beneficial to both long-term growth and cyclical resilience.

2.2. Forecast performance during the financial crisis and its aftermath: GDP growth

As the crisis intensified, it spread across countries rapidly, in a manner well beyond that suggested by the linkages built into standard forecasting and simulation models (Bini Smaghi, 2010). Global trade collapsed in late 2008 and the early part of 2009, and

Figure 1. The evolution of year-on-year projections for OECD GDP growth

Q4-on-Q4 percentage change in GDP at each forecast date shown



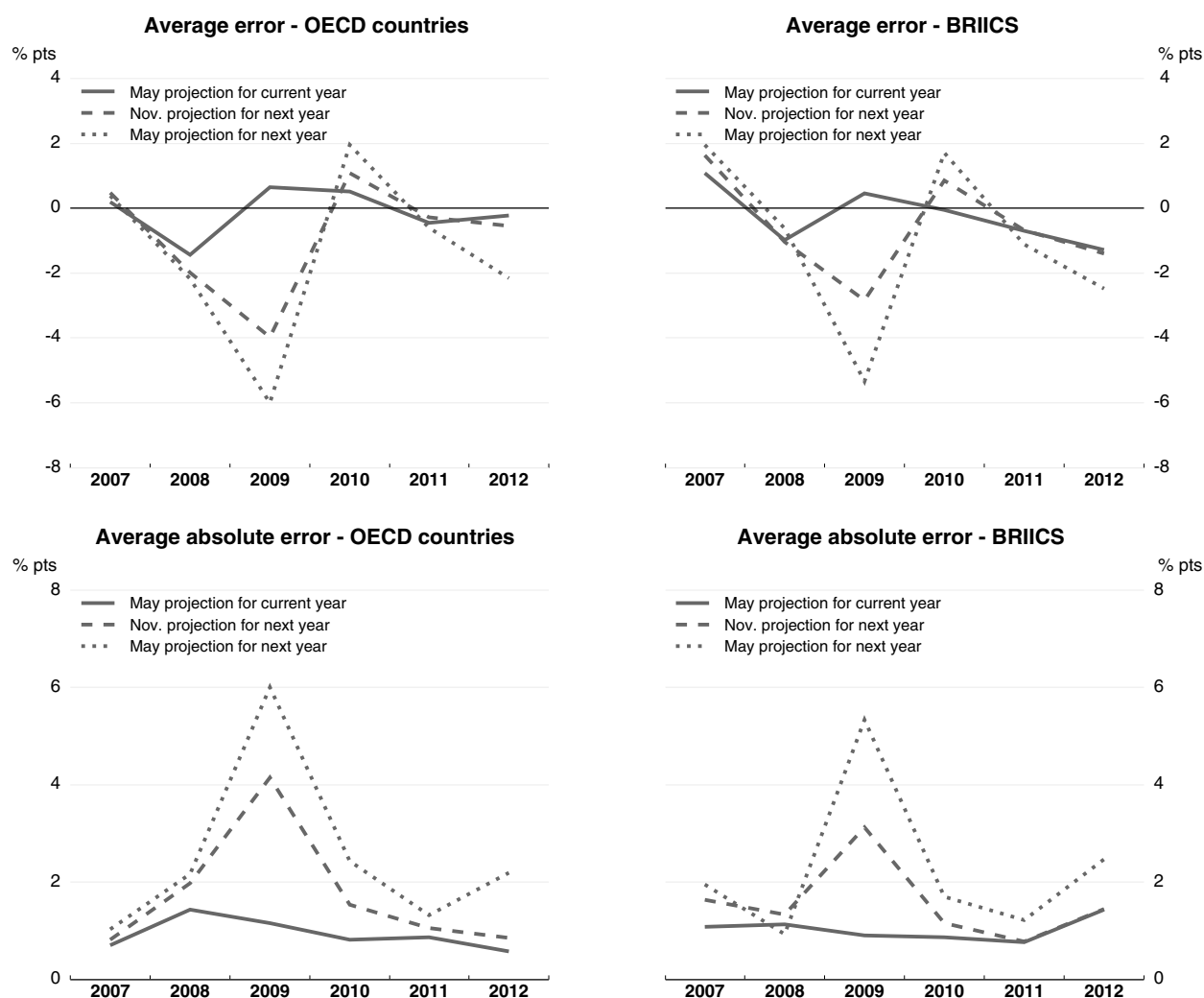
Note: Outturn is defined as the published figure at May the following year.

Source: OECD Economic Outlook Databases.

private sector sentiment fell sharply worldwide. The cumulative errors in the May 2008 projections for GDP growth in 2008 and 2009 were large. Forecasts were revised down consistently and very rapidly when the sub-prime crisis erupted and Lehman Brothers subsequently failed, with growth outturns in 2009 nonetheless still proving substantially weaker than had been projected (Figure 2).⁴ The onset of the euro area crisis in 2010, with the re-pricing of sovereign debt and banking sector risks, subsequently contributed to a further period of growth disappointments, particularly in Europe.

Thus, on average across countries, calendar year GDP growth was overestimated across 2007-12.⁵ As might be expected, given the larger information set available at the time of the forecast, the errors in (current year) projections of current year growth are smaller than in (current year) projections of GDP growth for the year ahead. On average, growth was overestimated in both the OECD and BRIICS economies, though the errors in the latter were slightly smaller than in the OECD economies (Figure 3). The largest errors were made in the vulnerable euro area economies. The large errors from the year-ahead projections for some countries during the crisis gives the distribution of the errors a negative skew; the dispersion of these errors was largest in 2009 and 2010 (Figure 4). The errors from the current year projections are more symmetric overall and the range is much narrower, even in 2009.

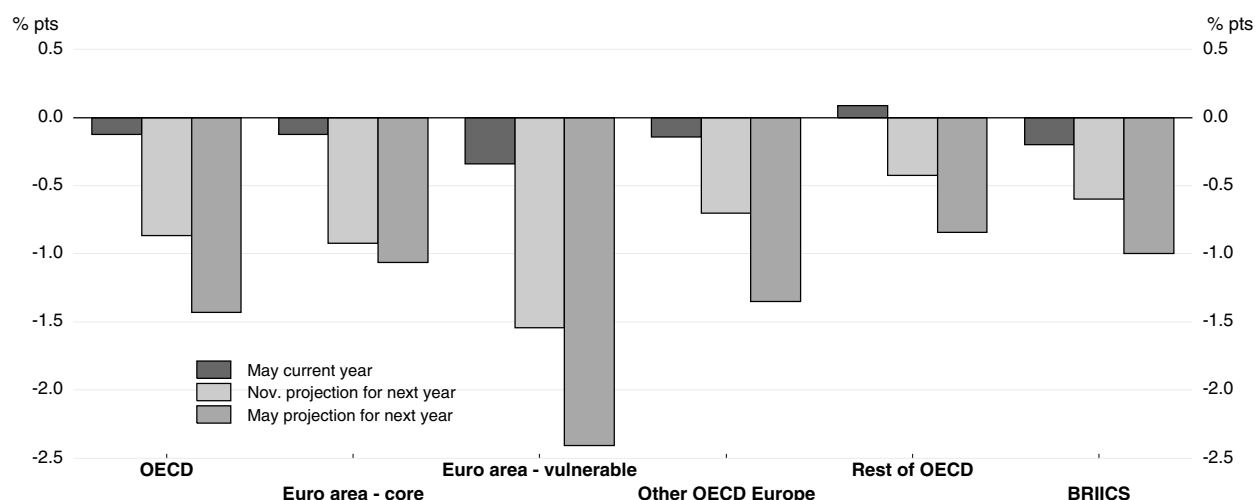
Figure 2. **Errors in calendar year GDP growth projections in recent years**



Note: An unweighted average of projection errors across countries is shown. See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outturn as at May the following year. A negative (positive) average error indicates over(under)-prediction. There is only partial coverage for Chile, Estonia, Indonesia, Israel, Slovenia and South Africa.

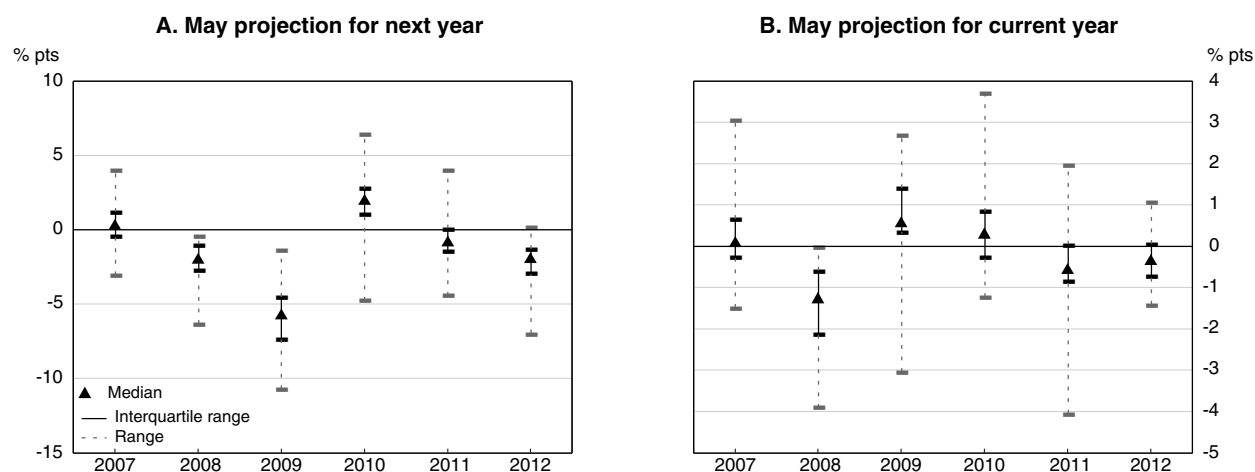
Source: OECD Economic Outlook Databases; and OECD calculations.

Figure 3. **Average errors of calendar year GDP growth projections by country group**
2007-12, percentage points



Note: See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outcome as at May the following year. A negative (positive) average error indicates over(under)-prediction. There is only partial coverage for Chile, Estonia, Indonesia, Israel, Slovenia and South Africa. Countries included in the “euro area – vulnerable” group are Greece, Ireland, Italy, Portugal and Spain.
Source: OECD Economic Outlook Databases; and OECD calculations.

Figure 4. **Dispersion of calendar GDP growth projections for OECD countries**



Note: See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outcome as at May the following year. A negative (positive) average error indicates over(under)-prediction. Countries with partial coverage are excluded (Chile, Estonia, Israel and Slovenia).
Source: OECD Economic Outlook Databases; and OECD calculations.

The differences in forecasting performance across countries narrow if the projection errors are adjusted for the considerable cross-country differences during this period in average GDP growth rates and the volatility of GDP growth. On average, in the May current-year projections, the RMSE for OECD countries is a little under one-half of the volatility of GDP growth over 2007-12, whereas for the projections of growth a year ahead it is around the same magnitude as GDP growth volatility (Appendix A). For the non-OECD countries, the adjusted measure of accuracy is similar to the OECD countries. A necessary but not

sufficient requirement for forecast efficiency is that accuracy should improve as the forecast horizon shortens. This is the case for the average error, average absolute error and RMSE for almost all countries (Appendix A).

Statistical tests of the properties of the OECD country growth projections are summarised in Pain et al. (2014), though the results need to be viewed with some caution, given the small sample size available. The main findings include:

- The projections for OECD countries are found to be biased during this period, particularly the projections for the following year.
- The projections in this period were informative, in the sense that they contain statistically-significant information correlated with the outturns. However, various tests of weak efficiency for the OECD countries are not passed at any of the horizons (other than the requirement that the projections become more accurate as the forecast horizon narrows).
- Stronger tests of efficiency are also rejected. The forecasts through the sample period were downwardly revised, by 0.6 percentage points on average in November and then 0.7 percentage points the following May when the year was underway. There is a significant positive relationship between successive revisions, which is consistent with smoothing.⁶
- The May projections for GDP growth in the same year and the following year contain additional statistically-significant information over a naïve forecast that uses only the previous year's growth rate.
- Similar results are found for other sets of forecasts, with forecast-encompassing tests suggesting that consensus forecasts do not contain additional information over the OECD projections at any horizon.⁷

Another way of judging the usefulness of the projections is by their ability to get the “story” right – or directional accuracy. In the sample period 2007-12, there were 56 instances of a year-on-year rise in GDP growth in an OECD country (growth in year t was higher than growth in year $t-1$) and 116 instances of declines.⁸ The projections correctly anticipated around 90% of the growth pick-ups, both in the same year and the year ahead. But declines were more difficult: around 90% of declines were correctly projected in the same year, but the performance of the projections made in May for growth in the following calendar year was no better than a coin toss. This feature of OECD projections was also reported in an earlier assessment for the pre-crisis period (Vogel, 2007) and is also present in the 40-year dataset discussed later (see Table 5 in Section 4).

3. Factors correlated with the growth forecast errors

3.1. Factors correlated with errors over the whole period and during the downturn

This section highlights some of the individual variables correlated with the growth projection errors over the whole period and during the downturn phase (forecasts made in May 2008 for 2008-09) and the subsequent recovery (forecasts made in May 2010 for 2010-11 and May 2011 for 2011-12). Pair-wise correlations between the country forecast errors, four broad sets of pre-crisis factors and four indicators of contemporaneous economic developments are set out in Table 1.⁹ Given the small sample size and the potential collinearity of country characteristics, some caution is warranted in interpreting these results. Nonetheless, some of the reported associations suggest potential sources of error in the projections that could be investigated in more detail in the future.¹⁰

Table 1. **Correlation between growth projection errors and pre-crisis conditions**

	Full period: 2007-12		Sub-periods		
	Average error	RMSE	Cumulative error over two years shown (at publication date indicated)		
	From projections at May for the next year		2008-09 (May-08 EO)	2010-11 (May-10 EO)	2011-12 (May-11 EO)
Openness					
Trade openness	-0.3*	0.4**	-0.4**	0.0	-0.3
Financial openness [^]	0.3	-0.3	0.4*	0.0	0.0
Foreign banks' assets	-0.3*	0.5***	-0.5***	0.1	0.0
Economy-wide regulations					
Product market regulation	-0.1	0.3*	-0.2	0.2	-0.2
Employment protection legislation	-0.1	0.4*	-0.4*	0.3	-0.2
Financial structure					
Stock market capitalisation	0.2	-0.3	0.3*	-0.1	0.0
Regulatory capital	0.1	0.4**	-0.3*	0.5***	0.3
Financial services [^]	0.1	-0.3	0.3	-0.3*	0.3
Pre-crisis imbalances					
House price growth (2000-07)	-0.2	-0.1	0.1	-0.2	-0.1
Private-sector credit growth (2000-07)	-0.2	0.0	0.0	-0.2	-0.1
Current account balance	0.3*	-0.1	0.0	0.2	0.0
Contemporaneous developments					
Equity prices	0.3	0.0	0.0	0.7***	0.8***
Non-performing loans	-0.7***	0.3*	-0.3	-0.5***	-0.5***
Business confidence (manufacturing)	0.1	0.0	0.2	0.4*	0.2
Consumer confidence	0.3*	-0.1	0.0	0.6***	0.4**

Note: Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outturn as at May the following year. A negative (positive) error indicates over(under)-prediction. Negative correlations for average errors indicate more negative growth surprises are correlated with more positive changes in the other variable. *, **, *** denote significance at the 10%, 5%, or 1% level, respectively. ^ denotes variables with extreme values that have been omitted because they change the nature of the relationship. The first three columns exclude countries with partial coverage: Chile, Estonia, Israel and Slovenia. Contemporaneous developments are the change over the relevant period being considered.

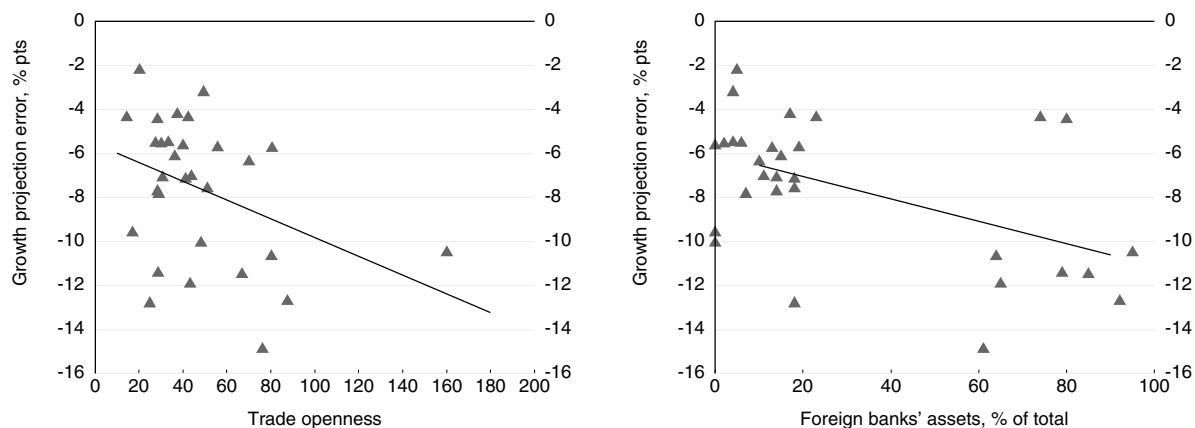
Sources: IMF; Financial Soundness Indicators; OECD Economic Outlook Databases; OECD Going for Growth (2012); OECD House Price Database; OECD Structural Analysis Database; World Bank Global Financial Development Database; World Bank World Development Indicators; and OECD calculations.

First, growth was weaker than expected and projection errors higher in countries that are more open to external developments and exposed to shocks from other economies. Factors such as openness to international trade and the share of total national banking assets held by foreign-owned banks are found to be significantly negatively correlated with forecast errors over the period as a whole and particularly so during the downturn (Figure 5). This suggests that the projections may have failed to fully reflect the increasing globalisation of real and financial activity prior to the crisis, which had raised the potential for cross-border and cross-market transmission of economic and financial shocks. In particular during the crisis, there was a tendency for foreign-owned banks – an increasingly important presence in domestic banking markets in the pre-crisis period – to cut credit extensions or reduce new lending in their host economies in order to meet lower risk targets imposed by their parent banks.¹¹

Second, larger forecast errors over 2007-12 have occurred in countries with more stringent pre-crisis labour and product market regulations (Figure 6). This finding is difficult to interpret, but may in part reflect the weight given to pre-crisis empirical

evidence that tight regulations could help to cushion economic shocks (Duval et al., 2007), and insufficient attention being paid to the extent to which tighter regulations could delay necessary reallocations across sectors in the recovery phase. A third possibility is that it reflects a correlation between restrictive regulation and the pre-crisis build-up of imbalances, although indicators such as pre-crisis house price and credit growth are not found to be strongly correlated with the forecast errors.¹²

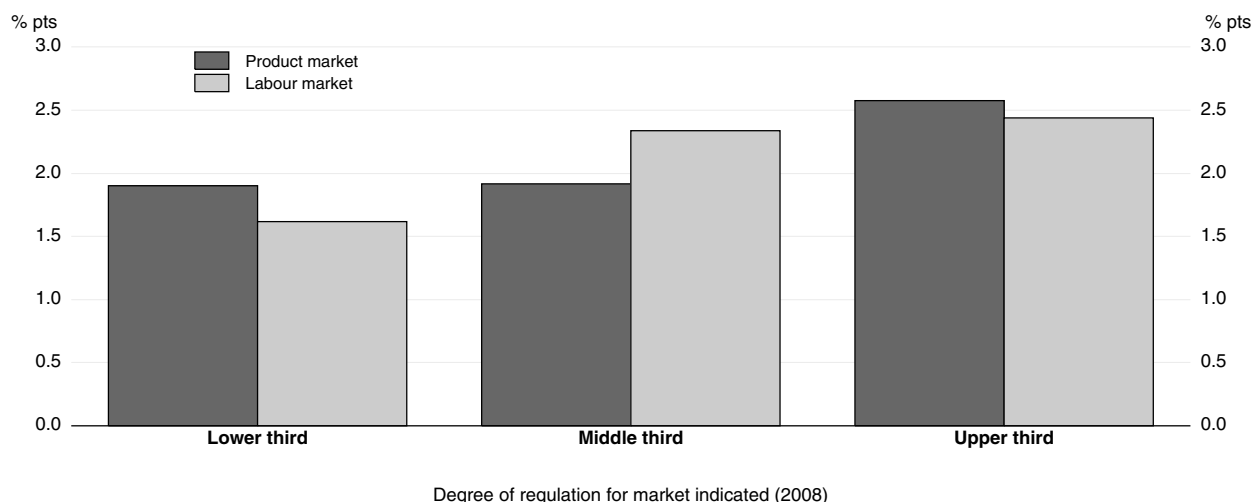
Figure 5. Relationship between growth projection errors for 2008-09 and other characteristics
Cumulative errors for 2008-09 calendar year projections made at May 2008



Note: See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less forecast growth at each forecast horizon, where actual growth is the published outturn as at May the following year. A negative (positive) error indicates over(under)-prediction. Countries with partial coverage for the period are excluded (Chile, Estonia, Israel and Slovenia).

Source: OECD Economic Outlook Databases; World Bank Global Financial Development Database; and OECD calculations.

Figure 6. Relationship between accuracy of growth projections and regulation
Average RMSE of growth projections for 2007-12 made at November a year earlier



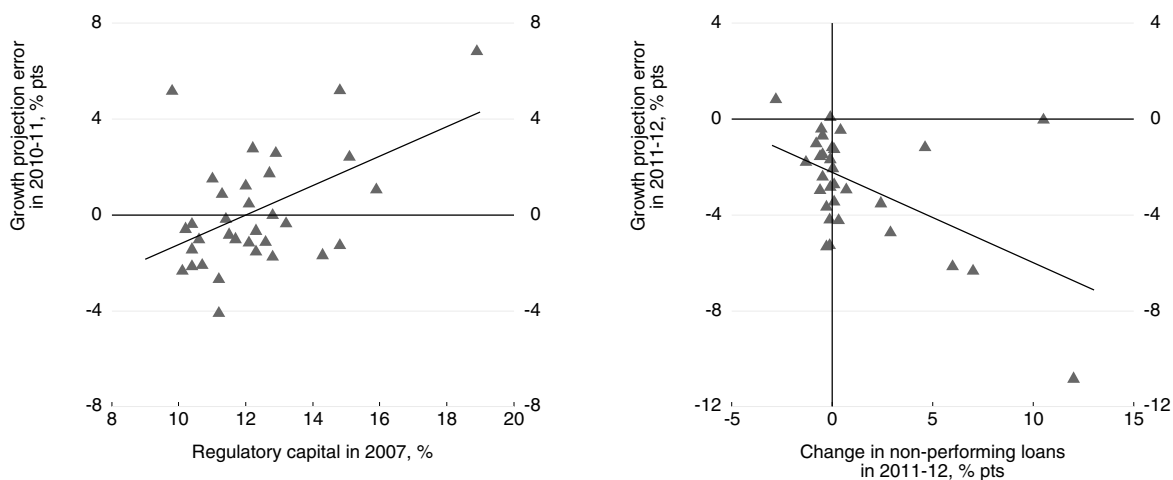
Note: See Box 1 for a description of the calculation of errors and the different projections considered. Countries with partial coverage for the period are excluded (Chile, Estonia, Israel and Slovenia).

Source: OECD, *Going for Growth* (2012); OECD Economic Outlook Databases; and OECD calculations.

Third, growth was weaker than expected in the recovery in countries in which banks had a low ratio of regulatory capital to risk-weighted assets in 2007 (Figure 7, left panel). This may suggest that insufficient account was taken of the post-crisis deleveraging undertaken by weakly capitalised banks. Although, over the period as a whole, sizeable projection errors (i.e. large RMSEs) have been made in countries in which banks have comparatively high levels of regulatory capital, this primarily reflects the higher underlying growth volatility of these economies in this period.¹³

Finally, during the post-crisis recovery period there are strong correlations between projection errors and changes in equity prices, private sector confidence and non-performing loans in the banking sector (Figure 7, right panel). This again suggests that the projections may have underestimated the negative impacts of impaired banking systems and weak confidence.¹⁴

Figure 7. **Relationship between errors in growth projections in the recovery and the financial sector**



Note: See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outcome as at May the following year. A negative (positive) error indicates over(under)-prediction.

Source: IMF, Financial Soundness Indicators; OECD Economic Outlook Databases; World Bank Global Financial Development Database; and OECD calculations.

Perhaps surprisingly, indicators of pre-crisis vulnerabilities such as house price increases and private credit growth are not generally associated with negative growth surprises since the crisis began.¹⁵ A possible explanation might be that these vulnerabilities have been generally well understood and incorporated successfully into forecast judgements. However, larger downside surprises on average have occurred in countries with a pre-crisis current account deficit.

Changes in the OECD composite leading indicators (CLIs) ahead of May 2008 were negatively correlated with the errors in the growth projections made at that time, with the CLIs continuing to improve in some countries even at the point at which GDP was past its peak (see Appendix 7 in Pain et al., 2014). This suggests that the usual relationship between many high-frequency variables – including activity, labour market and survey data – broke down during this period.

3.2. Fiscal consolidation, bond spreads and growth disappointments in 2010-11 and 2011-12

In 2010-11 and 2011-12, the pick-up in the pace of the global recovery projected in the May 2010 and May 2011 *Economic Outlook* ultimately failed to materialise. As mentioned above, one factor behind this may have been an underestimate of the impact of impaired banking systems and bank deleveraging during these years (Figure 7). But these years also saw the deepening of the euro area crisis and the onset of widespread fiscal consolidation, providing other possible explanations for the projection errors. IMF studies point to a negative relationship between the errors in its Spring 2010 *World Economic Outlook* projections for GDP growth over 2010-11 and projected fiscal consolidation during this period (IMF, 2012; Blanchard and Leigh, 2013). An implication drawn from this is that fiscal multipliers were underestimated considerably, thereby contributing to an overestimate of growth prospects.¹⁶

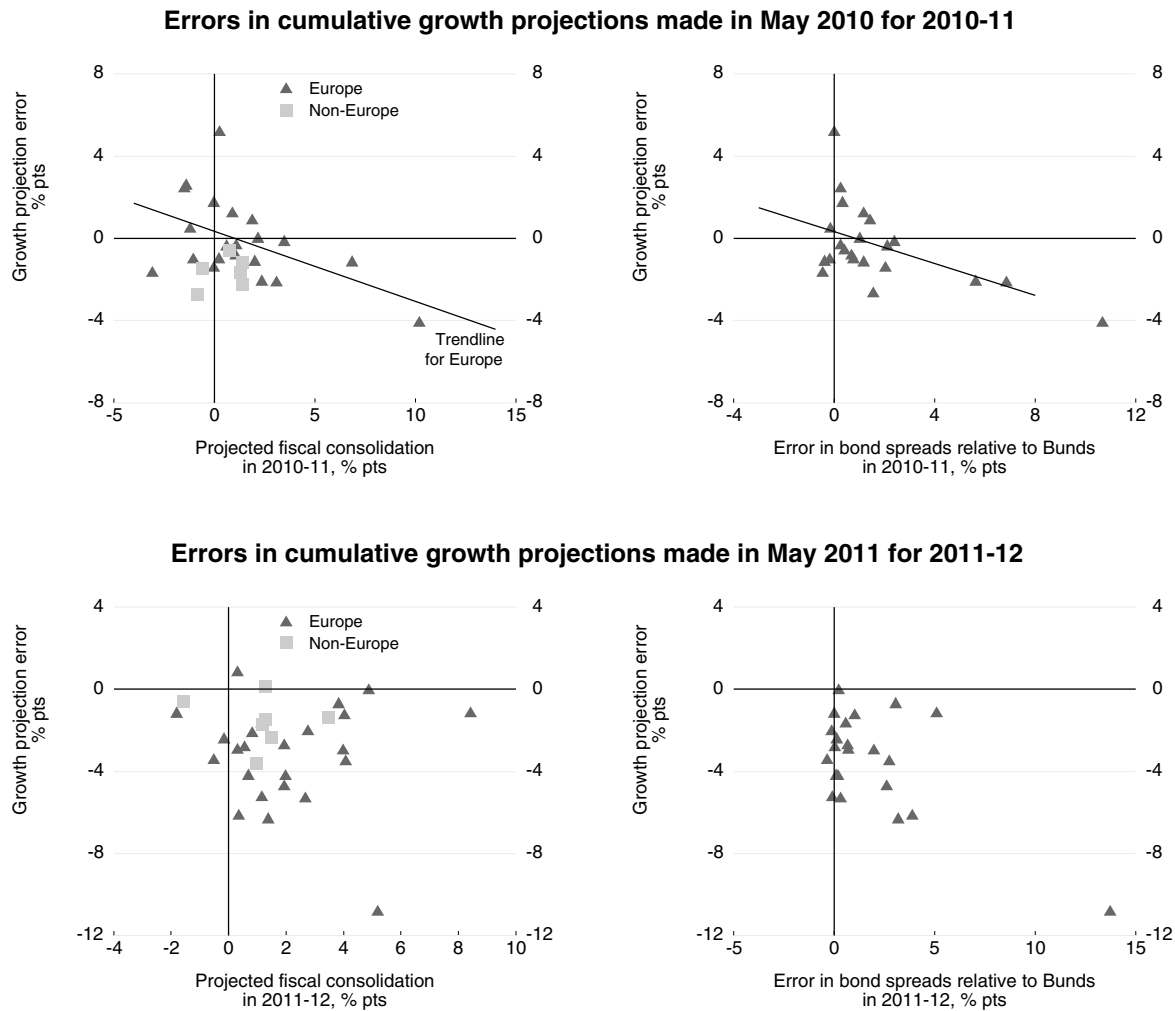
The OECD projections published in the *Economic Outlook* in May 2010 for cumulative GDP growth during 2010-11 offer only partial support for this hypothesis. Projected fiscal consolidation is indeed negatively correlated with the growth forecast errors in 2010-11 (Figure 8, upper left panel). But this relationship stems from the downside growth surprises in the European economies, and in particular, Greece; without Greece, the relationship remains negative but no longer statistically significant. Outside Europe, the relationship is harder to discern. The projections from May 2011 for cumulative growth 2011-12 are not related to fiscal consolidation (as was the case in the *World Economic Outlook* forecasts). To the extent that the earlier relationship reflects an initial underestimation of the multiplier effects of fiscal consolidation, the lack of relationship in the projections made one year later suggests that forecasters revised their judgement as the impact of consolidation in an environment with limited monetary policy space became clearer.

A second possibility is that the correlation between fiscal consolidation and growth forecast errors in 2010-11 could rise from an underestimate of the amount of fiscal consolidation, rather than from an underestimate of the fiscal multiplier. On average, in the European economies, cumulative fiscal consolidation in 2010-11 was 0.8% of GDP greater than projected in the May 2010 *Economic Outlook*.¹⁷ The actual level of fiscal consolidation is also negatively correlated with the growth forecast errors in 2010-11, but this relationship is again found to be dependent on the inclusion of Greece in the sample considered.

A further aspect of the published OECD projections in these years was that they were based on a “muddling through” assumption, with the euro area crisis assumed to diminish slowly and government bond spreads between other European countries and Germany assumed to narrow over the projection period. In fact, spreads widened in many countries over the projection period. There is a clear correlation between the errors in the assumptions about bond spreads and the growth forecast errors in 2010-11 (Figure 8, upper right panel), with growth weaker than projected in countries whose bond spreads relative to Germany were higher than had been assumed. This negative correlation is robust to the exclusion of Greece.

Because bond spreads are clearly endogenous to growth outcomes and the possibility of further fiscal consolidation, it is difficult to identify conclusively whether incorrect assumptions about fiscal multipliers or bond spreads are more important. However, econometric evidence suggests that the errors in the bond spread assumptions appear to

Figure 8. **Relationship between errors in growth projections in the recovery and economic developments**



Note: See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less forecast growth at each forecast horizon, where actual growth is the published outturn as at May the following year. A negative (positive) error indicates over(under)-prediction. The trendline for the error in bond spreads to Bunds excludes Greece. Trendlines are not shown where the relationship is not statistically significant at the 10% level.

Source: *Economic Outlook Databases*; and OECD calculations.

be a more robust explanation of forecast errors than the possible underestimation of fiscal multipliers or the underestimation of the amount of fiscal consolidation. Considered together, there is no longer any significant relationship between the forecast errors and projected fiscal consolidation, but the errors in bond spreads remain negatively, and in most cases significantly, related to the forecast errors for 2010-11 (Table 2). A similar result is reported in European Commission (2012). Repeating the exercise for 2011-12 reinforces the finding that projected fiscal consolidation is not significantly related to the errors. The relationship with bond spreads is not robust and varies according to the estimation techniques used and whether Greece is included in the sample.

These findings raise questions about the appropriateness of the forecast assumptions made during the recovery, particularly the prevailing assumption that the euro area crisis would ease over the projection period. This issue is discussed further below. Other factors also contributed to the projection errors over 2010-12, including the underestimation of the impact of impaired banking systems and bank deleveraging (Figure 7).

Table 2. Explaining the errors in the 2010-11 growth projections

Dependent variable: Cumulative error in the GDP growth projections made in May 2010 for 2010-11

Country sample	Europe				Europe excl. Greece			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS
Error in bond spreads (to Bunds)	-0.43*** (0.087)		-0.47** (0.210)	-0.36*** (0.126)	-0.44** (0.181)		-0.49* (0.243)	-0.36** (0.135)
Projected fiscal consolidation		-0.31** (0.136)	0.02 (0.207)	-0.05 (0.155)		-0.19 (0.168)	0.01 (0.213)	-0.05 (0.193)
Constant	0.52 (0.484)	0.22 (0.488)	0.54 (0.508)	0.47 (0.482)	0.52 (0.563)	0.16 (0.499)	0.57 (0.599)	0.46 (0.553)
Obs	21	21	21	21	20	20	20	20
R-squared	0.34	0.22	0.34	0.34	0.17	0.05	0.16	0.17
F-statistic of excluded instruments from first stage	61.0	-	21.7	-	16.1	-	12.9	-

Note: The sample is restricted to those countries with values for both variables. Robust standard errors are reported in parenthesis. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

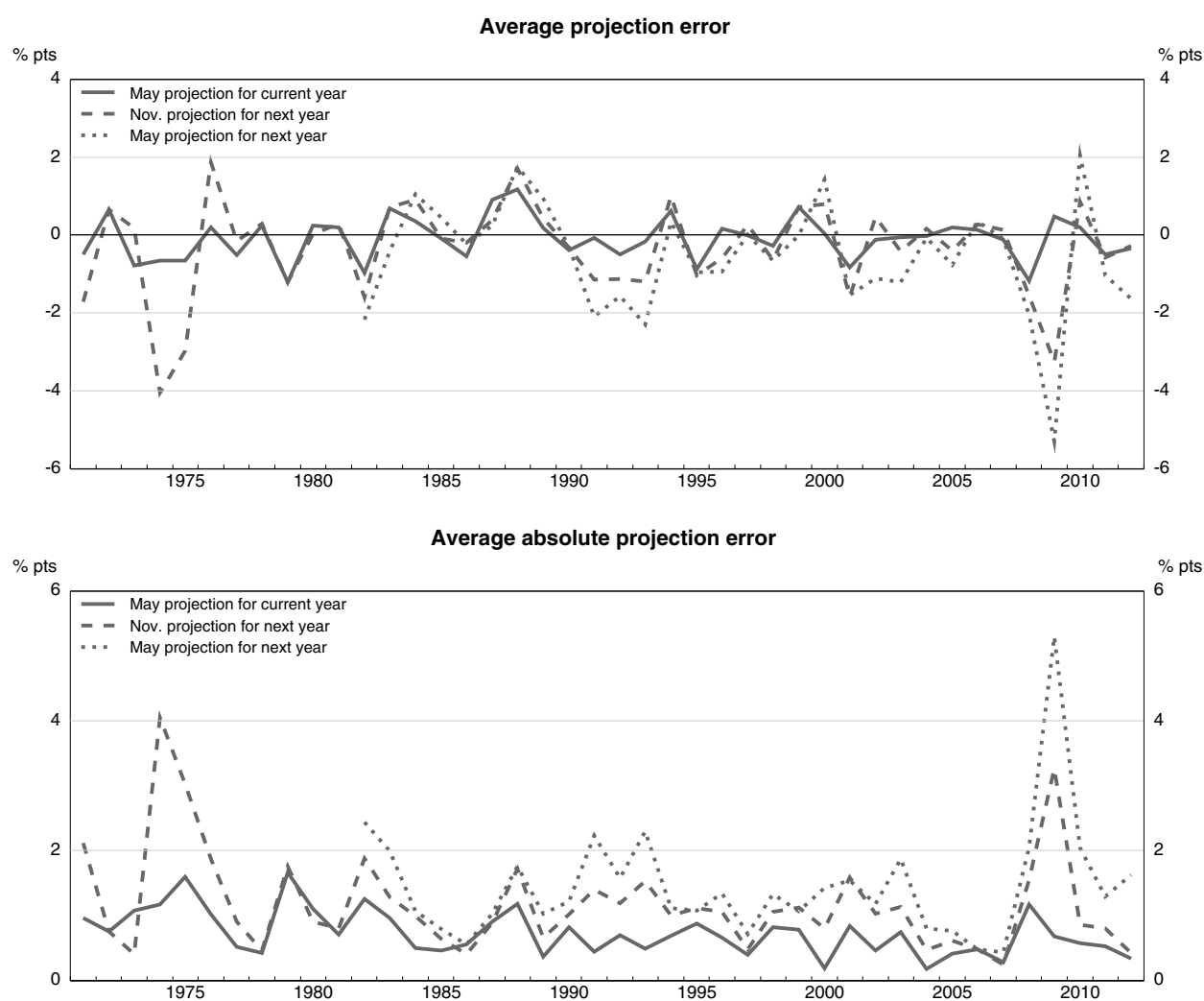
Source: OECD calculations.

4. A broader perspective on the size of recent growth projection errors

4.1. A longer-term perspective: OECD forecasts over 1971-2012

Examining the characteristics of the OECD growth projections for the G7 countries – for which there are data from 1971 – helps to place the characteristics of recent forecast errors in context, and allows an assessment of whether they are markedly different from past experience. Key findings include:

- The errors made during the crisis are large, but not unprecedented. Errors of a magnitude comparable with those in 2009 occurred in the early 1970s, following the first oil crisis (Figure 9).¹⁸
- The average absolute projection error in the G7 economies declined from the late 1970s to the mid-2000s. In part, this reflects the moderation in the volatility of output growth. At shorter horizons the decline might also reflect improved data availability, especially regarding near-term developments, and improvements in forecasting procedures.
- In many cases, weak efficiency of the projections for G7 countries over four decades cannot be rejected (Table 3).¹⁹ The near-term forecasts perform best; fewer of the May projections for growth in the next year are weakly efficient, although the sample is shorter.
- Overall, the near-term projections for G7 countries have been unbiased. The average error in the May projections for growth in the same year is close to zero (or not statistically different from zero) for each country. In contrast, the November and May projections for growth the following calendar year have, on average, been overly optimistic, although in many cases the bias is not statistically significant. In the May projections for GDP growth in the following year, the extent of over-prediction

Figure 9. **Errors in GDP growth forecasts for G7 countries over the long run**

Note: In each year the error is an unweighted average across countries. See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less forecast growth at each forecast horizon, where actual growth is the published outturn as at May the following year. A negative (positive) average error indicates over(under)-prediction. May projections for GDP growth in the following year are available only from 1982.

Source: OECD, *Economic Outlook* publications and databases; and OECD calculations.

over 2007-12 was markedly worse than the average over the preceding 25 years (from when data are available).²⁰

- On average across the G7 countries, the RMSEs over 2007-12 are higher than over the previous 15-year period of relative stability and low inflation, from 1991-2006 (Table 4, left panel). However, this difference disappears if the RMSEs are standardised by the volatility of GDP growth in the different samples. This is true also over a longer period from 1971 onwards; indeed the errors made during the period from 2007-12 are smaller than those made in the decade after 1971, reflecting the difficulties in forecasting the impact of the oil price shocks, as well as the exceptionally high volatility in GDP growth since the start of the financial crisis.

Table 3. The long-run GDP growth forecasts are weakly efficient
p-values from tests of forecast efficiency

	May projection for current year	November projection for following year	May projection for following year
	1971-2012	1971-2012	1982-2012
Canada	0.342	0.349	0.000
France	0.104	0.064	0.002
Germany	0.113	0.381	0.020
Italy	0.376	0.161	0.008
Japan	0.036	0.057	0.390
United Kingdom	0.687	0.373	0.427
United States	0.264	0.893	0.878

Note: Based on joint tests of whether in country-level regressions of the outcome on the projection and a constant, the coefficient for the projections is one and the constant is zero, with shading indicating significance at the 10% level. Newey-West standard errors are used. May projections are only available from 1982.

Source: OECD calculations.

Table 4. RMSE of calendar year GDP growth projections over the longer run
Unweighted average of G7 countries

	Percentage points of GDP			Ratio to standard deviation of growth		
	May projection for current year	November projection for next year	May projection for next year	May projection for current year	November projection for next year	May projection for next year
1971-2012	1.0	1.7	n.a.	0.4	0.7	n.a.
1971-1981	1.3	2.2	n.a.	0.5	0.8	n.a.
1982-1990	1.0	1.3	1.7	0.6	0.8	1.0
1991-2006	0.7	1.3	1.6	0.5	0.9	1.2
2007-2012	0.7	1.6	2.7	0.3	0.6	1.1

Note: Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outcome as at May the following year. May projections for the following year are only available from 1982.

Source: OECD Economic Outlook publications and databases; and OECD calculations.

The general tendency to overestimate growth in the year ahead, on average, is common to many forecasters (Abreu, 2011). Two reasons appear to be a common failing to predict downturns and to predict their size: directional accuracy is asymmetric, with a much lower share of decelerations and recessions predicted a year in advance; and errors are larger in recessions. These difficulties have been found across forecasters, across countries and over longer periods of time (Zarnowitz, 1991; Loungani, 2001; Abreu, 2011; González Cabanillas and Terzi, 2012).

In contrast, the directional accuracy of the long sample of G7 growth projections does not appear very different from that of the broader sample of countries over 2007-12 reported earlier (Table 5). There is a high degree of accuracy in projecting growth pick-ups, but little success predicting declines in growth (especially in the year ahead, with the success rate approximately as good as a coin toss). A similar finding is reported by Vogel (2007), using data for 1991-2006.

Table 5. **Directional accuracy of May growth projections**

	Increases in growth	Decreases in growth
OECD countries: 2007-12		
Number in period	56	116
% correct: Projections for same year	86	88
% correct: Projections for next year	91	46
G7 countries: 1982-2006		
Number in period	82	78
% correct: Projections for same year	79	83
% correct: Projections for next year	74	45

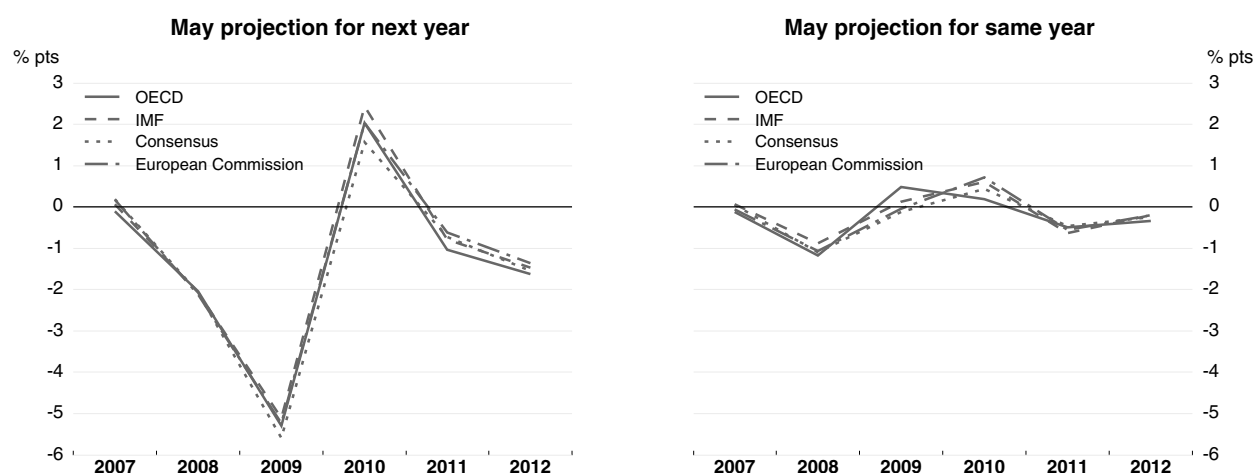
Note: The threshold for increase or decrease is a change in growth exceeding 0.1 percentage points (from rounded data).

Source: OECD, *Economic Outlook* publications, databases and calculations.

4.2. Comparisons with other forecasters

Overall, the size and profile of the OECD projection errors are very similar to the errors in the projections made by other international organisations, such as the IMF and the European Commission (González Cabanillas and Terzi, 2012), and also the errors that are implied by consensus forecasts (Ahir and Loungani, 2014) (Figure 10). All of these overestimated growth considerably in 2009 and were then surprised by the extent of the initial bounce-back in 2010 and the subsequent weakness of the recovery. On average, all forecasters over-predicted growth during the period 2007-12. No set of projections clearly outperformed the others during this period; for instance, the RMSEs of the calendar year growth projections made in May the same year were almost identical across forecasters for each of the G7 economies, possibly reflecting a collective “groupthink” underlying the projections and their assumptions.

A comparison of the OECD growth projections and consensus forecasts for G7 countries over a longer period reveals similar patterns (Pain et al., 2014), with little significant difference

Figure 10. **Comparison of average projection errors for growth in G7 countries**

Note: An unweighted average of projection errors across countries is shown. See Box 1 for a description of the calculation of errors and the different projections considered. Errors are calculated as actual growth less forecast growth at each forecast horizon, where actual growth is the published outcome as at May the following year. A negative (positive) average error indicates over(under)-prediction. European Commission and IMF projections are released in April.

Source: Consensus Economics; European Commission European Economic Forecast publications; IMF *World Economic Outlook Databases*; OECD *Economic Outlook Databases*; and OECD calculations.

in the overall accuracy of the OECD and consensus forecasts over the past two decades. Forecast-encompassing tests (Fair and Shiller, 1990) suggest that the OECD May current-year projections and November year-ahead projections encompass the consensus forecasts (i.e. contain all the relevant information found in the consensus forecasts). However, neither of the separate May projections for growth in the following year encompasses the other. This is the case for both the pre-crisis period and also over 2007-2012.

4.3. The nature of the projections

The OECD forecasts are conditional projections rather than pure forecasts. The projections are consistent with the advice given about monetary policy settings and rest on explicit and implicit assumptions, such as fiscal policy changes and whether the euro area crisis will be successfully contained. Some potentially endogenous factors, such as commodity prices and nominal exchange rates, are also switched off over the projection period.²¹ This raises a number of difficult issues.

One issue is that the projections, and especially the accompanying commentary, sometimes point to unsustainable developments and the need for policy changes. If these are built into the projections, but the advice is not followed (or vice versa), an error in the projections will likely occur. It is not clear how or whether this type of error can be controlled for in the projection process.

A related issue, as noted above, is that the projections made during the euro area crisis placed a lot of weight on the assumption that the euro area crisis would be contained and subsequently ease, with accompanying declines in sovereign bond spreads and improvements in confidence. Yet it is not clear what else could have been assumed in projections for a period of over two years made public by an inter-governmental organisation. One possible answer – and the route chosen by the OECD at the time – is to make extensive use of scenario analyses to quantify possible adverse outcomes if key assumptions did not hold.

A third issue relates to the interpretation of the forecasts since the onset of the crisis. Given the skewed distribution of possible projections, depending on the assumptions chosen, and the emergence of fat negative tail risks, the nature of the published projections has changed. Whereas previously they were best seen as mean projections, more recently they have become modal forecasts, i.e. reflecting the most likely outcome. This has, however, resulted in a communications problem, as many users of the projections still view them as mean projections with a balanced distribution of risks.

5. Recent changes in forecasting techniques and procedures and the unfinished agenda

The challenges encountered in forecasting in recent years have led to changes in forecasting techniques and procedures in the OECD and in other international organisations, some of them still in progress. Key developments are outlined below, drawing on OECD experience and consultations with experts from the IMF, World Bank, European Commission and European Central Bank.²²

5.1. Increases in the centralisation or “top-down” component of the forecast process

Reflecting the common errors made in all country forecasts in recent years, and the extent to which these have been associated with stronger than expected global financial,

trade and sentiment interconnections, increasing weight has been placed on central strategic guidance at an early stage of the forecast process, at the OECD and elsewhere.

- One aspect has been early identification of key global developments and risks and their likely quantitative implications for global activity and global trade. This ensures a consistent view amongst country specialists of forces acting on the forecasts and outlook for the major economies.
- Initial early guidance in the form of top-down centralised projections is now provided at the OECD and elsewhere, to help ensure that projections for individual countries are based on a common general storyline. These projections draw together the key points from standalone analyses such as indicator models and assessments of financial market developments.

5.2. Enhanced monitoring of near-term activity developments

Although it took time for the full effects of the sub-prime and Lehman Brothers collapse to start to appear in published activity data, they were reflected swiftly in high-frequency tendency surveys. Developments such as the generalised collapse of confidence during the Great Recession and more recently when the euro area crisis intensified have also pointed to the potential usefulness of early survey-based signals. This has been reflected in a number of ongoing developments:

- The longstanding use of statistical composite leading indicators by the OECD has been augmented by the use of empirical “nowcasting” models, exploiting high-frequency information to gain an early picture of key activity developments (Box 2). Examples used by the OECD in forming projections for the *Economic Outlook* include the suite of quarterly GDP growth models for G7 economies introduced a decade ago (Sédillot and Pain, 2003), and the more recently-developed suite of indicator models for global trade (Guichard and Rusticelli, 2011). The OECD growth indicator models during the crisis provided a very useful real-time signal of a significant slowdown and then major contraction. Elsewhere, the indicator model approach is being extended by using high-frequency information to model private sector expenditure. Assessments are also being undertaken of possible non-linearities in the relationship between indicator variables and growth outcomes at times of extreme stress.
- Greater use has also been made of the anecdotal information gathered from outside business contacts. At times of fast-moving financial developments these may provide an early signal of changes in factors such as credit conditions, and thus near-term activity developments. One example given in the interviews was an early signal about the impact of a freezing of trade credits at the height of the crisis.
- Increasing attempts have been made, though not yet in the OECD, to utilise high-frequency information on real-time activity developments provided by internet-based indicators (“big data”). Internet-based search measures, based on Google Trends, have been used to identify early signals about specific housing and labour market developments (Hellerstein and Middeldorp, 2012; McLaren and Shanbhogue, 2011), as well as policy uncertainty (Baker et al., 2012). Recent studies have also sought to use real-time financial transactions data (such as the volume of SWIFT banking transactions, or data on credit card transactions) as an early indicator of GDP growth and trends in global trade finance (Gill et al., 2012).²³ The usefulness of such indicators when combined with other longstanding indicator variables has yet to be explored.

Box 2. Indicator models, composite leading indicators and OECD projections

Two models available to OECD forecasters are the short-term indicator models, which are used to predict GDP growth for two quarters, and the OECD Composite Leading Indicators (CLIs), which are designed to predict turning points in growth relative to trend. As explained below, both tools are designed to extract information about growth from high frequency and timely data sources. By responding reasonably rapidly, the reliance on the indicator models likely reduced overall forecast errors. But it does seem that better use of the information in the CLIs could have reduced the forecast errors in the November projections for growth in the following year (but not the other horizons).

Although both tools combine recent data to produce signals about economic growth, they are different by construction.

- The short-term indicator models (Sédillot and Pain, 2003 and Mourougane, 2006) are regression-based models that efficiently combine available monthly and quarterly data to predict current quarter and quarter-ahead GDP growth for the G7 countries and the euro area. The input data include “hard” data, such as industrial production and retail sales, and “soft” data, such as business and consumer sentiment. Because the dependent variable is GDP growth, the process of producing a forecast is straightforward. Other work has shown the main gain from the indicator approach is for current-quarter forecasts, where estimated indicator models appear to outperform autoregressive time series models, both in terms of size of error and directional accuracy.
- The OECD CLIs are long-running series produced for all OECD countries and some non-OECD countries. For each country, 5-10 series are used that best predict turning points in the business cycle. The CLI is typically available with a two-month lag. Because the results are primarily qualitative, it is possible that the information contained in these series is not used optimally in the forecasting round.

The performance of the indicator models in the crisis

The OECD growth indicator models during the crisis provided a very useful real-time signal of a significant slowdown and then major contraction (see figure below). Nonetheless, large errors were recorded in the fourth quarter 2008 and early 2009 and even when the third quarter 2008 was underway, the contraction in growth was not foreseen. But both the downturn and bounce-back were highlighted relatively quickly. Overall, the indicator models were useful in compiling the projections at that time and likely reduced the size of the overall errors made.

The real-time performance of the indicator models during the crisis can also be assessed relative to that in the pre-crisis period and the simulated out-of-sample performance when the models were first developed. The root mean squared errors (RMSE) for the G7 countries are shown below (see table). There are slight differences in the timing and information sets for each forecast, but in almost all cases one to two months of hard indicator data and two to three months of survey data were available for each economy. Four main points emerge from the 2007-12 period:

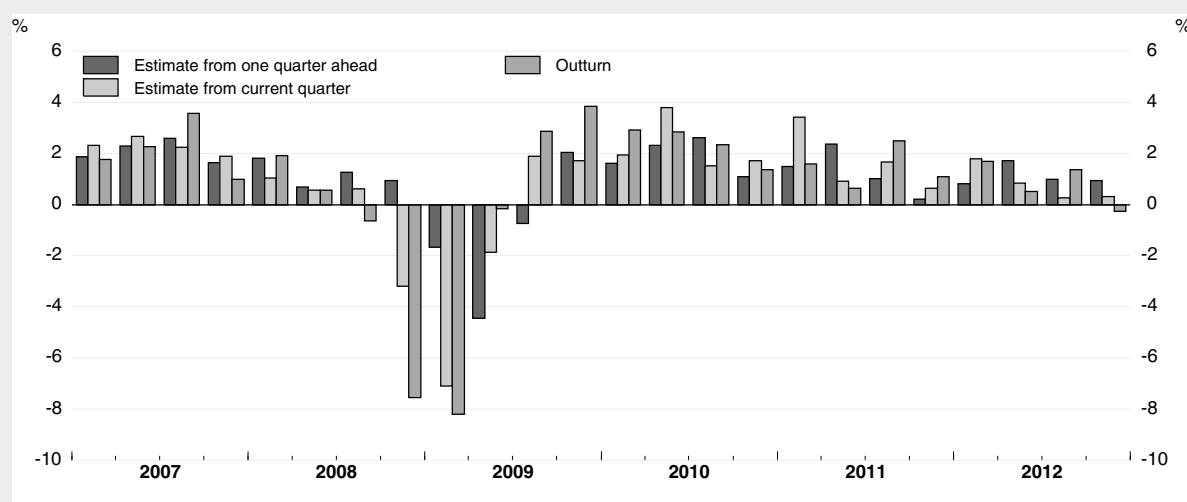
- The current quarter real-time forecast errors are comparable with those before the crisis (Sédillot and Pain, 2005), both in terms of the cross-country pattern of the errors and the size of the errors. A notable exception is the United Kingdom, where the real-time errors since 2007 have been more than double those in the pre-crisis period. The errors were also larger in Japan. But overall, the data suggest that the current quarter models have remained a useful guide to current economic activity.
- The real-time forecast errors in the quarter-ahead forecasts were generally larger than might have been expected based on those from the pre-crisis period and the initial out-of-sample exercise, particularly in Japan, Germany and the United Kingdom. However, the cross-country differences in the magnitude of the errors are broadly comparable with the pre-crisis period. The generally-higher real-time errors over 2007-12 appear to be largely due to the relatively greater volatility of growth outcomes in the recent period: when the RMSEs are scaled by the standard deviation of growth outcomes, the errors over 2007-12 are all smaller, except for the United Kingdom.

Box 2. Indicator models, composite leading indicators and OECD projections (cont.)

- In all economies the real-time errors for the one-quarter ahead forecasts have been greater than those for the current quarter forecasts. The difference between the two is smallest for France.
- The published forecasts appear to have incorporated the information from the indicator model. The current quarter errors from the published projections are positively and significantly correlated with those from the indicator model. And there is only a statistically significant relationship (at the 10% level) between the errors from the published current quarter forecasts and the current quarter indicator model projections if 2008Q4 is included.

OECD indicator model estimates and G7 GDP growth

Annualised quarter-on-quarter percentage changes, 2007Q1-2012Q4



Note: Outturn is defined as the published figure at the following *Economic Outlook*. The indicator model estimates shown are real-time estimates, typically from the last month of the quarter. Data are weighted using nominal GDP at PPP rates in 2010.

Source: OECD Economic Outlook Databases; and OECD calculations.

The current quarter indicator model has performed well in recent years

RMSE of annualised quarter-on-quarter GDP growth projections, percentage points

	Current quarter			One-quarter ahead		
	Simulated RMSE		Actual RMSE	Simulated RMSE		Actual RMSE
	1998Q1-2002Q4	2003Q1-2005Q1		1998Q1-2002Q4	2003Q1-2005Q1	
United States	1.6	1.4	1.5	2.1	1.7	2.3
Japan	2.3	2.3	4.3	2.5	2.8	5.9
Germany	1.7	1.3	1.9	2.2	1.5	4.2
France	1.0	1.0	1.4	1.6	1.2	2.0
Italy	1.0	1.8	2.0	1.6	2.1	3.3
United Kingdom	1.0	0.5	2.0	1.0	0.9	2.7
Canada	0.8	-	2.0	1.8	-	2.7

Note: The simulated RMSEs and actual out-of-sample RMSEs for 2003Q1-2005Q1 are derived from Sédillot and Pain (2005, Table 4). The simulated RMSEs for Canada are derived from Mourougane (2006, Table 3) based on the availability of one month of GDP data. The outturn data for 2003-05 are the first available outturn estimate of GDP growth; the outturn data for 2007-12 are the first available published outturn data reported in following issues of the *OECD Economic Outlook*.

Source: Sédillot and Pain (2005); Mourougane (2006); and OECD calculations.

Box 2. Indicator models, composite leading indicators and OECD projections (cont.)**The usefulness of the OECD composite leading indicators since the crisis began**

A simple way of assessing whether the real-time CLIs contained information that might have helped to reduce the real-time projection errors is to regress the errors in GDP growth projections on the change in the CLIs that would have been available to forecasters at the time. One-month, 3-month, 6-month and 12-month changes up to the projection date in the CLIs for the G7 countries were considered for the period 2007-12 (see Appendix 7 of Pain et al., 2014).

Some caution is, of course needed in interpreting the results, but the following points emerge from this exercise:

- It does not appear that there was some information in the CLIs that could have been used to reduce the projection errors made in May 2008 as the downturn got underway.
- However, in the early stages of the recovery, at a time of considerable uncertainty, the errors in the November 2009 and May 2010 growth projections are found to be positively associated with the changes in the CLIs at the time of the forecast.

5.3. Enhanced monitoring of financial market developments and greater integration into the forecast process

Increasing attention is now paid to financial market developments in the construction of the projections and in empirical analysis.

- As the financial crisis progressed, the OECD Secretariat developed new financial conditions indices (FCIs) for the United States, Japan, the euro area and the United Kingdom (Guichard and Turner, 2008; Guichard et al., 2009). These weight together a wide range of financial variables that have a well-established link with GDP growth 12 to 18 months later.²⁴ The FCIs have been used extensively in the forecasting rounds since 2008-09 and as a guide to possible GDP effects in scenario analyses.
- A limitation of the aggregate FCIs is that they do not pick up all recent financial market developments, notably financial fragmentation in the euro area. Thus some organisations have also sought to directly incorporate country-specific information on bank lending rates and lending conditions in the set of variables being projected.
- In the OECD and elsewhere, discussions with internal financial market specialists and/or outside experts have been strengthened. In the OECD there are regular contacts between the Economics Department and the Directorate for Financial and Enterprise Affairs. Elsewhere, developments include the establishment of new divisions/units covering financial market developments and macro-financial linkages and considering key risks.
- Work is underway in some institutions on the difficult tasks of augmenting existing macroeconomic models with more detailed relationships of banking sector behaviour and strengthening linkages to reflect global financial interconnectedness. The macroeconomic models available at the time of the crisis typically ignored the banking system and failed to account for the possibility of bank capital shortages and credit rationing having an impact on macroeconomic developments. However, incorporating the financial sector into macroeconomic models is proving to be a major challenge.

5.4. Enhanced focus on risk assessments and global spillovers

Reflecting the uncertainty about the basic assumptions underlying the projections and the speed and depth of cross-country spillovers since the crisis began, there is now an enhanced focus on risk assessments and global spillovers in all international organisations:

- Greater information is now being provided about the distribution of risks around the central projections. At the OECD, while not providing a numerical risk distribution, the risk profile for the *Economic Outlook* projections has typically been characterised qualitatively as being balanced or skewed or bimodal and the projections themselves characterised as being modal rather than average projections.²⁵ Other international forecasting organisations have presented their assessed numerical risk distribution in the form of a fan chart or in the form of forecast ranges for key variables.
- Greater use is being made of quantitative scenario analyses to illustrate alternative outcomes and their global implications. Two notable examples in the OECD are the scenario used to illustrate a severe downside outcome arising from adverse euro area developments in the latter half of 2011 and the scenario illustrating the impact of the US debt ceiling becoming binding in the November 2013 *Economic Outlook*. The first of these scenarios made use of new forecasting tools, such as the financial conditions indices and models linking uncertainty and confidence with final expenditure, as well as a standard macroeconomic model, to calibrate the potential near-term activity impact of deteriorating financial conditions and a collapse in confidence brought about by a further worsening of the euro area crisis. The second used a global macroeconomic model to illustrate the extent to which a low-probability, but high-impact event would quickly spread to other countries.

Notes

1. The main analysis in this paper is based on calendar year growth rates. Figure 1 uses Q4/Q4 growth rates to illustrate the sudden deterioration in the financial crisis in the last quarter of 2008 and the impact this had on forecasts at the time.
2. See for example: OECD (2005, 2006), van den Noord (2006), Dieter et al. (2009) and Shigehara and Atkinson (2012).
3. Empirical analysis and macro-model simulations of an increasingly globalised financial system did, however, highlight the resulting potential for shocks in one economy to spill over to others to an ever greater extent (Hervé et al., 2007).
4. With the projection errors in each year tending to be in the same direction for almost all countries, the average absolute error across countries in Figure 2 is similar in magnitude to the average errors.
5. The projection errors are derived by subtracting the forecast from the outcome, so that an overestimate of growth results in a negative projection error.
6. This suggests the forecasts are not consistent with full information rational expectations.
7. This is based on the tests of the errors for G7 countries in Appendix 3 of Pain et al. (2014).
8. Changes in growth of less than 0.1 percentage points (from rounded data) are excluded.
9. The set of explanatory variables here is related to that in other studies that seek to explain cross-country differences in growth outcomes during the crisis, including Rose and Spiegel (2010), Cecchetti et al. (2011) and Berkmen et al. (2012). This paper focuses on a smaller set of countries (OECD member countries) and projection errors during the recovery as well as the downturn, and also considers some contemporaneous developments.
10. Multivariate regressions would better allow causal relationships to be established but are difficult in this small sample. Some preliminary results using the RMSEs of the country growth forecasts are presented in Table A.5.4 of Pain et al. (2014), although caveats still apply. The key finding is that for the November year-ahead projection errors, product market regulation remains correlated with accuracy after controlling for openness, with both variables significant at the 10% level.

11. The evidence points to the enhanced exposure of host economies to source economy shocks during the downturn outweighing the potential for foreign-owned banks, with access to internal capital resources, to help shield host economies from domestic shocks, which would limit negative growth surprises. A related finding is reported by Cetorelli and Goldberg (2012). Claessens and van Horen (2012) also show that foreign banks reduced their domestic credit more sharply than domestic banks during the crisis.
12. Related measures of business and labour market regulation produced by the World Bank and the Fraser Institute were found to have similar relationships (not shown here).
13. The negative relationship early in the crisis is more difficult to interpret since other studies have shown that countries with higher capital ratios outperformed those with lower ratios during that time (Cecchetti et al., 2011). One difference here is the focus on errors, rather than outcomes. Another is the use of regulatory capital (rather than total capital) to risk-weighted assets.
14. It is possible that the correlations could also reflect a reverse causal link, with weaker growth outcomes leading to more impaired loans and deteriorating confidence. However, qualitatively similar results are obtained if confidence changes available at the time of the forecasts are used instead.
15. Cecchetti et al. (2011) find that lower pre-crisis levels of private-sector credit to GDP and lower rates of growth in private-sector credit from 2005-07 were associated with better growth outcomes in the downturn (across a broader sample of countries). In contrast, there is evidence of a positive, but not significant, relationship between the errors in the OECD growth projections and pre-crisis level of private-sector bank credit.
16. The IMF studies focus primarily on European countries, although the main findings remain even if the sample is broadened to other advanced economies.
17. Actual fiscal consolidation was, however, strongly positively correlated with projected fiscal consolidation.
18. The average error and average absolute error of projections made in November 1973 for growth in 1974 were both larger than any projection error from the November year-ahead forecasts during the financial crisis.
19. Based on empirical estimates of equations 1 and 2 from Box 1.
20. The difference between the errors from the same-year May projections for growth in recent years compared with before the crisis is smaller in magnitude but is still statistically significant for three countries (France, Italy and the United Kingdom).
21. In 2008 and 2009, commodity prices were assumed to be fixed in the projections. In fact, they acted as an important automatic stabiliser for global activity, falling sharply as the crisis intensified and rebounding as the recovery began. If these developments had been incorporated in the OECD projections, then it is possible that the underestimation of the depth of the recession in 2008-09 and the extent of the bounce-back in 2009-10 would have been larger.
22. As part of the consultation it was agreed that individual comments and remarks would not be attributable to any individual institution. Additional relevant information is taken from Kenny and Morgan (2011) and González Cabanillas and Terzi (2012).
23. One example of the usefulness of such indicators was provided by the early signal of the extent of supply disruption in Japan following the earthquake and tsunami in March 2011. As noted in the May 2011 OECD *Economic Outlook*, an indicator of product availability compiled by the Billion Prices Project at MIT showed a fall of around 15% in online product availability in March.
24. The variables used include changes in credit availability, corporate bond spreads, short- and long-term interest rates, exchange rates and household wealth-income ratios.
25. Attempts have also been made to provide an indication of the uncertainty associated with the projections from the growth indicator models (Laurent and Kozluk, 2012; Rusticelli, 2012).

References

- Abreu, I. (2011), "International Organisations' vs. Private Analysts Forecasts: An Evaluation", *Bank of Portugal Working Papers*, No. 20/2011, Lisbon.
- Ahir, H., and P. Loungani (2014), "'There will be Growth in the Spring': How Well do Economists Predict Turning Points?", *Vox*, 14 April.

- Baker, S.R., N. Bloom and S.J. Davis (2012), "Has Economic Policy Uncertainty Hampered the Recovery?", *Becker Friedman Institute for Research In Economics Working Paper*, No. 2012-003.
- Berkmen, S.P., G. Gelos, R. Rennhack and J.P. Walsh (2012), "The Global Financial Crisis: Explaining Cross-country Differences in Output Impact", *Journal of International Money and Finance*, Vol. 31, No. 1, pp. 42-59.
- Bini Smaghi, L. (2010), "What has the Financial Crisis Taught Us? The Global Dimension and International Policy Cooperation", speech at the 21st Century Forum, Beijing, September.
- Blanchard, O. and D. Leigh (2013), "Growth Forecast Errors and Fiscal Multipliers", *International Monetary Fund Working Papers*, 13/1, Washington, DC.
- Cecchetti, S.G., M.R. King and J. Yetman (2011), "Weathering the Financial Crisis: Good Policy or Good Luck?", *BIS Working Papers*, No. 351, Basel.
- Cetorelli, N. and L. Goldberg (2011), "Global Banks and International Shock Transmission: Evidence from the Crisis", *IMF Economic Review*, Vol. 59/1.
- Claessens, S. and N. van Horen (2012), "Foreign Banks: Trends, Impact and Financial Stability", *International Monetary Fund Working Papers*, 12/10, Washington, DC.
- Dieter, H., L. Seabrooke and E. Tsingou (2009), "The Global Credit Crisis and the Politics of Financial Reform", *Garnet Policy Brief*, No. 9.
- Duval, R., J. Elmeskov and L. Vogel (2007), "Structural Policies and Economic Resilience to Shocks", *OECD Economics Department Working Papers*, No. 567, OECD Publishing, <http://dx.doi.org/10.1787/140152385131>.
- Duval, R. and L. Vogel (2008), "Oil Price Shocks, Rigidities and the Conduct of Monetary Policy: Some Lessons from a New Keynesian Perspective", *OECD Economics Department Working Papers*, No. 603, OECD Publishing, <http://dx.doi.org/10.1787/242188534400>.
- European Commission (2012), "Box 1.5: Forecast Errors and Multiplier Uncertainty", *European Economic Forecast – Autumn 2012, European Economy*, No. 7/2012, Brussels.
- Fair, R.C. and R.J. Shiller (1990), "Comparing Information in Forecasts from Econometric Models", *American Economic Review*, Vol. 80/3.
- Gill, T., D. Perera and D. Sunner (2012), "Electronic Indicators of Economic Activity", *Reserve Bank of Australia Bulletin*, June 2012, pp. 1-12.
- González Cabanillas, L. and A. Terzi, "The Accuracy of the European Commission's Forecasts Re-examined", *European Economy. Economic Papers*, No. 476, Brussels.
- Guichard, S., D. Haugh and D. Turner (2009), "Quantifying the Effect of Financial Conditions in the Euro Area, Japan, United Kingdom and United States", *OECD Economics Department Working Papers*, No. 677, OECD Publishing, <http://dx.doi.org/10.1787/226365806132>.
- Guichard, S. and E. Rusticelli (2011), "A Dynamic Factor Model for World Trade Growth", *OECD Economics Department Working Papers*, No. 874, OECD Publishing, <http://dx.doi.org/10.1787/5kg9zbvuwq2-en>.
- Guichard, S. and D. Turner (2008), "Quantifying the Effect of Financial Conditions on US Activity", *OECD Economics Department Working Papers*, No. 635, OECD Publishing, <http://dx.doi.org/10.1787/236860073636?site=fr>.
- Hellerstein, R. and M. Middelorp (2012), "Forecasting with Internet Search Data", *Liberty Street Economics*, Federal Reserve Bank of New York.
- Hervé, K., N. Pain and F. Sédillot (2007), "Globalisation and the Macroeconomic Policy Environment", *OECD Economics Department Working Papers*, No. 552, OECD Publishing, <http://dx.doi.org/10.1787/185006265382?site=fr>.
- IMF (2012), "Box 1.1: Are We Under-estimating Short-term Fiscal Multipliers?", *World Economic Outlook*, October, pp. 41-43, Washington, DC.
- Kenny, G. and J. Morgan (2011), "Some Lessons from the Financial Crisis for Economic Analysis", *ECB Occasional Paper*, No. 130, Frankfurt.
- Laurent, T. and T. Kozluk (2012), "Measuring GDP Forecast Uncertainty Using Quantile Regressions", *OECD Economics Department Working Papers*, No. 978, OECD Publishing, <http://dx.doi.org/10.1787/5k95xd76juvg-en?site=fr>.

- Loungani, P. (2001), "How Accurate Are Private Sector Forecasts? Cross-country Evidence From Consensus Forecasts of Output Growth", *International Journal of Forecasting*, Vol. 17, No. 3, pp. 419-32.
- McLaren, N. and R. Shanbhogue (2011), "Using Internet Search Data as Economic Indicators", *Bank of England Quarterly Bulletin*, Q2 2011, pp. 134-40.
- Mourougane, A. (2006), "Forecasting Monthly GDP for Canada", *OECD Economics Department Working Papers*, No. 515, OECD Publishing, <http://dx.doi.org/10.1787/421416670553>.
- OECD (2005), "OECD Economic Outlook No. 78", *OECD Economic Outlook: Statistics and Projections* (database), OECD Publishing, <http://dx.doi.org/10.1787/data-00088-en>.
- OECD (2006), "OECD Economic Outlook No. 79", *OECD Economic Outlook: Statistics and Projections* (database), OECD Publishing, <http://dx.doi.org/10.1787/data-00087-en>.
- Pain, N., C. Lewis, T.-T. Dang, Y. Jin and P. Richardson (2014), "OECD Forecasts During and After the Financial Crisis: A Post Mortem", *OECD Economics Department Working Papers*, No. 1107, OECD Publishing, <http://dx.doi.org/10.1787/5jz7311qw1s1-en>.
- Rose, A.K. and M.M. Spiegel (2010), "Cross-country Causes and Consequences of the Crisis: An Update", *NBER Working Papers*, No. 16243.
- Rusticelli, E. (2012), "Non-parametric Stochastic Simulations to Investigate Uncertainty around the OECD Indicator Model Forecasts", *OECD Economics Department Working Papers*, No. 979, OECD Publishing, <http://dx.doi.org/10.1787/5k94kq50b2jd-en>.
- Sédillot, F. and N. Pain (2003), "Indicator Models of Real GDP Growth in Selected OECD Countries", *OECD Economics Department Working Papers*, No. 364, OECD Publishing, <http://dx.doi.org/10.1787/275257320252>.
- Sédillot, F. and N. Pain (2005), "Indicator Models of Real GDP Growth in the Major OECD Countries", *OECD Economic Studies*, No. 40, OECD Publishing, www.coe-rexecode.fr/content/download/26371/270575/version/1/file/OECD_Paper_20060614.pdf.
- Shrestha, M. and M. Marini (2013), "Quarterly GDP Revisions in G-20 Countries: Evidence from the 2008 Financial Crisis", *International Monetary Fund Working Papers*, No. 13/60, Washington, DC.
- Shigehara, K. and P. Atkinson (2011), "Surveillance by International Institutions: Lessons from the Global Financial and Economic Crisis", *OECD Economics Department Working Papers*, No. 860, OECD Publishing, <http://dx.doi.org/10.1787/5kgchzchkd2-en>.
- van den Noord, P. (2006), "Are House Prices Nearing a Peak? A Probit Analysis for 17 OECD Countries", *OECD Economics Department Working Papers*, No. 488, OECD Publishing, <http://dx.doi.org/10.1787/816770752416>.
- Vogel, L. (2007), "How Do the OECD Growth Projections for the G7 Economies Perform?: A Post-Mortem", *OECD Economics Department Working Papers*, No. 573, OECD Publishing, <http://dx.doi.org/10.1787/111804483765>.
- Zarnowitz, V. (1991), "Has Macro-Forecasting Failed?", *NBER Working Papers*, No. 3867.

APPENDIX A

Summary of projection errors

This appendix summarises the errors in the GDP growth projections for each country and for groups of countries. As in the main text, the key statistics are:

- The average error, which is positive or negative, depending on whether growth was under- or over-predicted on average during the period. This statistic is also a measure of bias.
- The average absolute error and root mean squared error (RMSE), which are indicators of accuracy. They can be scaled to facilitate comparisons across countries or time – to do this the average absolute error is scaled by average absolute growth during the same period and the RMSE is scaled by the volatility of growth in that period (the standard deviation).

Errors in annual GDP growth projections

In the 2007-2012 period, the average errors in the projections for GDP growth in the following year were negative for almost all countries – i.e. growth was weaker than projected. In contrast, the errors in the May projections for growth in the same year were, on average, close to zero, and negative in less than half of the OECD economies (see Table A.1). There is a wide range of outcomes around these means. The largest over-predictions of growth amongst those countries where projections were made throughout 2007-2012 were in Greece and Ireland amongst the OECD economies and in Russia for the BRIICS.

The average errors mask offsetting positive and negative errors in the projections. As a result, the average absolute errors or RMSEs (Tables A.2 and A.3) are larger than the average errors. Average absolute errors are typically around 1 percentage point for the May current year projections, but rise as the forecast horizon is lengthened to over 2 percentage points for most countries. The RMSEs present a similar picture. By both metrics, OECD countries outside of Europe had lower errors in the projections made for growth in the year ahead.

Scaling these measures growth outcomes provides a more standardised basis with which to compare accuracy across countries. When scaled by absolute GDP growth outcomes, the average absolute error in the current year projections is around 20-30% of the typical absolute growth rate of GDP (Table A.2, column 4). In contrast, for the average OECD economy, the average absolute error in the May projections for growth in the following year is approximately the same size as the absolute growth rate (Table A.2, column 6). On the basis of the RMSEs relative to the volatility of growth, the projections for the OECD countries were typically as accurate as those for the BRIICS (Table A.3, columns 4 to 6). Projection errors are, on average, around one-half of the volatility of GDP growth in the May current year projections, but broadly equal to the standard deviation of growth in the May projections for growth in the following year.

Table A.1. **Average errors of GDP growth projections for 2007-12**

Percentage points

	May projection for current year	November projection for next year	May projection for next year
OECD countries			
Australia	0.2	-0.3	-0.5
Austria	0.2	-0.1	-0.5
Belgium	0.0	-0.4	-0.9
Canada	-0.4	-0.5	-1.2
Chile ¹	0.5	0.7	0.5
Czech Republic	-0.3	-1.8	-2.4
Denmark	-0.8	-1.5	-1.8
Estonia ¹	1.4	-1.0	1.8
Finland	-0.8	-1.4	-1.8
France	-0.4	-0.6	-1.2
Germany	0.1	-0.2	-0.6
Greece	-1.2	-2.6	-3.7
Hungary	-0.6	-1.4	-3.0
Iceland	0.3	0.7	-1.2
Ireland	-0.3	-1.9	-3.5
Israel ¹	0.7	0.5	1.2
Italy	-0.5	-1.5	-2.1
Japan	0.0	-1.3	-1.5
Korea	-0.2	-0.9	-1.4
Luxembourg	-0.8	-1.5	-2.4
Mexico	0.1	-0.9	-1.6
Netherlands	0.1	-0.7	-1.2
New Zealand	0.4	-0.2	-1.0
Norway	-0.2	-0.7	-0.5
Poland	0.2	0.1	0.0
Portugal	0.2	-0.8	-1.4
Slovak Republic	0.1	-0.8	-1.6
Slovenia ¹	-1.2	-4.1	-2.3
Spain	0.1	-0.9	-1.3
Sweden	0.0	-1.0	-1.3
Switzerland	0.3	0.1	-0.2
Turkey	0.0	-0.7	-1.5
United Kingdom	-0.4	-1.0	-1.7
United States	-0.2	-0.4	-1.0
<i>OECD aggregate projection</i>	<i>-0.1</i>	<i>-0.7</i>	<i>-1.2</i>
Non-OECD countries			
Brazil	-0.1	-0.4	-0.9
China	0.1	-0.1	-0.6
India	-0.7	-0.7	-0.9
Indonesia ¹	0.4	0.1	-0.2
Russia	-0.8	-1.7	-2.3
South Africa ¹	-0.7	-0.7	-1.8
Country group averages²			
OECD	-0.1	-0.9	-1.4
Euro area	-0.2	-1.1	-1.5
Euro area: Core	-0.1	-0.9	-1.1
Euro area: Vulnerable	-0.3	-1.5	-2.4
Other OECD Europe	-0.1	-0.7	-1.4
OECD outside Europe	0.1	-0.4	-0.8
BRIICS	-0.2	-0.6	-1.0

Note: Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outcome as at May the following year. A negative (positive) average error indicates over(under)-prediction. All projections are calendar year except India, which is fiscal year. "OECD aggregate projection" refers to the published projection for the OECD area.

1. Projections only available for part of the period 2007-12.

2. Countries included in the "euro area: vulnerable group" are Greece, Ireland, Italy, Portugal and Spain.

Source: OECD, *Economic Outlook Databases*; and OECD calculations.

Table A.2. **Average absolute errors of GDP growth projections for 2007-12**

	Average absolute error (percentage points)			Ratio of average absolute error to average absolute growth		
	May projection for current year	November projection for next year	May projection for next year	May projection for current year	November projection for next year	May projection for next year
OECD countries						
Australia	0.8	0.8	1.1	0.3	0.3	0.4
Austria	0.4	1.2	1.9	0.2	0.5	0.8
Belgium	0.7	1.1	1.9	0.4	0.6	1.0
Canada	0.4	0.9	2.0	0.2	0.4	0.9
Chile	0.9	0.9	0.5	0.2	0.2	0.1
Czech Republic	0.7	2.4	3.3	0.2	0.8	1.1
Denmark	1.1	1.8	2.4	0.6	0.9	1.3
Estonia	1.5	5.1	2.7	0.2	0.7	0.6
Finland	1.6	2.8	3.2	0.5	0.9	1.0
France	0.6	0.7	1.6	0.4	0.5	1.2
Germany	0.8	1.4	2.4	0.3	0.5	0.9
Greece	1.2	2.7	3.8	0.3	0.6	0.9
Hungary	0.8	2.1	4.0	0.4	1.0	2.0
Iceland	1.2	1.7	2.3	0.4	0.5	0.7
Ireland	1.5	2.3	3.7	0.6	0.9	1.4
Israel	1.0	1.2	2.2	0.3	0.4	0.5
Italy	0.7	1.5	2.4	0.3	0.8	1.3
Japan	0.9	2.0	2.6	0.4	0.8	1.1
Korea	1.3	1.7	2.3	0.4	0.5	0.7
Luxembourg	1.3	2.0	3.8	0.5	0.8	1.6
Mexico	0.8	2.2	2.6	0.2	0.5	0.6
Netherlands	0.6	1.2	2.2	0.3	0.5	1.0
New Zealand	1.4	1.4	2.1	0.7	0.7	1.0
Norway	0.6	1.1	1.0	0.3	0.5	0.5
Poland	0.9	1.0	1.9	0.2	0.2	0.5
Portugal	0.7	1.1	2.1	0.4	0.6	1.2
Slovak Republic	0.7	2.4	3.3	0.1	0.5	0.6
Slovenia	1.2	4.1	2.7	0.4	1.4	2.2
Spain	0.3	1.1	1.8	0.1	0.6	1.0
Sweden	1.5	2.3	3.2	0.5	0.7	1.0
Switzerland	0.7	0.8	1.6	0.4	0.4	0.8
Turkey	1.7	3.5	5.0	0.3	0.7	1.0
United Kingdom	0.5	1.1	2.1	0.3	0.6	1.2
United States	0.3	0.6	1.7	0.1	0.3	0.8
<i>OECD aggregate projection</i>	<i>0.4</i>	<i>1.0</i>	<i>1.9</i>	<i>0.2</i>	<i>0.5</i>	<i>0.9</i>
Non-OECD countries						
Brazil	1.1	2.0	2.7	0.3	0.6	0.7
China	0.8	0.9	1.2	0.1	0.1	0.1
India	1.4	1.9	2.2	0.2	0.3	0.3
Indonesia	0.4	0.5	1.3	0.1	0.1	0.2
Russia	1.3	2.5	3.2	0.2	0.4	0.6
South Africa	0.7	0.8	1.8	0.2	0.3	0.6
Country group averages						
OECD	0.9	1.7	2.5	0.3	0.6	1.0
Euro area	0.9	1.8	2.4	0.3	0.6	1.1
Euro area: Core	0.8	1.9	2.3	0.3	0.6	1.0
Euro area: Vulnerable	0.9	1.8	2.8	0.3	0.7	1.1
Other OECD Europe	1.0	1.8	2.7	0.4	0.7	1.0
OECD outside Europe	0.8	1.2	1.7	0.3	0.5	0.8
BRIICS	1.0	1.6	2.3	0.2	0.3	0.4

Note: Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outturn as at May the following year. All projections are calendar year except India, which is fiscal year. Projections are only available for part of the period 2007-12 for Chile, Estonia, Israel, Slovenia, Indonesia and South Africa; these are excluded from country group averages of the ratio of average absolute error to average absolute growth. Countries included in the "euro area: vulnerable" group are Greece, Ireland, Italy, Portugal and Spain.

Source: OECD, *Economic Outlook Databases*; and OECD calculations.

Table A.3. **RMSE of GDP growth projections for 2007-12**

	RMSE (percentage points)			Ratio of RMSE to standard deviation of growth outcomes		
	May projection for current year	November projection for next year	May projection for next year	May projection for current year	November projection for next year	May projection for next year
OECD countries						
Australia	0.9	0.9	1.2	0.9	0.9	1.1
Austria	0.5	1.5	2.4	0.2	0.6	1.0
Belgium	0.7	1.4	2.5	0.3	0.7	1.1
Canada	0.5	1.3	2.5	0.2	0.6	1.1
Chile	0.9	1.1	0.5	2.3	2.6	1.6
Czech Republic	0.9	3.1	4.3	0.2	0.9	1.2
Denmark	1.3	2.3	2.9	0.5	0.9	1.1
Estonia	1.8	6.7	2.9	0.2	0.7	1.1
Finland	1.8	3.8	4.5	0.4	0.9	1.0
France	0.7	1.0	2.0	0.4	0.6	1.2
Germany	0.9	2.0	3.0	0.3	0.6	1.0
Greece	1.8	3.1	4.5	0.4	0.7	1.0
Hungary	0.9	2.6	4.7	0.3	0.9	1.6
Iceland	1.5	1.9	3.0	0.4	0.5	0.7
Ireland	2.0	3.1	5.1	0.5	0.8	1.4
Israel	1.4	1.4	2.7	0.8	0.8	3.2
Italy	0.8	2.1	3.2	0.3	0.8	1.3
Japan	1.3	2.6	3.4	0.4	0.8	1.1
Korea	1.5	1.9	2.7	0.7	0.9	1.3
Luxembourg	1.8	2.8	4.5	0.6	1.0	1.5
Mexico	1.0	3.2	4.3	0.2	0.7	1.0
Netherlands	0.7	1.7	2.9	0.2	0.6	1.1
New Zealand	1.7	1.8	2.2	0.8	0.9	1.1
Norway	0.7	1.4	1.5	0.4	0.8	0.8
Poland	1.1	1.0	2.2	0.6	0.6	1.2
Portugal	1.0	1.4	2.5	0.5	0.7	1.2
Slovak Republic	0.9	3.8	4.8	0.2	0.8	1.0
Slovenia	1.5	5.4	3.2	0.4	1.3	1.8
Spain	0.3	1.4	2.4	0.1	0.6	1.0
Sweden	2.0	2.9	4.0	0.5	0.8	1.1
Switzerland	0.8	1.0	1.8	0.5	0.6	1.2
Turkey	1.8	4.1	5.6	0.3	0.8	1.1
United Kingdom	0.6	1.7	2.9	0.2	0.6	1.1
United States	0.4	0.8	1.9	0.2	0.4	1.0
<i>OECD aggregate projection</i>	<i>0.5</i>	<i>1.4</i>	<i>2.5</i>	<i>0.2</i>	<i>0.6</i>	<i>1.1</i>
Non-OECD countries						
Brazil	1.3	2.2	3.0	0.4	0.7	1.0
China	0.9	1.1	1.4	0.6	0.7	1.0
India	1.8	2.1	2.5	0.9	1.0	1.2
Indonesia	0.6	0.6	1.5	0.7	0.7	7.2
Russia	1.4	4.3	6.0	0.2	0.8	1.1
South Africa	0.7	0.9	1.8	2.3	3.1	4.4
Country group averages						
OECD	1.2	2.6	3.4	0.4	0.7	1.1
Euro area	1.2	2.9	3.4	0.4	0.7	1.1
Euro area: Core	1.2	3.1	3.3	0.3	0.7	1.1
Euro area: Vulnerable	1.3	2.4	3.7	0.4	0.7	1.1
Other OECD Europe	1.2	2.4	3.5	0.4	0.7	1.1
OECD outside Europe	1.1	1.8	2.5	0.5	0.7	1.1
BRIICS	1.2	2.4	3.6	0.5	0.8	1.1

Note: Errors are calculated as actual growth less projected growth at each forecast horizon, where actual growth is the published outturn as at May the following year. All projections are calendar year except India, which is fiscal year. Projections are only available for part of the period 2007-12 for Chile, Estonia, Israel, Slovenia, Indonesia and South Africa; these are excluded from country group averages of the ratio of the RMSE to the standard error of growth. Countries included in the "euro area: vulnerable" group are Greece, Ireland, Italy, Portugal and Spain.

Source: OECD, *Economic Outlook Databases*; and OECD calculations.

The effect of the global financial crisis on OECD potential output

by

Patrice Ollivaud and David Turner*

Potential output losses from the global financial crisis are estimated by comparing recent OECD published projections with a counter-factual assuming a continuation of pre-crisis productivity trends and a trend employment rate which is sensitive to demographic trends. Among the 19 OECD countries which experienced a banking crisis over the period 2007-11 the median loss in potential output in 2014 is estimated to be about 5½ per cent, compared with a loss in aggregate potential output across all OECD countries of about 3½ per cent. The loss does, however, vary widely across countries, being more than 10% for several smaller European, mainly euro area, countries. The largest adverse effects come from lower trend productivity, which is a combination of both lower total factor productivity and lower capital per worker. Despite large increases in structural unemployment in some countries, the contribution of lower potential employment is limited because the adverse effect on labour force participation is generally much less than might have been expected on the basis of previous severe downturns. This may partly reflect pension reforms and a tightening up of early retirement pathways. Pre-crisis conditions relating to over-heating and financial excesses, including high inflation, high investment, large current account deficits, high total economy indebtedness and more rapid growth in capital-per-worker are all correlated with larger post-crisis potential output losses. This suggests that underlying the potential output losses was a substantial misallocation of resources, especially of capital, in the pre-crisis boom period. On the other hand, more competition-friendly product market regulation is associated with smaller losses of potential output, suggesting that it facilitates a reallocation of resources across firms and sectors in the aftermath of an adverse shock and so helps to mitigate its consequences.

JEL classification: E32; E44.

Keywords: Banking crisis, financial crisis, global financial crisis, potential output.

* The authors are economist and Head of Division, respectively, from the OECD Macroeconomic Analysis Division of the Economics Department (patrice.ollivaud@oecd.org; david.turner@oecd.org). A previous version of this paper, based on earlier vintage of data and projections, was prepared for the 2014 Bank of Korea International Conference “Strengthening Growth Potential in the Aftermath of the Global Financial Crisis”, June 2-3, 2014, Bank of Korea, Seoul, Korea. The authors thank Jonathan Ostry, Enrique Alberola, Peter Hoeller, Boris Cournede, Davide Furceri, Yvan Guillemette, Christine Lewis, Phil Bagnoli and Mark Pearson for valuable comments and discussions as well as Ines Gomez Palacio for help with documentation preparation. The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

1. Introduction and summary

The global financial crisis has been extremely costly in terms of lost output and jobs. A crucial judgement for macroeconomic policy is how much of this loss will persist once cyclical recovery is complete. Specifically, what level of loss in medium-term output will remain relative to the trajectory that might have been expected in the absence of the crisis.¹ The extent of such losses is a particularly important consideration for fiscal policy given the large increase in government deficits and debt experienced by many OECD countries; the greater the loss, the less that fiscal balances will improve with any recovery and so the greater the need for discretionary action to correct remaining fiscal imbalances. The extent of the loss in capacity to produce output without generating unsustainable macroeconomic imbalances is also relevant for monetary policy; the greater the loss, the sooner any recovery will run into capacity constraints and so the sooner the need to normalise policy rates and end unconventional monetary easing. An understanding of which countries have been most severely impacted by the crisis may also cast light on which pre-crisis conditions and policies make economies vulnerable to such shocks and so suggest ways of improving resilience in the future.

Evaluating the effect of the crisis on medium-term output prospects is inherently problematic, given the impossibility of knowing what the counter-factual is and because of the difficulties of disentangling the effect of the crisis from other effects, including policy changes. Moreover, even seven years after the crisis, with many economies judged to be operating well below capacity, any assessment can only be preliminary. The approach in the current paper is based on OECD published estimates of potential output (OECD, 2014a), which is that estimated trend level of output which can be sustained without generating inflationary pressures. Losses related to the crisis are evaluated on a country-by-country basis relative to a counter-factual assuming a continuation of pre-crisis labour productivity trends, but with alternative assumptions regarding the different components of the trend employment rate. There is only one country (Germany) out of 19 OECD countries which experienced a banking crisis over the period 2007-11, which did not incur a loss in potential output on this basis. Nevertheless, this approach suggests (unsurprisingly) that other factors, including policy changes, have had an influence on potential output, particularly where post-crisis developments in some components of potential output for some countries are positive relative to the counter-factual. However, a strength of the approach is that in such cases it allows an explicit adjustment to be made when there are good reasons to suspect other influences on potential output are important.

The main findings of the paper are as follows:

- Among the 19 OECD countries which experienced a banking crisis over the period 2007-11, the median loss in potential output in 2014 is estimated to be about 5½ per cent, which is slightly higher than OECD estimates of previous major banking crises.
- The detail of the analysis, by country and component of potential output, confirms that there are important effects other than just the crisis, including policy effects, acting on

potential output. Adjusting the estimates, to exclude such effects where they seem most obvious, suggests that the loss to aggregate OECD potential output is about 3½ per cent instead of 2% when not adjusting for such effects. Although there is little effect of these adjustments on the median estimate for countries which experienced a banking crisis, the median loss for other OECD countries is doubled to 4%.

- These effects are smaller than other widely-quoted estimates in the literature relating to the medium-term output effects of past financial crises. It is argued that this is because other studies tend to over-estimate the pre-crisis trend growth rate as a basis for the counter-factual against which output losses are evaluated.
- The estimated effects of the current crisis in this paper do, however, vary widely across countries. The most severely affected are overwhelmingly smaller European, mainly euro area, countries; the reduction in potential output in 2014 being more than 10% for the Czech Republic, Estonia, Greece, Hungary, Iceland, Ireland, Portugal and Slovenia. For Germany and Japan, the crisis is not estimated to have had any adverse effect on potential output.
- The largest adverse effects come from lower productivity, which is a combination of both lower total factor productivity and lower capital per worker. The decline in capital per worker accounts for more than 3 percentage points of the decline in potential output in 2014 for Estonia, Iceland, Hungary, Korea and Portugal, with these countries also being among those in which the pre-crisis growth of capital per worker was most rapid.
- The contribution of lower potential employment to the loss in potential is typically much smaller than from productivity, although this contribution is about 5 percentage points or more for Iceland, Ireland, Greece, Portugal, and Spain, with most of this explained by higher structural unemployment. Adverse effects on labour force participation are modest and generally much less than might have been expected on the basis of previous severe downturns, which may partly reflect pension reforms and a tightening up of early retirement pathways. Within Europe, changes in net migration flows can also explain much of the estimated effect of the crisis on movements in trend participation. Conversely, the United States is one country where there has been a marked decline in trend participation, but this is attributed mostly to demographic effects and not the effect of the crisis.
- In terms of pre-crisis conditions, both higher inflation and positive output gaps (a measure of over-heating), are correlated with subsequent potential output losses. In addition, pre-crisis conditions relating to financial excesses, namely high investment, larger current account deficits and high total economy indebtedness are also correlated with larger potential output losses. This provides support to arguments that macroeconomic policy should be concerned with financial developments over and above their immediate implications for inflation. Additionally, faster growth in the capital-per-worker component of potential growth is correlated with post-crisis potential output losses. These results therefore suggest that underlying the potential output losses was a substantial misallocation of resources, especially of capital, in the pre-crisis boom period, with rapid growth in capital contributing to unsustainable pre-crisis productivity growth.
- More competition-friendly product market regulation is associated with smaller crisis-related losses of potential output. This suggests that more flexible product markets

facilitate a reallocation of resources across firms and sectors in the aftermath of an adverse shock and so helps to mitigate its consequences.

The remainder of the paper is organised as follows: Section 2 reviews multi-country studies which have been used to evaluate the medium-term output costs of previous financial crises; Section 3 describes the method used in the current study and presents the results of the analysis; Section 4 examines which pre-crisis macro and structural conditions and policies are correlated with the potential output loss.

2. A selective review of recent studies of the medium-term output losses following financial crises

While a number of studies have examined the immediate consequences of past financial crises, there are fewer that have focused on the medium-term losses to output. The focus here is on three such studies (Cerra and Saxena, 2008; Furceri and Mourougane, 2009; and IMF, 2009), because differences between them are helpful in motivating the approach in the current paper as well as explaining differences in the results. While these studies are all relatively recent, they do not include an assessment of the effect of the current financial crisis. However, a more recent study by Ball (2014) is also considered because it does focus on the recent financial crisis and is also similar to the current paper in using OECD estimates of potential output to evaluate the medium-term output losses from the crisis.

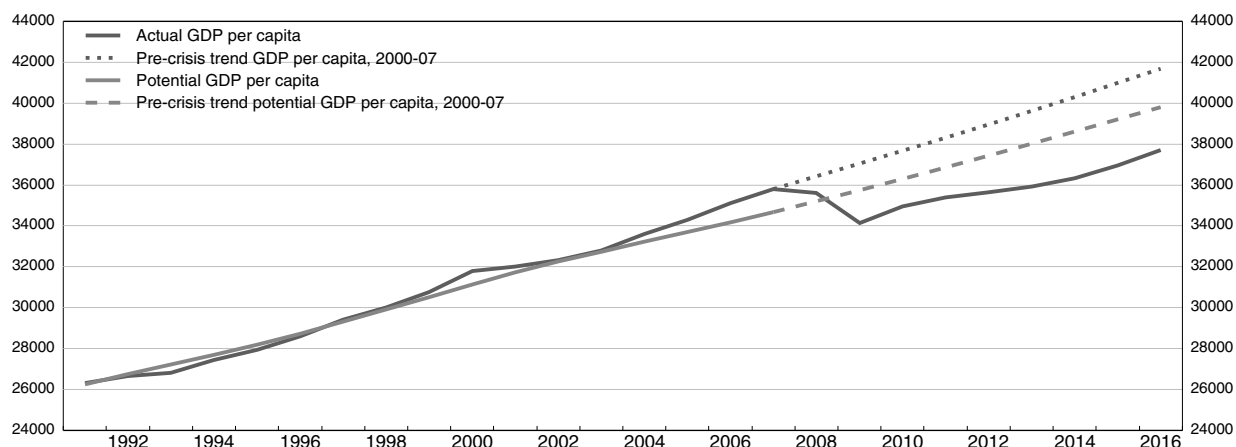
The papers by Cerra and Saxena (2008) and Furceri and Mourougane (2009), hereafter referred to as CS and FM, form a natural comparison pair because both adopt a similar methodology, namely estimating an autoregressive panel model to explain GDP growth rates augmented by crisis dummies. The average effect of a crisis is then evaluated by calculating the impulse response function from the estimated coefficients. There are, however, important differences between the two papers, most notably CS consider a much broader coverage of developing and developed countries and separately distinguish the effect of banking, currency and political crises, whereas FM focus on banking and currency crises and only consider OECD countries. In order to have a better basis for comparison across the two studies and because of the greater relevance to the estimates of the effect of the current crisis, the focus here is only on a comparison of the effects of banking crises in high income countries in CS and the effect of the so-called “Big Five” banking crises (Reinhart and Rogoff, 2008) in FM.² The results from the two studies are strikingly different even for what should be a similar set of crises, with the adverse effects more than 3½ times larger in CS compared with MF; thus after ten years, CS estimate that banking crises in high-income countries reduce output by 15%,³ whereas FM estimate a reduction of nearly 4% for the “Big Five” crises.⁴

The large difference in these estimates is mostly due to the different dependent variables used in each of the studies; whereas CS use actual observed GDP growth rates, FM use potential output growth rates. Using actual GDP growth rates is likely to bias upwards implicit estimates of both the pre-crisis growth rate and the pre-crisis level of GDP which is sustainable, and hence lead to an exaggerated estimate of the post-crisis loss in output. This is because prior to a financial crisis, there is typically an extended period of above average growth, usually associated with a number of symptoms of over-heating.

To illustrate the difference which using actual and potential GDP can make in establishing a counter-factual trend and hence to calculations of the medium-term output

loss following a crisis, the example of OECD aggregate GDP over the recent crisis is considered. Assuming the start of the crisis to be 2008 (the bankruptcy of Lehman Brothers), for the purposes of this illustration a pre-crisis OECD trend growth rate is calculated as the average growth rate over the period 2000-07 of alternatively *actual* or *potential* OECD GDP per capita (Figure 1). Alternative post-crisis counter-factuals are then calculated by extending either actual or potential OECD GDP per capita from 2007 according to their respective pre-crisis growth rates, where potential output estimates are taken from recent *OECD Economic Outlook* projections (OECD, 2014a). The counter-factual based on actual GDP per capita is progressively higher than that based on potential output; the difference in 2007 is initially about 3%, and after ten years (the horizon used by both CS and FM) the difference is about 5%. This difference in the counter-factual trend is likely to translate into a corresponding difference in estimates of the medium-term output loss.

Figure 1. **OECD GDP per capita compared with various pre-crisis trends**
In 2010 PPP US dollars



Note: OECD GDP per capita over the period 2014-16 is based on November 2014 *OECD Economic Outlook* projections.

Source: OECD calculations based on *OECD Economic Outlook*, November 2014 long-term database.

The approach used in IMF (2009) is to calculate medium-term output losses relative to a pre-crisis trend calculated using actual GDP, although it is acknowledged that estimating the pre-crisis trend is “tricky” particularly because of the problem of “insulating the analysis from the impact of any immediate pre-crisis boom or slump”. The method adopted by the IMF is to estimate a linear trend through the actual GDP series during a seven year pre-crisis period that ends three years before the onset of the crisis. While this approach has the merit of being transparent, it will not necessarily insulate the pre-crisis trend from any immediate pre-crisis boom if there is a protracted run-up prior to the financial crisis as typically seems to be the case (Borio, 2012). For example, applying the IMF methodology to the previous example of OECD GDP over the period of the current crisis (so calculating the pre-crisis trend on actual GDP per capita over the period 1998-2005), leads to an even *higher* counter-factual trajectory for output (than using the trend in actual GDP per capita over 2000-07) and so an even larger medium-term output loss.

The use of actual GDP to establish pre-crisis trends probably explains why the IMF findings of the medium-term output loss, following a financial crisis, are closer to those of

CS than FM and probably over-estimate the hit; for high-income countries the average reduction in GDP after seven years is about 10%.⁵ Applying the IMF methodology to the current example of OECD aggregate GDP over the current financial crisis, and treating the most recent OECD projections as being realised, would give a similar result in 2015 (i.e. seven years after the crisis) with a loss in medium-term output of about 10%. However, this estimate is likely to be an exaggeration of the permanent output loss, not only because the pre-crisis trend in output may be overstated, but also because many countries are still currently operating well below capacity. Indeed, while it might be reasonable to assume that seven years after a typical cyclical downturn GDP has returned to its medium-term trend, following a financial crisis the downturn is likely to be more protracted (see, for example, Haugh et al., 2009).

The justification for using potential rather than actual GDP to estimate medium-term output losses in the remainder of the paper is twofold. First, because of the difficulties of establishing a pre-crisis trend and second, because GDP in many OECD countries does not appear to have returned to trend and is not expected to do so in the next year or so.⁶ In addition, as explained in more detail in the next section, the counter-factual used in the paper is not as crude as a simple extrapolation of the pre-crisis trend in potential output per capita (as used in Figure 1), but rather adopts different assumptions regarding different components of potential output. Using potential output does raise issues about the methodology of constructing potential output (described in Appendix A), but these estimates do have the merit of following the same procedure across all countries and are subject to checking from country specialists.

In this paper, one important change which is made relative to the approach of both CS and FM, is that in common with the approach in IMF (2009), the effect of the crisis is evaluated on GDP per capita rather than just GDP. This is a simple and intuitive improvement in modelling growth given that changes in population growth are likely to imply changes in trend GDP growth of a similar order of magnitude, independently of the incidence of crises. Not taking into account such effects risks attributing the effect of a slowdown in population growth to the effect of a crisis.⁷

An alternative approach to estimate the crisis hit is proposed by Ball (2014) who compares recent estimates of potential output from the OECD and IMF with estimates which were made in 2007 in order to compute the damage the crisis has done to potential output (see Box 4.3 in OECD, 2010a for a similar approach). Ball calculates an average loss, weighted by economy size, across OECD countries of 8.4%. A problem, however, with comparing different vintages of potential output projections, is that both the OECD and IMF have tended to revise down estimates of both the level and growth rate of pre-crisis potential output (so pre-crisis output gaps are now more positive), implicitly recognising that there was more over-heating going on than was recognised at the time. This in turn means that the adverse effects from the crisis on potential output are likely to be exaggerated by comparing different vintages of projections.⁸

3. Analysing the effect of the current crisis on potential output

3.1. Methodology

To derive the counter-factual post-crisis trajectory for potential output per capita, the pre-crisis growth rate in trend productivity is extrapolated, but each of the components of the potential employment rate are treated differently, as explained below.

The current OECD method of estimating potential output, see Appendix 1 and Johansson et al. (2013) for details, assumes a Cobb-Douglas production function, which can be simplified so that potential output (Y^*) is represented in terms of potential employment (N^*), the capital stock (K) and labour-augmenting technical progress (E^*), so that:

$$y^* = \alpha (n^* + e^*) + (1 - \alpha) k, \quad [1]$$

where lower case letters denote logs and α is the wage share.

Growth in potential output per head of population (P) can be explained in terms of two components: trend labour productivity and a potential employment rate, as follows:

$$\Delta(y^* - p) = \Delta(y^* - n^*) + \Delta(n^* - p). \quad [2]$$

For the purposes of the post-crisis counter-factual, the trend labour productivity component is extrapolated at the same average growth rate as over the pre-crisis period 2000-07.⁹ This trend productivity component can also be split into two components (each of which are assumed to follow their pre-crisis growth rate), which can be derived from combining [1] and [2], to give an effect from labour efficiency (or equivalently an effect from total factor productivity) and an effect from changes in capital per worker, represented as:

$$\Delta(y^* - n^*) = \alpha \Delta e^* + (1 - \alpha) \Delta(k - n^*). \quad [3]$$

A different assumption is adopted for the potential employment component (the second term in [2]), which can itself be split into three parts:

$$\Delta(n^* - p) = \Delta(n^* - lf^*) + \Delta(lf^* - pwa) + \Delta(pwa - p), \quad [4]$$

where lf^* is the trend labour force and PWA is the population of working age (here taken to be the population aged 15 to 74). Each of the three right-hand-side terms is treated in a different way as follows:

- The first term on the right-hand side of [4] corresponds to the change in the structural unemployment rate. For the purposes of projecting the counter-factual, the structural unemployment rate is assumed to remain at its pre-crisis (2007) level, hence the contribution from this term to the post-crisis counter-factual growth rate is zero.
- The second term corresponds to the change in the trend labour force participation rate. One possibility would be to hold the trend labour force participation rate constant, however this would not allow for the effect of changes in the demographic composition of the working-age population on the aggregate participation rate. One simple way of allowing for this would be to hold age- and gender-specific participation rates constant and then project the aggregate participation rate allowing for the evolving demographic structure of the population of working age. However, a drawback of such an approach is that it does not allow for social trends whereby each generation or cohort may have a different tendency to work than its predecessors. Allowing for such effects is particularly important, for example, to capture the tendency for female participation to increase over time. Such effects are modelled here by holding the probability of entry and exit of the labour force for each age and gender cohort at its pre-crisis (2007) level.¹⁰ The evolution of the participation rate by age-gender cohort after 2007 is then influenced by what the same cohort did when younger. The counter-factual is then derived by aggregating across age-gender cohorts, although obviously it cannot take into account policy changes implemented after 2007.
- The final term captures an effect on the employment rate as the share of the working age population in total population changes. For most OECD countries, ageing populations mean that this share is declining and so acting as a drag on the employment rate. For the

purpose of the post-crisis counterfactual, this term is set to its realised values and then extended with the most recent vintage of projections. The essential point is that this demographic effect is the same in the most recent estimate of potential output and in the counter-factual, so that it does not contribute to estimates of the effect of the crisis.¹¹

In the first instance, the effect of the crisis is evaluated as the difference between the counter-factual trend and the latest estimates and projections of potential output consistent with the projections published in the November 2014 *OECD Economic Outlook*. This deflection of potential output relative to the pre-crisis trend can then be decomposed into an effect on trend labour productivity (which can be sub-divided into effects from total factor productivity and capital per worker) and potential employment (which can be sub-divided into effects from the structural unemployment rate and the trend participation rate). The potential output projections published in the November 2014 *OECD Economic Outlook* are generated using the framework described in the Appendix, although in some cases these are judgementally modified by country experts. Thus while the counterfactual is estimated from a set of simple mechanical rules, the most recent potential output projections should instead be interpreted as the OECD's current evaluation of potential output underlying recently published macroeconomic projections.

3.2. Empirical results

The method described above is applied to all individual OECD countries, with the results suggesting wide variation in the estimated effect of the crisis on individual countries (Table 1). Among the 19 countries which are identified by Laeven and Valencia (2012) as having experienced a banking crisis over the period 2007-11 (Austria, Belgium, Denmark, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Netherlands, Portugal, Slovenia, Spain, Sweden, Switzerland, the United Kingdom and the United States), the median loss in potential output in 2014 is about 5½ per cent, with only one country (Germany) not experiencing a loss. By contrast, among OECD countries which did not experience a banking crisis, the median potential output loss is only 2% in 2014, with six countries estimated to have experienced a gain in potential output relative to the counter-factual.

Table 1. **Estimated post-crisis deflection of potential output of OECD countries in 2014**

Percentage reduction relative to a pre-crisis counter-factual scenario

	TUR	DEU*	ISR	AUS	CHL	JPN	POL	FRA*	CAN	AUT*	OECD	NZL	USA*	KOR	CHE*	ITA*	BEL*
(1) Productivity	-4.2	-0.8	-0.6	0.8	-0.9	-0.8	-6.7	-1.1	-0.9	-3.5	-2.4	-0.9	-2.2	-5.0	-4.1	-3.0	-4.2
(1a) Capital/worker	-1.1	-0.7	0.0	5.0	1.5	-2.2	-0.8	-0.1	1.4	-1.6	-1.2	1.3	-1.5	-3.2	-1.6	-2.4	-0.6
(1b) Total factor productivity	-2.6	0.1	-0.2	-3.5	-1.6	1.5	-5.5	-0.9	-2.0	-1.8	-1.3	-1.8	-0.5	-1.3	-2.4	-0.7	-3.4
(2) Employment	9.6	5.5	4.7	0.8	1.8	1.4	7.0	0.5	0.1	2.0	0.6	-1.0	-0.2	1.9	0.9	-0.8	-0.2
(2a) Labour force participation	9.7	2.9	1.3	1.6	0.9	1.5	4.8	1.6	0.3	2.2	0.7	0.5	-0.5	1.9	1.2	0.6	-0.1
(2b) Structural unemployment	-0.1	2.7	3.5	-0.8	0.9	-0.1	2.3	-1.1	-0.2	-0.1	-0.1	-1.6	0.3	0.0	-0.2	-1.4	-0.1
(3) Total	5.5	4.7	4.1	1.6	0.9	0.6	0.3	-0.6	-0.9	-1.5	-1.9	-2.0	-2.4	-3.1	-3.2	-3.8	-4.3
Memorandum item:																	
Output gap	-1.9	-1.2	-0.1	-0.8	-1.2	0.3	-0.4	-2.2	-0.4	-2.8	-2.3	0.0	-3.3	-1.1	-1.1	-6.0	-1.2

Table 1. **Estimated post-crisis deflection of potential output of OECD countries in 2014** (cont.)
Percentage reduction relative to a pre-crisis counter-factual scenario

	LUX*	NOR	NLD*	MEX	DNK*	ESP*	SWE*	GBR*	FIN	SVK	IRL*	CZE	PRT*	SVN*	HUN*	ISL*	EST	GRC*
(1) Productivity	-4.4	-5.2	-3.3	-2.4	-3.3	2.1	-6.8	-7.6	-10.1	-8.3	-5.3	-12.5	-3.1	-8.6	-16.3	-9.1	-14.8	-15.2
(1a) Capital/worker	-2.1	-0.9	-0.3	-2.0	-1.4	1.3	-1.1	-1.9	0.8	0.8	-1.0	-1.6	-3.9	-2.2	-3.3	-8.6	-8.8	-1.2
(1b) Total factor productivity	-2.2	-4.0	-2.9	0.0	-1.8	1.0	-5.6	-5.5	-10.8	-8.7	-4.0	-10.9	0.9	-6.3	-13.0	-0.4	-6.0	-14.3
(2) Employment	-0.3	0.5	-1.4	-2.6	-2.3	-8.4	0.2	0.6	1.5	-1.1	-4.9	1.6	-8.0	-4.6	2.7	-4.9	-1.4	-8.6
(2a) Labour force participation	1.6	0.6	-0.3	-1.8	-1.4	-2.9	0.8	1.2	1.3	1.5	-2.3	1.0	-4.4	-2.5	4.5	-4.5	-2.7	-2.4
(2b) Structural unemployment	-1.9	-0.1	-1.1	-0.8	-0.9	-5.6	-0.5	-0.5	0.1	-2.6	-2.6	0.7	-3.5	-2.2	-1.9	-0.3	1.3	-6.2
(3) Total	-4.7	-4.7	-4.7	-5.0	-5.6	-6.4	-6.6	-6.9	-8.6	-9.4	-10.2	-10.9	-11.1	-13.2	-13.6	-13.9	-16.2	-23.8
Memorandum item:																		
Output gap	-3.1	-0.4	-3.1	-1.2	-2.6	-6.2	-2.2	-0.8	-3.2	-1.2	-5.3	-2.6	-6.0	-4.1	0.3	-2.4	-1.2	-12.9

Notes: An asterisk (*) denotes that the country is judged to have experienced a banking crisis between 2007 and 2011 based on Laeven and Valencia (2012). Estimated effects of the crisis are measured relative to a counter-factual scenario in which trend productivity continues at its pre-crisis (2000-07) trend growth rate, structural unemployment rates remain at their pre-crisis (2007) levels and trend participation rates are projected to allow for evolving demographics by holding labour force entry and exit rates constant at pre-crisis levels.

Row (3) = Row (1) + Row (2); Row (1) = Row (1a) + Row (1b); Row (2) = Row (2a) + Row (2b).

Source: OECD calculations based on OECD Economic Outlook, November 2014 long-term database.

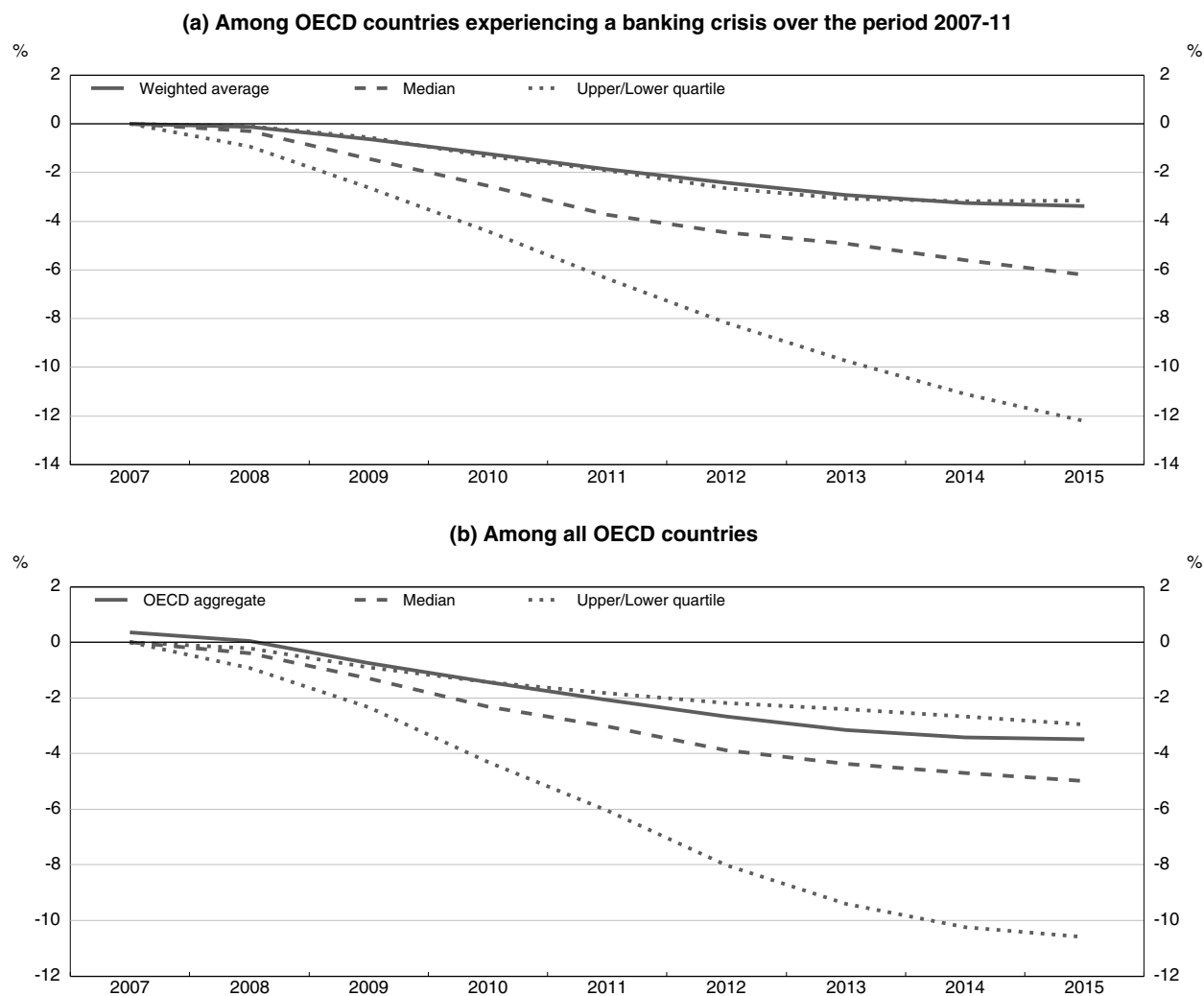
A strength of the current approach to evaluating the effect of the crisis on OECD potential output, is that results for individual countries and components of potential output are readily transparent. In some cases, this might seem to raise awkward questions, particularly when the effect of the crisis is calculated to have had a significant positive effect on some component of potential output. However, such effects serve as a reminder that while the crisis may have had a major impact on potential output, there were other effects (most notably from policy reforms) on potential output over the post-crisis period. This also raises the possibility of adjusting the estimate of the crisis hit to potential to exclude effects which can plausibly be attributed to factors other than the crisis. On this basis, an alternative adjusted estimate of the effect of the crisis on potential output is calculated by excluding the following effects:

- The positive effect on trend productivity due to faster post-crisis growth in capital per worker in Australia, Canada and Chile, which is probably explained by mining-related investment booms.
- Falls in structural unemployment in a number of countries – most notably Israel, Poland and Germany – which can be attributed to structural reforms rather than the crisis. For Israel, the decline of structural unemployment is partly explained by the lagged effect of policy reforms, notably a cut in social security payments and unemployment benefits (Elkayam and Ilek, 2013). In the case of Germany, this can be attributed to the lagged effects of the Hartz reforms, which pre-dated the crisis.
- The large positive effects on trend labour force participation for Turkey, Poland and Hungary which may be attributed to the effect of structural reforms. In the case of Turkey, a surge in participation can be attributed to the introduction of major job creation incentives in favour of disadvantaged groups (youth and women) which were very successful, particularly in a context where female participation was initially very low (OECD, 2010b). In the case of Poland, the strong rise in participation relative to the counterfactual is mostly explained by a gradual tightening in eligibility criteria for disability benefits and for early-retirement (OECD, 2014c). In the case of Hungary, the improvement in the labour force participation rate is partly explained by new tax/benefit policies aiming to make work pay and by the stepping up of public works schemes (OECD, 2014d).

Excluding these effects gives an estimate of the adverse effect of the crisis on aggregate OECD potential output in 2014 which is revised up to about 3½ per cent, although the median for all OECD countries is unchanged at about 4¾ per cent (Figure 2[b]). For countries having experienced a banking crisis over 2007-11, the median is also unchanged at about 5½ per cent in 2014, while for other OECD countries the median loss is doubled to 4%. The adjusted measures of potential output loss are used in the remainder of the paper because they better reflect the likely effect of the crisis on potential output.

Figure 2. Estimated effects of the crisis on the potential output of OECD countries

Percentage reduction in potential output relative to a pre-crisis counter-factual scenario

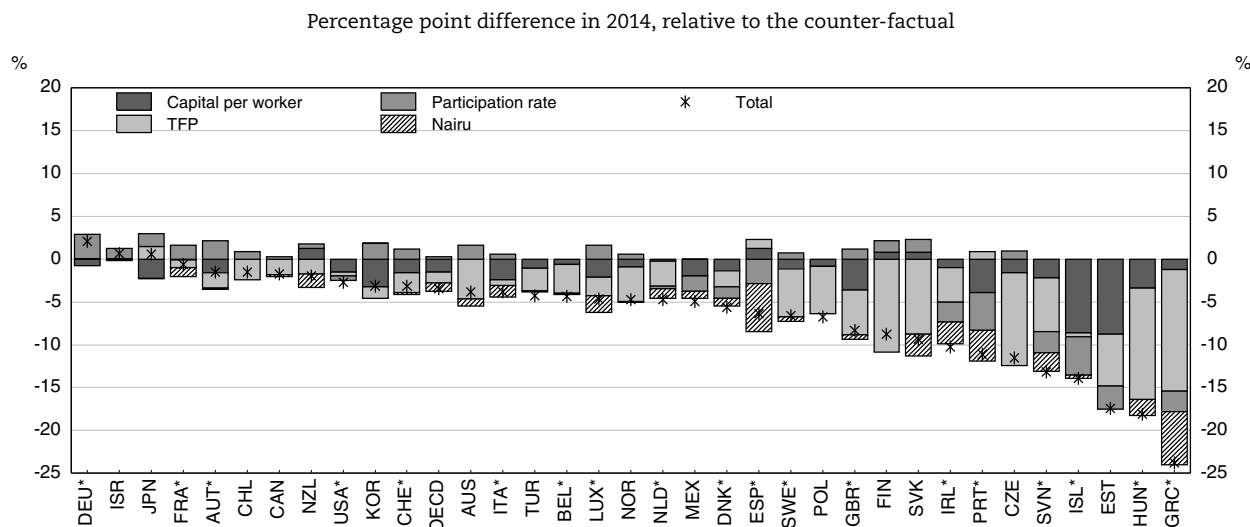


Note: Estimated effects of the crisis are measured relative to a counter-factual scenario in which trend productivity continues at its pre-crisis (2000-07) trend growth rate, structural unemployment rates remain at their pre-crisis (2007) levels and trend participation rates are projected to allow for evolving demographics by holding labour force entry and exit rates constant at pre-crisis levels. Positive effects on particular components of potential output for particular countries are excluded if these can be linked with policy reforms or other developments not associated with the crisis (see text for details).

Source: OECD calculations based on OECD *Economic Outlook*, November 2014 long-term database.

The estimated impact of the crisis on individual OECD countries is heterogeneous and, especially for those most severely affected, is mostly explained by the loss in labour productivity (Table 1, Figure 3 and Figure 4). For the Czech Republic, Estonia, Greece, Hungary, Iceland, Ireland, Portugal and Slovenia, the estimated negative impact of the crisis is above 10%, mostly reflecting lower trend productivity.¹² Lower productivity is both a result of lower total factor productivity (tfp) and lower capital per worker, with the former typically being more important. Most countries which experienced a banking crisis experienced a substantial reduction in tfp, although important exceptions are France, Germany and the United States.¹³ The decline in capital per worker accounts for more than 3 percentage points of the decline in potential output in 2014 for Estonia, Iceland, Hungary, Korea and Portugal, with these countries also being among those in which the pre-crisis growth of capital per worker was most rapid.¹⁴

Figure 3. **Contributions to the effect of the crisis on potential output for individual OECD countries**



Note: An asterisk (*) next to the country code denotes that the country is judged to have experienced a banking crisis between 2007 and 2011 based on Laeven and Valencia (2012). See also note to Figure 2.

Source: OECD calculations based on OECD *Economic Outlook*, November 2014 long-term database.

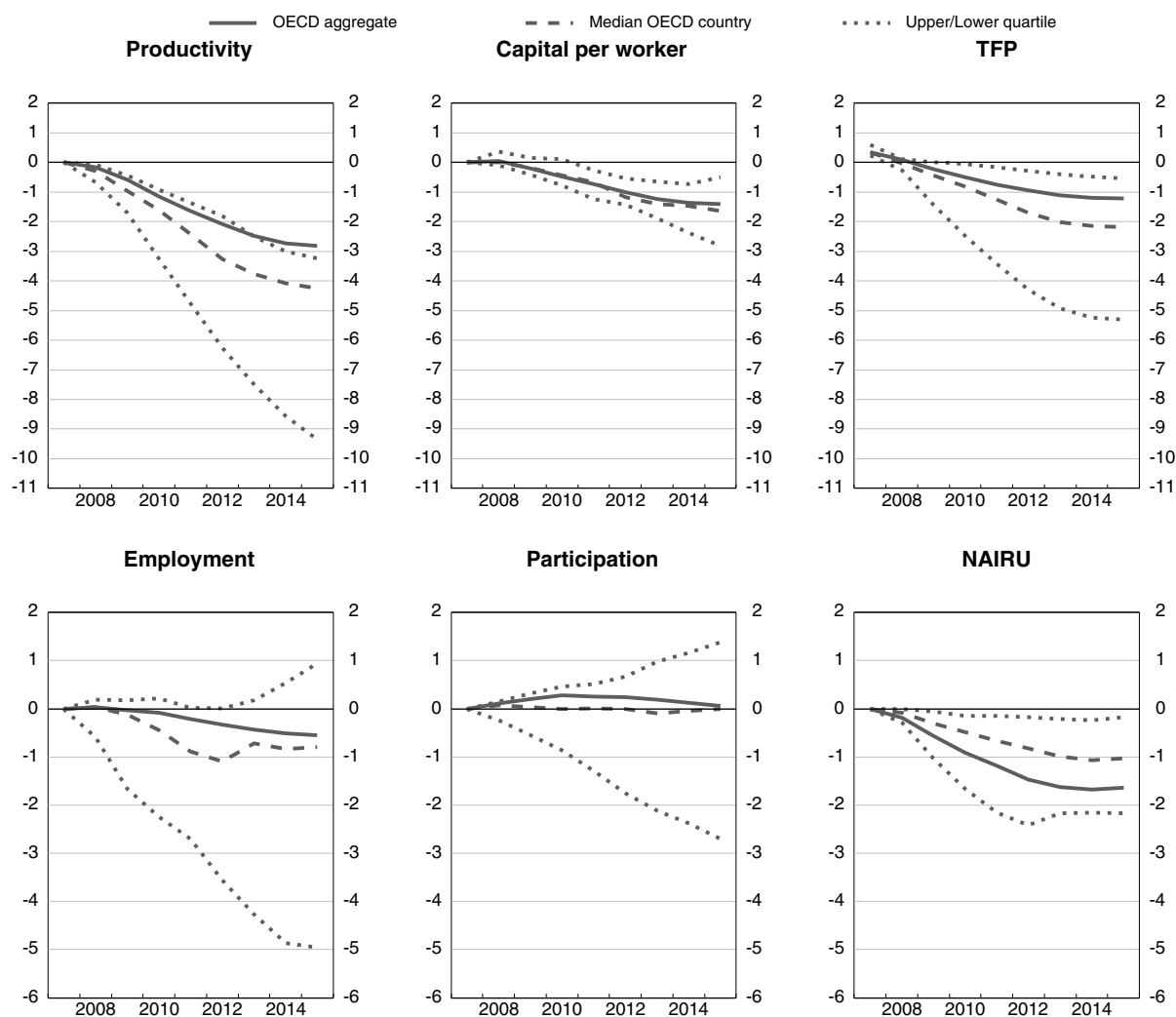
The contribution of lower potential employment to the estimated effect of the crisis is typically much smaller, although this contribution is more than 4 percentage points for Iceland, Ireland, Greece, Portugal, Spain and Slovenia. For Greece, Portugal and Spain most of this effect is explained by a rise in structural unemployment. While the loss related to potential employment is usually less than for trend labour productivity, the adverse fiscal impact of the former may be relatively more important (OECD, 2014a).

For the majority of OECD countries, trend participation rates have held up remarkably well when assessed against the counter-factual; for the median country, trend participation is close to the counter-factual (Figure 4). The absence of strong negative effects on trend participation for many OECD countries contrasts with the effect of previous severe downturns in OECD countries, which have been estimated to typically reduce participation rates by 1½ to 2½ percentage points over the medium term (Duval et al., 2011). The more resilient response from participation in the current crisis reflects policy

reforms which increase incentives to work at older ages, including pension reforms, the phasing out of early retirement schemes and the tightening of eligibility criteria for other social transfer schemes that operated as *de facto* early retirement schemes (OECD, 2013). It is also possible that declines in the value of defined-contribution pensions due to falls in equity prices have encouraged workers to postpone retirement. Nevertheless, there are some countries where the decline in the participation rate has substantially contributed to the negative effect of the crisis on potential output, mostly in countries where the increase in structural unemployment has been greatest. Thus, the adverse effect of the crisis on participation is estimated to have subtracted between 2 and 4½ percentage points from potential output in 2014 for Estonia, Iceland, Ireland, Greece, Portugal, Slovenia and Spain.

Figure 4. **Effects of the crisis on potential output components for banking crisis countries**

Breakdown by components, percentage deviation from the baseline



Notes: See note to Figure 2.

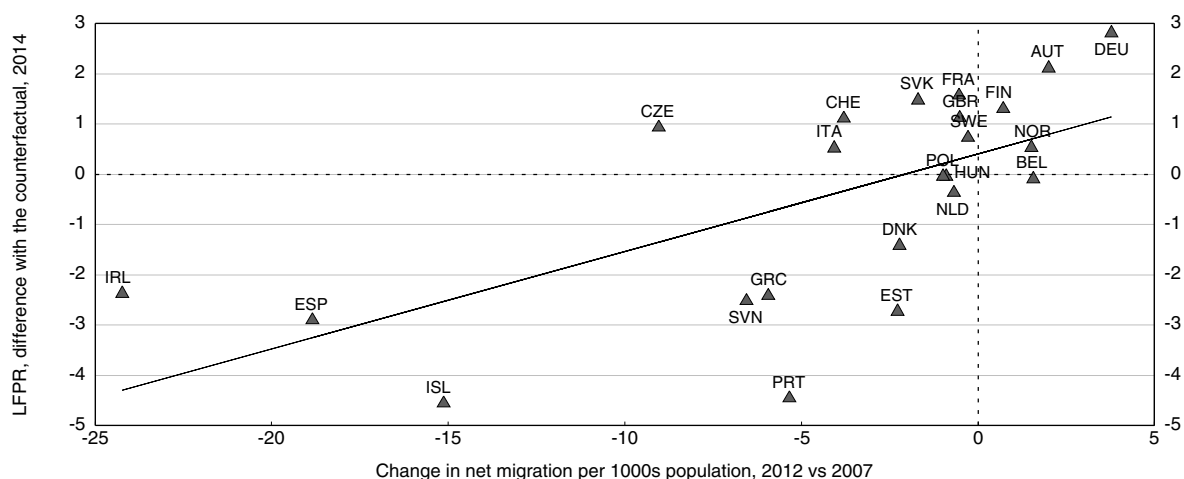
Source: OECD calculations based on OECD *Economic Outlook*, November 2014 long-term database.

The large fall in labour force participation in the United States since the start of the crisis is worth highlighting because there has been much discussion about the extent to which this is cyclical or explained by demographic or other trend factors.¹⁵ In terms of the analysis presented here, the fall in the trend participation rate is nearly all explained by demographic factors and so coincides closely with the counter-factual and hence relatively little of the fall in the trend participation rate is attributed to the crisis.¹⁶

Another factor which helps to explain the cross-country pattern of participation effects attributed to the crisis is the evolution of net migration.¹⁷ In the aftermath of the crisis, migration flows have changed substantially, particularly inside Europe, in favour of countries in which the effect of the crisis on the labour market has been less severe. For example, Germany experienced a major increase in net inward migration, whereas countries like Spain or Ireland, which were attracting many immigrants before the crisis, are now having large net outward migration rates. This is confirmed by a positive correlation between the estimated effect of the crisis on the participation rate for 23 European countries (row [2a] in Table 1) with the change in net migration rate between 2007 and 2012 (Figure 5).

Figure 5. **The crisis hit to participation and the change in net migration**

Percentage point difference



Source: OECD calculations based on OECD Economic Outlook, November 2014 long-term database; Eurostat.

4. Relating the crisis hit to potential output with pre-crisis conditions

This section attempts to cast light on the cross-country diversity of the hit to potential output, by examining which pre-crisis conditions are correlated with it. A similar exercise was undertaken by IMF (2009) in explaining medium-term output losses from previous financial crises (but excluding the present one) and thus provides a basis for comparison. For the purpose of the current paper, the country-specific effect of the crisis in 2014 is regressed on a range of variables summarising pre-crisis conditions relating to the macro conjuncture, indebtedness, pre-crisis trend growth and structural settings (Table 2). As the sample is modest (at most 34 OECD countries), each explanatory variable was tested separately with the addition of only a single control variable. The level of the output gap in 2007 is used as a control variable; it is statistically significant in nearly all cases and

negatively signed, implying that those countries which were over-heating more prior to the crisis tended to be those which experienced bigger losses in potential output.

In terms of the pre-crisis macroeconomic conjuncture, larger current account deficits and high inflation are both correlated with larger potential output losses. High levels of pre-crisis investment or housing investment are also correlated with larger potential output losses, but only if the output gap control variable is omitted. However, the pre-crisis fiscal balance is not found to be significant in explaining potential output losses. These results are broadly similar to the findings of IMF (2009), which found that high investment, large current account deficits and high inflation (relative to an historical average) all helped to explain cross-country differences in medium-term output losses from past financial crises. Overall, this provides support to arguments that macroeconomic policy should be concerned with financial developments over and above their immediate implications for inflation and fiscal balances.

The extent of total pre-crisis total economy indebtedness, either in terms of low net worth or high external debt, is correlated with the loss in potential output, as is higher household net worth (although the latter result is less statistically significant), but not government debt. These results confirm the findings of Sutherland and Hoeller (2013) in suggesting that a high level of debt creates weaknesses and vulnerabilities that both amplify shocks and limit countries scope to react during a crisis.

A higher pre-crisis trend growth is a predictor of larger subsequent output losses, although the only component which is significantly correlated with larger post-crisis potential output losses is the capital-per-worker component. These results, together with those that link the scale of output losses to indebtedness and other variables associated with financial excesses, suggest that underlying the potential output losses is a substantial misallocation of resources in the pre-crisis boom, especially capital but also labour, as well as the adverse effect of debt and capital overhangs during the bust (see Borio (2013) for a discussion of typical balance sheet recessions). By the same token, the pre-crisis potential growth rate is likely to have been exaggerated by the boost to productivity from rapid capital accumulation, which was unlikely to be sustained.

More flexible product markets are associated with smaller crisis-related losses in potential output, which may be because this allows for an easier reallocation of resources across firms and sectors in the aftermath of an adverse shock. Thus, countries with more competition-friendly product market regulation experienced smaller losses in potential output; a difference in the OECD PMR indicator by one standard deviation being associated with a difference in output losses of 2½ percentage points. On the other hand, while trade openness and financial openness may be related to the international transmission of shocks, they do not appear to be correlated with medium-term output losses.¹⁸

Table 2. The relationship between the loss in potential output and pre-crisis conditions

	Standard deviation of explanatory variable	Coefficient	Output gap, 2007	Adj. R2	Observations
Pre-crisis macro conjuncture					
Fiscal balance, 2007, % of GDP	4.3	0.210 0.167	-1.322 ***	0.49 -0.02	32
Current account balance, 2007, % of GDP	7.1	0.293 *** 0.467 ***	-0.944 ***	0.53 0.33	34
Real interest rate, average 2003-07	1.6	-0.320 1.057	-1.405 ***	0.43 0.05	30
Housing investment, % of GDP ¹	1.7	-0.077 -0.823 **	-1.287 **	0.47 0.13	21
Total investment, % of GDP ¹	2.6	-0.206 -0.668 **	-1.057 ***	0.41 0.18	24
CPI inflation, 2003-07, % points	1.1	-1.757 *** -2.595 ***	-1.121 ***	0.56 0.25	32
Pre-crisis indebtedness (2007, % of GDP)					
Gross government debt	34	-0.008 0.032	-1.336 ***	0.44 0.00	31
Net government debt	49	0.000 0.021	-1.332 ***	0.45 0.00	31
Net foreign debt	58	-0.035 *** -0.051 ***	-1.047 ***	0.54 0.26	34
Net worth, total economy	55	0.044 *** 0.066 ***	-1.004 ***	0.60 0.37	30
Net worth, household	66	0.027 ** 0.048 ***	-1.065 ***	0.53 0.28	30
Pre-crisis structural conditions					
Product Market Regulation, 2008 ²	0.3	-8.265 *** -8.499 **	-1.274 ***	0.58 0.12	29
Trade openness, 2007 ³	56	-0.005 -0.024	-1.222 ***	0.41 0.03	34
Financial openness, 2007 ⁴	487	0.001 -0.001	-1.265 ***	0.40 -0.03	31
Pre-crisis potential output growth rate, (average 2000-07, % pa)					
Potential output per capita growth rate	1.2	-1.253 * -2.506 ***	-1.007 ***	0.47 0.26	34
Trend labour efficiency growth rate	1.3	-0.603 -1.212	-1.193 ***	0.43 0.05	34
Potential employment growth rate	0.9	1.319 1.462	-1.238 ***	0.45 0.02	34
Capital per worker growth rate	1.7	-1.019 ** -1.932 ***	-0.912 ***	0.47 0.32	33

Notes: The table reports regression results where the dependent variable is the estimated percentage point effect of the crisis on potential output in 2014, so for most countries it has a negative sign. For each of the explanatory variables, listed in the first column, two regressions are run: the first (in the upper row) also includes the pre-crisis output gap as a control variable; the second regression (in the lower row) excludes the output gap variable. Statistical significance at the 1, 5 and 10% level is denoted by ***, **, and *, respectively.

1. Nominal housing and total investment as a percentage of GDP are measured as the difference between their averages over 2003-07 compared with 1980-2007.
2. A lower value of the Product Market Regulation indicator implies more competition-friendly regulation (see OECD, 2014b for details).
3. Trade openness is defined as the sum in 2007 of exports and imports divided by GDP.
4. Financial openness is defined as the sum in 2007 of liabilities and assets divided by GDP.

Source: Authors' calculations based on OECD Economic Outlook, November 2014 long-term database.

Notes

1. Note that the paper here is concerned with estimating the permanent medium-term loss to the level of output once any cyclical recovery is complete. However, even if this loss turns out to be small, the cyclical loss in output (measured by a cumulative output gap) until output returns to a sustainable trend is, for many OECD countries, likely to be larger than for any other downturn experienced in the post-war period.
2. The “Big Five” banking crises are Spain (1977), Norway (1987), Finland (1991), Sweden (1991) and Japan (1992). Judging from the source CS use to identify banking crises, namely Caprio and Klingebiel (2003), this corresponds with the “high-income” banking crises identified by CS.
3. For upper-middle income countries, CS estimate that banking crises reduce GDP after ten years by about 12%.
4. CS do conduct further sensitivity analysis, including attempting to allow for the endogeneity of crises, which allows for the effect of previous growth affecting the probability of a crisis, although they consider that this additional analysis implies that their main results concerning output losses are likely underestimated.
5. For IMF (2009), the 90% confidence interval around the output loss after seven years for high-income countries is roughly between 5% and 15%. The average output losses after ten years for both low-income and middle-income countries are similar to the 10% estimate for high-income countries.
6. A further illustration of the problem of using actual GDP to generate pre-crisis trends is that it can generate estimates of large output losses even when no crisis has occurred, because strong cyclical upswings are confused with strong pre-crisis trends. For example, applying the IMF (2009) methodology to Canada, France and the United Kingdom, but treating every single year over the period 1965-2000 as a potential “crisis”, gives 27 years in which output losses exceed 10%. However, according to the various sources used by all the studies referenced in this paper, none of these countries experienced a financial crisis over this period. While a majority of these large output losses occurred in the 1970s following the first oil shock, six episodes occurred during the 1990s.
7. Moreover, changes in population growth can have quantitatively important effects on GDP growth over the medium term horizons considered in this paper. For example in the case of Japan, the difference between the annual average population growth in the decade following the 1992 banking crisis compared with the decade preceding it, was about -0.2% per annum, which if cumulated over a decade would imply a difference in the level of GDP of 2%.
8. Another problem with comparing different vintages of potential output estimates is that there are often substantial revisions to the underlying GDP data between different vintages, especially when there has been a change to the system of national accounts. Such differences are likely to be more important the further apart the vintage of projections which are compared. Ball attempts to allow for such revisions by re-normalising estimates of different vintages of potential output on 2007 levels. However data revisions, especially when they involve a switch to a new system of national accounts, can also involve significant revisions to historical GDP and potential GDP growth rates.
9. Assuming a constant trend productivity growth rate implicitly ignores the possibility of a secular decline in productivity growth of the form described by Gordon (2014), except to the extent that it is already a feature of the period 2000-07. However, the larger decline in productivity growth experienced by countries that suffered a banking crisis compared with those that did not, suggests that this is unlikely to be the main explanation over the horizon of interest in this paper.
10. See Annex 3 of Burniaux et al. (2004) for a more detailed explanation of this approach to projecting labour force participation.
11. Not omitting this term would risk wrongly attributing population ageing effects on the employment rate to the crisis, which in the case of some countries (including Estonia, Spain, Greece, Slovenia and Japan) could exaggerate the crisis hit to potential by between 2 and 3½ percentage points by 2014.
12. The positive productivity effect for Spain may seem surprising, but this probably reflects the consequence of a shift of production away from construction where productivity is low.
13. The absence of any significant effect of the crisis on US tfp is consistent with the findings of Fernald (2014).
14. There is a surprising positive post-crisis effect from capital-per-worker on potential output for Spain. This occurs because the substantial fall in the capital stock did not match the even larger

fall in potential employment of 8½ per cent, which is larger for Spain than for any other country and partly reflects the collapse of the labour-intensive construction sector.

15. See, for example, Erceg and Levin (2013) and Van Zandweghe (2012) for alternative views on the relative importance of demographic/trend and cyclical factors explaining the fall in US participation.
16. If, instead, all of the fall in the trend participation rate since 2007 was attributed to the crisis then the estimated negative effect of the crisis on US potential output would rise from 2½ to 5 percentage points in 2014.
17. For a more detailed analysis of the links between migration and the labour market situation in OECD countries, see OECD (2014e).
18. IMF (2009) has a contrary finding that greater financial openness is associated with smaller output losses, although the evidence is described as “weak”.

References

- Abiad, A., R. Balakrishnan, P.K. Brooks, D. Leigh and I. Tytell (2009), “What’s the Damage? Medium-term Output Dynamics After Banking Crises”, *IMF Working Paper* 09/245.
- Ball, L. (2014), “Long-term Damage from the Great Recession in OECD Countries”, *National Bureau of Economic Research Working Paper* No. 20185, May.
- Borio, C. (2012), “The Financial Cycle and Macroeconomics: What Have We Learnt?”, *BIS Working Paper*, No. 395, December.
- Borio, C. (2013), “On Time, Stocks and Flows: Understanding the Global Macroeconomic Challenges”, *National Institute Economic Review*, No. 225, August.
- Burniaux, J., R. Duval and F. Jaumotte (2004), “Coping with Ageing: A Dynamic Approach to Quantify the Impact of Alternative Policy Options on Future Labour Supply in OECD Countries”, *OECD Economics Department Working Papers*, No. 371, OECD Publishing.
- Cerra, V. and S.C. Saxena (2008), “Growth Dynamics: The Myth of Economic Recovery”, *American Economic Review*, 98(1), pp. 439-57.
- Duval, R., M. Eris and D. Furceri (2011), “The Effects of Downturns on Labour Force Participation”, *OECD Economics Department Working Papers*, No. 875, OECD Publishing.
- Elkayam, D. and A. Ilek (2013), “Estimating the NAIRU using both the Phillips and the Beveridge curves”, *Bank of Israel Discussion Paper*, No. 2013.10, December.
- Erceg, C. and A.T. Levin (2013), “Labor Force Participation and Monetary Policy in the Wake of the Great Recession”, *IMF Working Paper*, WP/13/245, July.
- Fernald, J. (2014), “Productivity and Potential Output, Before, During and After the Great Recession”, *Federal Reserve Bank of San Francisco, Working Paper* 2014-15.
- Furceri, D. and A. Mourougane (2009), “The Effect of Financial Crises on Potential Output: New Empirical Evidence from OECD countries”, *OECD Economics Department Working Papers*, No. 699.
- Gordon, R. (2014), “The Demise of U.S. Economic Growth: Restatement, Rebuttal, and Reflections”, *NBER Working Paper* No. 19895.
- Guichard, S. and E. Rusticelli (2011), “Reassessing the NAIRUs after the Crisis”, *OECD Economics Department Working Papers*, No. 918, OECD Publishing.
- Haugh D., P. Ollivaud and D. Turner (2009), “The Macroeconomic Consequences of Banking Crises in OECD Countries”, *OECD Economics Department Working Papers*, No. 683, OECD Publishing.
- IMF (2009), “What’s the Damage? Medium-Term Output Dynamics after Financial Crises”, *World Economic Outlook*, Chapter 4, October.
- Johansson, Å., Y. Guillemette, F. Murtin, D. Turner, G. Nicoletti, C. de la Maisonneuve, P. Bagnoli, G. Bousquet and F. Spinelli (2013), “Long Term Growth Scenarios”, *OECD Economics Department Working Papers*, No. 1000, OECD Publishing.
- Laeven, L. and F. Valencia (2012), “Systemic Banking Crises Database: An Update”, *IMF Working Paper*, WP/12/163, June.

- Morrisson, C. and F. Murtin (2010), "The Kuznets Curve of Education: A Global Perspective on Education Inequalities", *Centre for the Economics of Education Discussion Papers*, No. 116, London School of Economics.
- OECD (2010a). *OECD Economic Outlook*, No. 87, May, OECD Publishing.
- OECD (2010b). *OECD Economic Surveys: Turkey*, OECD Publishing.
- OECD (2013), *OECD Employment Outlook*, OECD Publishing.
- OECD (2014a), *OECD Economic Outlook*, No. 96, November, OECD Publishing.
- OECD (2014b), *Economic Policy Reforms 2014; Going for Growth Interim Report*, OECD Publishing.
- OECD (2014c), *OECD Economic Surveys: Poland 2014*, OECD Publishing.
- OECD (2014d), *OECD Economic Surveys: Hungary 2014*, OECD Publishing.
- OECD (2014e), *Matching Economic Migration with Labour Market Needs*, September, OECD Publishing.
- Reinhart, C.M. and K.S. Rogoff (2009), "The Aftermath of Financial Crises", *National Bureau of Economic Research. Working Paper*, No. 14656, January.
- Schreyer, P. (2003), "Capital Stocks, Capital Services and Multi-factor Productivity Measures", *OECD Economic Studies*, No. 37, 2003/2.
- Sutherland, D. and P. Hoeller (2012), "Debt and Macroeconomic Stability: An Overview of the Literature and Some Empirics", *OECD Economics Department Working Papers*, No. 1006, OECD Publishing.
- Van Zandweghe, W. (2012), "Interpreting the Recent Decline in Labour Force Participation", *FRB Kansas City Economic Review*, first quarter.

APPENDIX A

The OECD methodology for generating potential output

Potential output (Y^*) for the whole-economy is assumed to be determined by a Cobb-Douglas production function with constant returns to scale and factor inputs of physical capital (K), human capital (H) and potential employment (N^*), with technical progress of the Harrod-neutral labour-augmenting form, hereafter referred to as 'trend labour efficiency' (E^*):

$$y^* = \alpha (h + n^* + e^*) + (1 - \alpha) k, \quad [A.1]$$

where lower case letters denote logarithm and α is the wage share.

Each component is estimated and projected separately using different techniques.

The physical capital stock is defined in terms of the non-housing productive capital stock. It corresponds to the accumulation of past non-housing investment flows, taking into account that the efficiency of assets typically falls as its life expectancy diminishes (Schreyer, 2003) and can be expressed as:

$$K_t = K_{t-1}(1 - r_t) + I_t, \quad [A.2]$$

where r is the implicit rate of loss of productive capacity, and I is the non-housing investment flow. When available, historical data are extracted from the OECD Productivity Database. Otherwise a starting point for the capital output ratio is estimated based on average investment rate over the previous ten years and a loss rate at 4%. The projected capital stock is built up from short-term projections of investment consistent with the *Economic Outlook* projections.

Human capital is derived from empirical estimates of the return to average years of schooling with falling marginal returns at higher levels of education as in Morrison and Murtin (2010). The resulting human capital series is smoothed using a Hodrick-Prescott filter to better reflect the trend evolution of the variable.

Potential employment is computed as the combination of the structural unemployment rate (NAIRU), the working-age population (PWA) and the trend participation rate (LFPR*):

$$N^* = LFPR^*PWA(1 - NAIRU) \quad [A.3]$$

The NAIRU is estimated using a reduced form Philips-curve by means of a Kalman filter (see Guichard and Rusticelli, 2011). Beyond the estimation period the NAIRU is usually held constant, except where there is evidence that policy reforms may have an influence or where projections imply large changes in long-term unemployment, in which case rules-of-thumb and judgement are used to revise the projection.

The working-age population is defined as the population aged between 15 and 74. The data are sourced from either Eurostat (for EU countries) or United Nations (for other countries).

The trend participation rate for the population age 15 to 74 is computed in a two-step stage. First a reduced-form equation for the actual participation rate is estimated for each country using the unemployment gap and time trends, and the predicted values from this equation when the unemployment gap is closed are used to remove economic cycles from the actual participation rate. The resulting participation rate is then smoothed with a Hodrick-Prescott filter to obtain the trend participation rate. The trend participation rate is projected using a cohort model described in Johansson et al. (2013).

Labour efficiency is defined as the residual of the production function. Hence an initial (unsmoothed) estimate of labour efficiency is calculated using the Cobb-Douglas production function in [1] but using actual GDP in place of Y^* and actual employment in place of potential employment. Then a Hodrick-Prescott filter is applied to this unsmoothed labour efficiency to provide an estimate of the trend labour efficiency. To project labour efficiency as well as to reduce end-point problems associated with using an HP-filter, data on GDP and factor-inputs are extended with short-term forecasts of these variables. Beyond a two year horizon, labour efficiency is projected using a conditional convergence model described in Johansson et al. (2013).

One simplification which is made in the main paper is to combine the human capital and trend labour efficiency components into a single components which is referred to as labour efficiency.

The future of health and long-term care spending

by

Christine de la Maisonnette and Joaquim Oliveira Martins*

This paper proposes a new set of public health and long-term care expenditure projections until 2060, following up on the previous set of projections published in 2006. It disentangles health from long-term care expenditure as well as the demographic from the non-demographic drivers, and refines the previous methodology, in particular by better identifying the underlying determinants of health and long-term care spending and by extending the country coverage to include BRIICS countries. A cost-containment and a cost-pressure scenario are provided together with sensitivity analysis. On average across OECD countries, total health and long-term care expenditure is projected to increase by 3.3 and 7.7 percentage points of GDP between 2010 and 2060 in the cost-containment and the cost-pressure scenarios, respectively. For the BRIICS over the same period, it is projected to increase by 2.8 and 7.3 percentage points of GDP in the cost-containment and the cost-pressure scenarios, respectively.

JEL classification codes: H51, I12, J11, J14.

Keywords: Public health expenditures, long-term care expenditures, ageing populations, longevity, demographic and non-demographic effects, projection methods.

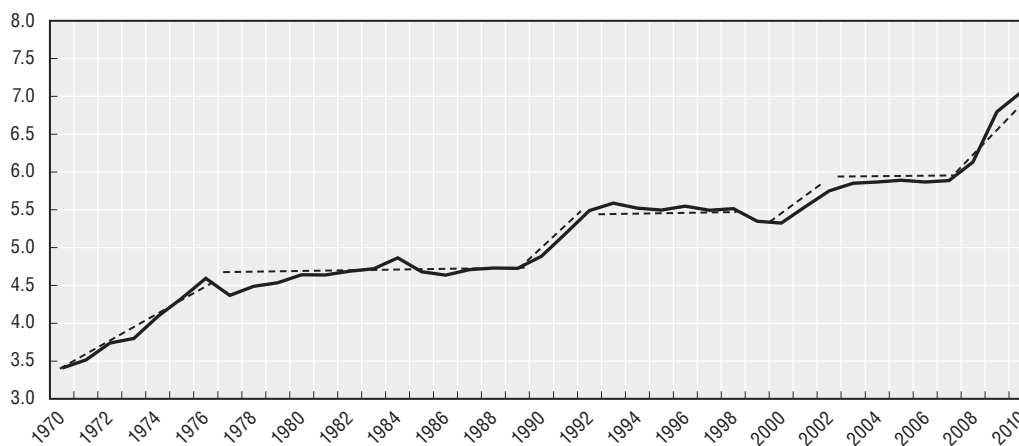
* The authors are, respectively, members of the Economics Department and the Directorate for Public Governance and Territorial Development. Correspondence: christine.maisonnette@oecd.org and joaquim.oliveira@oecd.org. They would like to thank Jørgen Elmeskov, Fabrice Murtin, Giuseppe Nicoletti and Jean-Luc Schneider as well as their colleagues from the Health Division of the Directorate for Employment, Labour and Social Affairs for their valuable comments on an earlier draft of the paper. They are also grateful to Luca Lorenzoni, Fabio Pammolli and Yuki Murakami for their thorough reading of the paper and their useful suggestions. Special thanks go to Ines Gomez Palacio for excellent technical and editing assistance. The views expressed in this paper are those of the authors and do not necessarily reflect those of the OECD and its member countries.

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

1. Introduction

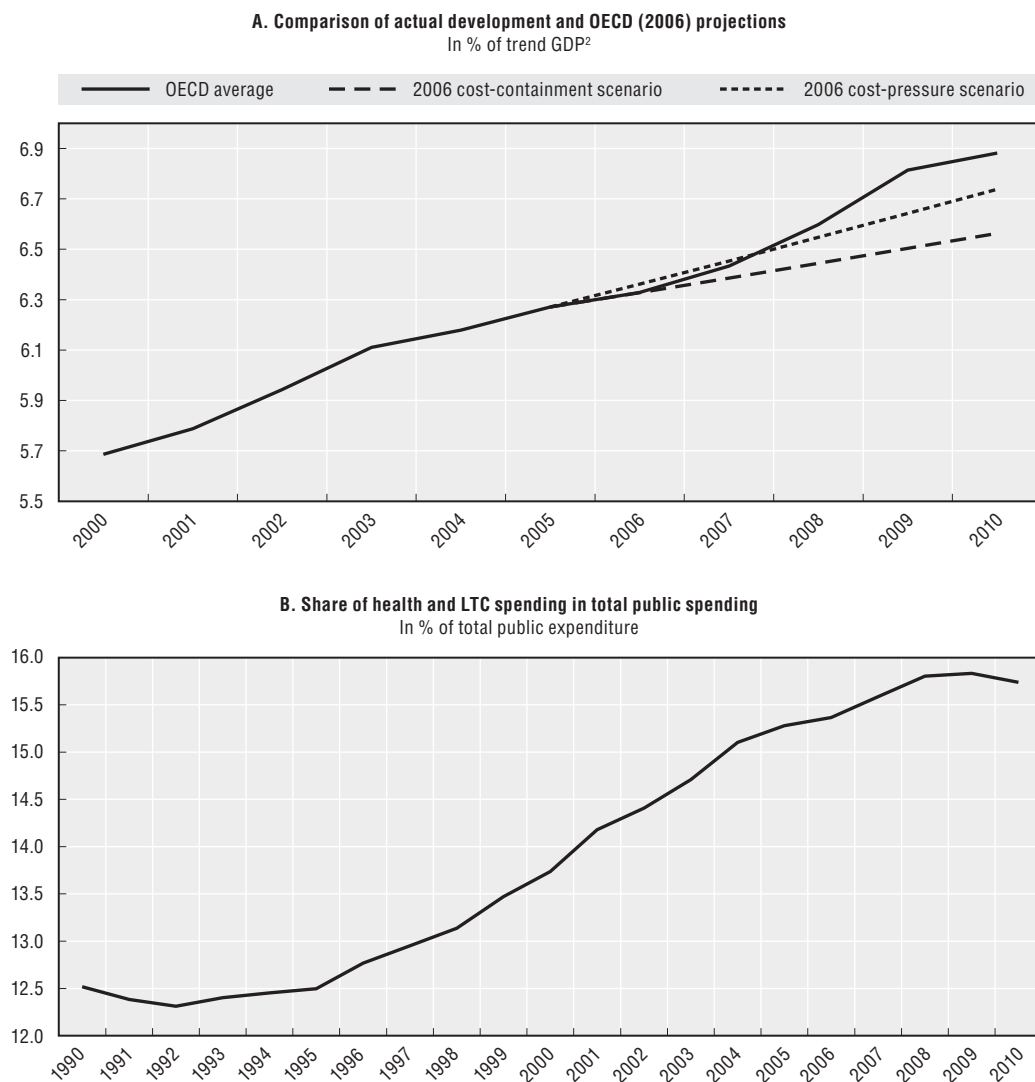
Spending on health and long-term care (henceforth LTC) is a first-order policy issue for most governments in OECD countries. These expenditures are putting pressure on public budgets, adding to that arising from other social spending programmes and, once interest rates normalise, servicing of higher debt levels than in the past. The ratio of public health and LTC expenditure to GDP¹ has been rising steadily for several decades. Since 1970, on average across OECD countries, the expenditure to GDP ratio has increased by 3.5 percentage points to reach around 7 percentage points in 2010. Interestingly, this ratio tends to increase during recession periods meaning that health expenditure keeps on rising while GDP stagnates or decreases (Figure 1). Despite a recent slowdown, as a share of GDP, public health and long-term care expenditure have still increased by slightly less than 1 percentage point from the level prior to the crisis.

Figure 1. **Evolution of public health and long-term care expenditure¹**
As a % of GDP



1. Unweighted average of available OECD countries.
Source: OECD Health Database (2012).

Looking towards the future, the OECD (2006) public health and LTC spending projections highlighted the growing pressures over the next 50 years. In an upside “cost-pressure” scenario, average health and LTC expenditure was projected to almost double again, to reach approximately 13% of GDP by 2050. But, even in the so-called “cost-containment” scenario, spending was still projected to rise to approximately 10% of GDP on average across OECD countries over the next 50 years. Comparing the 2006 OECD projections with actual data up to 2010 (Figure 2, panel A), the observed increase in the ratio of public health and LTC spending to GDP was above even the pessimistic “cost-pressure scenario”. On average for OECD countries, these expenditures have represented an increasing share of total public expenditures (Figure 2, panel B).

Figure 2. **Evolution of public health and long-term care expenditures**¹

1. Unweighted average of available OECD countries.

2. The trend GDP was derived from the OECD Economic Outlook Database.

Source: OECD Health Database (2011), OECD (2006), OECD Economic Outlook Database, No. 91.

The projections in this paper update and refine the OECD (2006) analysis by extending the country coverage and improving the method of estimating future developments in health and LTC expenditure. In addition to the 34 OECD countries, health and LTC expenditure is also projected for Brazil, China, India, Indonesia, Russia and South Africa (henceforth BRIICS). Concerning the methodology, as in 2006, the new projections separate health and LTC and, within each type of expenditure, demographic from non-demographic drivers. For health care, the main difference concerns the non-demographic drivers, with an attempt to better understand the residual expenditure growth by determining which share can be explained by the evolution of health prices and technology effects. Regarding LTC, a more precise estimation of the determinants of the number of dependants (people needing help in their daily life activities) is provided. In particular, the current set of projections treats health expenditure itself as a determinant of the dependency ratios

(number of dependent people in total population) by age groups. Finally, the new projections extend to 2060.

The projections are provided for public expenditure only. Nonetheless, the cross-country differences in treatment of tax expenditures,² which can be large in some cases, may introduce serious bias in the split between public and private expenditures. Indeed, countries interpret differently the OECD manual on the System of Health accounts on how to deal with tax expenditures. Australia and Germany, for instance, deduct tax expenditures from private insurance expenditures and report it as public expenditure. But, in the United States, tax expenditures are not considered as public expenditure (OECD, 2010).

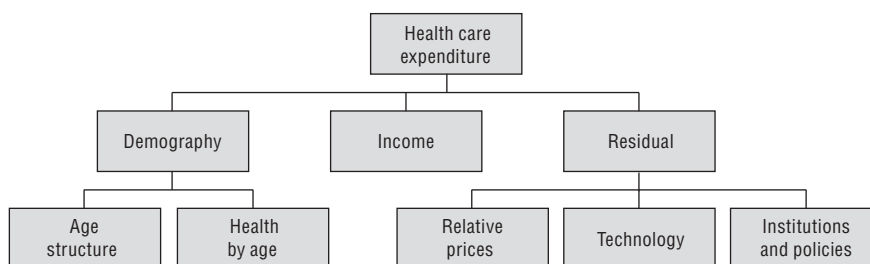
The paper is set out as follows. In Section 2, the main demographic and non-demographic drivers of public health expenditure are analysed and projected over the next half century. Cost-containment (with implicit policy actions) and cost-pressure (without policy actions) scenarios are presented, followed by a discussion of the sensitivity of the results to key assumptions. In Section 3, the same sequence applies to public LTC expenditures. In Section 4, different health and LTC expenditure scenarios are combined in order to provide a range of estimates for total expenditure. Section 5 concludes.

2. Health care

2.1. The determinants of public health care expenditure

The drivers of public health care expenditure are demographic and non-demographic (Figure 3). Demographic drivers relate broadly to the age structure of the population and the evolution of its health status, while a non-demographic driver is income. The relationship between health expenditures and income remains an unsettled issue, but independent of the precise income elasticity (measure of the responsiveness of health expenditure to a change in income) the combination of demographic and income effects fails to explain a large part of the total growth in public health-care expenditure in the past. Thus, a residual unexplained growth remains. Relative prices, technological progress and the underlying health policies and institutions are the most likely candidates for explaining this residual (see Box 3 in de la Maisonneuve and Oliveira Martins, 2013).

Figure 3. The determinants of public health expenditure



To quantify the relative contribution of each driver over the past 15 years, a simple accounting analysis was carried out (for data sources and methodology, see Annex A). The analysis accounts for changes in the population structure during the period 1995-2009, and assumes an income elasticity of health spending of either 0.8 or 1 (see below for a discussion of this hypothesis). Based on these assumptions an expenditure residual was derived. As can be seen from Table 1, demographic drivers explain relatively little of past

Table 1. **Decomposing growth in public health spending (1995-2009)¹**

	Real health spending (per capita in 2005 PPP USD)	Age effect	Income effect (income elasticity = 0.8)	Residual ²	<i>Memo item:</i> Residual with unitary income elasticity
	(Average annual % change)	(Average annual % contribution to change in spending)			
Australia	4.1	0.4	1.7	1.8	1.4
Austria	3.3	0.4	1.3	1.5	1.2
Belgium	4.2	0.4	1.2	2.7	2.4
Canada	2.6	0.6	1.3	0.8	0.5
Chile	7.7	0.6	2.2	4.8	4.3
Czech Republic	4.2	0.6	2.3	1.4	0.8
Denmark	3.7	0.2	0.8	2.7	2.5
Estonia	6.1	0.6	3.9	1.8	1.0
Finland	4.1	0.6	2.0	1.5	1.1
France	1.6	0.5	0.9	0.3	0.0
Germany	1.7	0.6	0.8	0.2	0.0
Greece	5.9	0.6	2.7	2.5	1.8
Hungary	2.0	0.5	2.1	-0.6	-1.1
Iceland	3.1	0.4	2.0	0.9	0.4
Ireland	6.5	0.3	2.9	3.3	2.6
Israel	1.6	0.4	2.4	-1.2	-1.7
Italy	3.1	0.6	0.4	2.1	2.0
Japan	2.7	1.2	0.8	0.7	0.5
Korea	11.0	1.1	3.1	6.5	5.7
Luxembourg	1.9	0.1	2.3	-0.5	-1.1
Mexico	2.5	0.5	1.1	1.0	0.7
Netherlands	5.2	0.5	1.4	3.3	2.9
New Zealand	6.3	0.4	0.3	5.5	5.5
Norway	3.5	0.1	1.3	2.1	1.7
Poland	6.9	0.7	3.6	2.6	1.7
Portugal	4.6	0.6	1.5	2.4	2.0
Slovak Republic	6.3	0.6	3.6	2.1	1.3
Slovenia	3.0	0.8	2.1	0.2	-0.3
Spain	3.4	0.5	1.5	1.4	1.0
Sweden	3.2	0.2	1.6	1.4	1.0
Switzerland	2.9	0.4	0.9	1.6	1.4
Turkey	7.7	0.5	2.6	4.5	3.9
United Kingdom	4.6	0.2	1.5	2.8	2.5
United States	3.6	0.3	1.1	2.3	2.0
Brazil	4.8	0.6	1.2	2.9	2.6
China	11.2	0.6	7.3	3.0	1.3
India	6.6	0.3	4.2	2.0	1.0
Indonesia	8.0	0.5	1.9	5.5	5.0
Russian Federation	3.7	0.4	3.3	0.1	-0.6
South Africa	3.1	0.4	1.6	1.2	0.8
OECD total average	4.3	0.5	1.8	2.0	1.5
BRICS average	6.2	0.5	3.2	2.5	1.7
Total average	4.6	0.5	2.0	2.0	1.5

1. Or latest available year. Figures may not add up due to rounding.

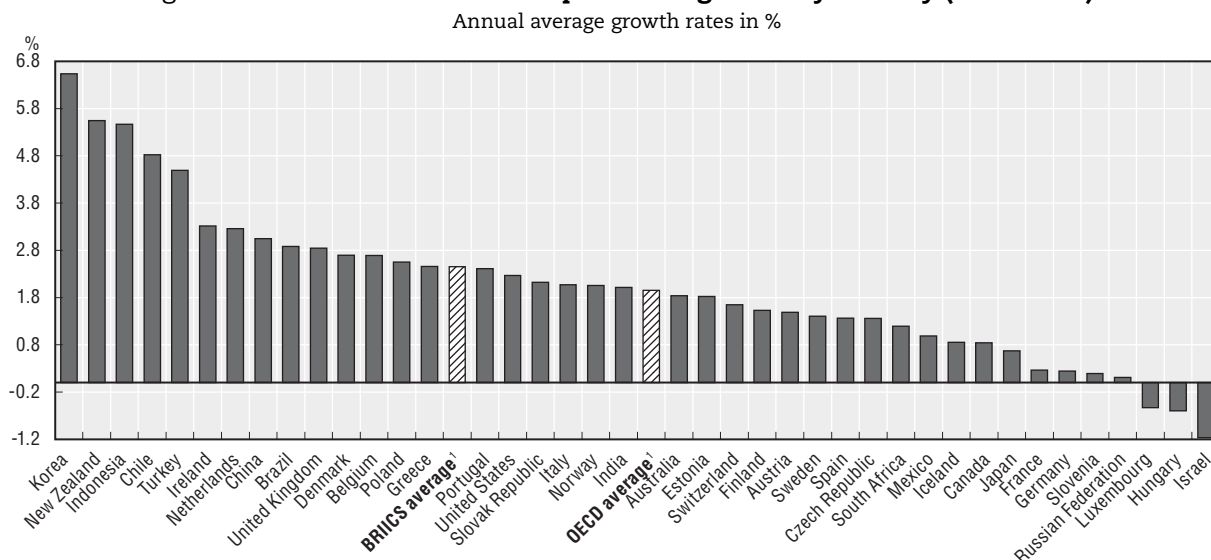
2. The residual expenditure growth is computed by subtracting the age effect and the increase in income (using an elasticity of 0.8) from the increase in real health spending. Accordingly, denoting h_t the real health spending, Age_t the age effect, and Y/N_t the real income per capita, the residual expenditure growth (Res) is expressed as follows: $Res = h_t - Age_t - 0.8 * Y / N$.

developments in health spending. Between 1995 and 2009, public health spending grew on average and in real terms (adjusted for overall inflation) by 4.3% per year in OECD countries, of which only 0.5 percentage points can be attributed to “pure demographic”

developments. Assuming an income elasticity equal to 0.8, income effects add 1.7 percentage points of growth on top of the 0.5 percentage points attributable to demographics. Therefore, a residual growth can be estimated at around 2 percentage points per year on average across OECD countries, with the BRIICS experiencing even more important residual growth effects. With an income elasticity of one, the residual on average for the OECD would still be around 1.5% per year. Only with a somewhat unrealistic income elasticity of 1.8, demographic and income effects could explain most of the observed expenditure growth.

The calculated expenditure residual growth displays large differences across countries (Figure 4). It ranges from above 5% per year for countries such as Korea and New Zealand, to values well below 0.5% for France, Germany, Slovenia and Russia, and is even negative for Israel. Extrapolating these country-specific idiosyncrasies over the next 50 years may generate unreliable results as most of them would lead to unsustainable expenditure ratios. For this reason, in this study, it was preferred to use a common average residual growth (see below).

Figure 4. **Health care residual expenditure growth by country (1995-2009)**



1. Unweighted average.

Source: OECD Health Database (2011) and authors' calculations.

2.2. Projection framework

The framework used to project public health expenditure combines both the demographic and non-demographic drivers described above. The evolution of each driver is first projected separately, and subsequently combined to compute the future growth of total health expenditure. Thus, the evolution of the future share of public health care spending in GDP depends on demographic and real income per capita projections, as well as on the extrapolated residual growth component. The demographic driver includes three elements – death-related costs, pure age effects based on population projections drawn from different sources and a healthy-ageing effect – which are described below. Future GDP is based on projections published in the *OECD Economic Outlook*, No. 91.³ As in the historical

accounting exercise, the income effects are captured by an elasticity of health expenditures that can take the value of 0.8 or 1, depending on the scenario.

The growth rates in spending ratios are first projected for each country and they are then adjusted to allow for a certain convergence across countries towards a common target level. *Ceteris paribus*, countries having a below-average initial level of public health expenditure to GDP ratios are projected to experience higher growth rates than those close to the average. This makes the projections more comparable across countries, as the effects of the different mechanisms at work during the projection period are isolated from the impact of the initial conditions. This is particularly important in the context of a projection method that does not assume country-specific residual growth in expenditures (as in the baseline assumed in this paper), or for countries that have in the base year a very low level of spending, such as emerging economies.

Therefore, the projected (for period t) health spending ratios were adjusted as follows:

$$\left(\frac{HE}{Y}\right)_{i,t} = \left[1 + \left(\frac{(HE/Y)_{OECD,0}}{(HE/Y)_{i,0}} \right) \cdot g_i \right] \cdot \left(\frac{HE}{Y}\right)_{i,0}$$

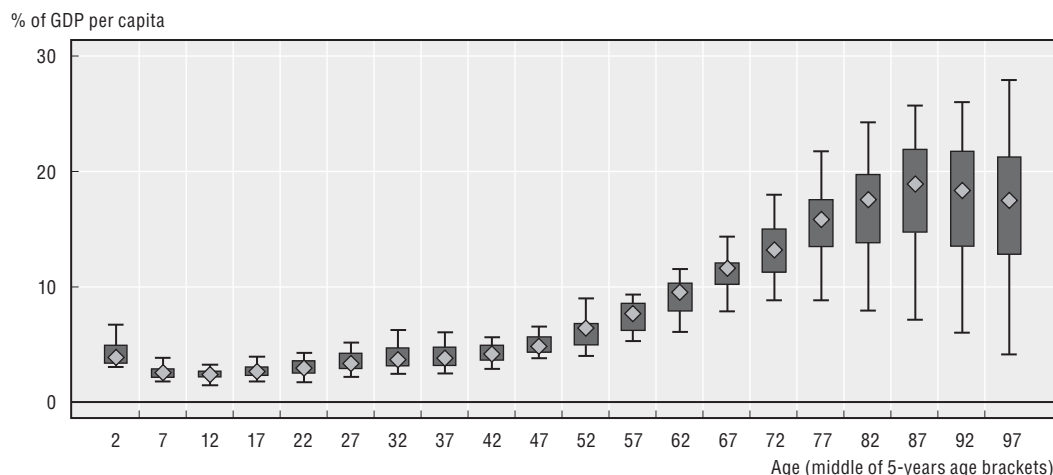
where g_i is the growth rate of health spending for country i (from period 0 to period t); $\left(\frac{HE}{Y}\right)_{i,0}$ is the health expenditure ratio for country i in the base period; and $\left(\frac{HE}{Y}\right)_{OECD,0}$ is the health expenditure ratio for the OECD average in the base period.

When the spending ratio of a given country is below (or above) the OECD average for the base year this adjustment will increase (decrease) the projected growth rate of expenditures to GDP, thus allowing for convergence to take place. In order to smooth out the impact of the recent crisis on expenditure, the base year spending ratios are computed as the average shares of public health care spending in GDP over 2006-10. This framework is used to project public health care expenditures over the period 2010-60.

The base information used to construct the health expenditure projection framework is the average health expenditure profile by age group (Figure 5). Average health

Figure 5. **Public health care expenditure by age groups**¹

% of GDP per capita



1. The graph shows the dispersion of health care expenditure across countries by age groups. The diamonds represent the median. The boxes are the 2nd and 3rd quartiles of the distribution of expenditure across countries. The whiskers are the 1st and 4th quartiles.

Source: European Commission, 2009 Ageing Report: Economic and Budgetary Projections for the EU27 Member States (2008-60).

expenditures are relatively high for young children; they decrease and remain stable for most of the prime-age period, and then start to increase rapidly in older age, the health care cost of people aged 90 and over being six times that of young people. Until the age of 65, health expenditure profiles are rather similar across countries, but from 65 onwards, they display a large heterogeneity. The standard deviation by age group increases from 1.8% for the 65-69 age group to more than 5% for people aged 90 and over.

From Figure 5, it could be expected a priori that an ageing population would be associated with increasing aggregate per capita public health care expenditures: the fact that the share of older people in the population is growing faster than that of any other age group, both as a result of longer lives and a lower birth rate, should generate an automatic increase in the average. However, this intuition finds little support in the data and assessing the effect of population ageing on health and health care has proved to be far from straightforward (Breyer et al., 2011).

Consistent with a large number of previous studies (Felder et al., 2000; Seshamani and Gray, 2004; Breyer and Felder, 2006; and Werblow et al., 2007; etc.), this paper assumes that what matters for health spending is not ageing but rather the proximity to death, i.e. the so-called “death-related costs” (DRC) hypothesis. This interpretation is consistent with the observed facts that health-care expenditure tends to increase in a disproportionate way when individuals are close to death, and mortality rates are obviously higher for older people. When the projected increase in life expectancy is accompanied by an equivalent gain in the number of years spent in good health, the health care spending is only driven by the proximity to death and not by an increase in the average age of the population. In other words, it is not ageing *per se* that pushes up average health expenditures, but rather the fact that mortality rates increase with age. The death-related costs hypothesis is, therefore, consistent with a so-called healthy-ageing regime, where longevity gains are all translated into years in good health (Box 1).

Box 1. Healthy ageing hypothesis

To take into account the “healthy ageing” hypothesis, the survivor expenditure curve is allowed to shift rightwards according to longevity gains, progressively postponing the age-related increases in expenditure. First, the curve by five-year age groups is interpolated in order to derive a yearly age profile. In this way, the expenditure curve can be shifted smoothly over time, in line with life expectancy gains.

Subsequently, the shift of the curve can be simulated by subtracting the increase in life expectancy at birth according to national projections from each current age. For example, a 70-year old person in Germany is projected to have the health status of a 67-year old person by 2025 and that of a 64-year old person by 2050.

By contrast, in a “pure demographic” scenario, the expenditure curves would not shift rightwards with ageing, reflecting the implicit assumption of unchanged health status at any given age. When these curves stay put in the presence of longevity gains, the share of life lived in “bad health” increases with life expectancy.

It should be noted that the population projections used in the analysis are pre-determined and do not take into account the effect of health spending on health status and longevity. Making population projections dependent on the level of health spending is beyond the scope of this project and could be the object of further research.

To account for the dominance of death-related costs and the healthy-ageing regime, the expenditure curve by age group needs to be split for each period into two segments of the population: survivors and non-survivors. The non-survivors' expenditure curve can be estimated by multiplying the estimated health costs in the proximity to death by age group by the number of deaths per age group. The "cost of death" was proxied here by the health expenditure per capita for the oldest age group (95+) multiplied by a factor that captures the tendency for the costs of death to be higher at younger ages. This factor is kept constant at 4 for people aged between 0 and 59 years and declines linearly (towards unity) afterwards.⁴

As regards non-survivors, two different demographic effects are at play. On the one hand, the number of deaths is set to rise due to the transitory effect of the post-War baby-boom. On the other hand, if mortality falls over time due to an ongoing increase in longevity, fewer will be at the very end of life in each given year, thus mitigating health care costs.⁵ The impact of demographic developments on public health care expenditures will depend on the relative size of these effects.

In principle, the survivors' expenditure curve for each country could be derived from the difference between the total cost curve and that of the non-survivor. However, given the uncertainties surrounding these data, it seemed preferable to estimate an average expenditure curve for survivors, to be then used for all countries. In this way, the projections are less sensitive to initial conditions and to country-specific data idiosyncrasies. This expenditure curve for survivors was estimated econometrically as a non-linear function of age. Each country-specific curve is then calibrated in order to fit the base year of the projections.

2.3. Non-demographic drivers

As mentioned above, overall demographic drivers explain relatively little of past developments in health spending, even assuming that the expenditure curve by age remains constant (or abstracting from "healthy ageing" effects). Therefore, the non-demographic drivers must play an important role: these are the growth in income and a residual growth component.

The effect of real income growth on public health expenditures has been the subject of a vast debate, but the precise value of the income elasticity is still uncertain. Empirical estimates tend to increase with the degree of income aggregation, implying that health care could be "an individual necessity and a national luxury" (Getzen, 2000). However, a high aggregate income elasticity (above unity) often found in macro studies may result from biases in estimates originating from a number of sources, such as failure to appropriately control for quality effects and account for the peculiar statistical properties of some of the variables. Based on the most recent findings from this literature (e.g. Acemoglu et al., 2009; and Holly et al., 2011), as well as estimations carried out by the authors, a real income elasticity equal to 0.8 has been used (a middle point across those estimates) and the sensitivity of the projections to this assumption was subsequently tested (by contrast, OECD, 2006 assumed a unitary income elasticity throughout).

Once income growth has been accounted for, several factors may explain the residual growth observed in the past: developments in relative prices, technological progress and the features of health institutions and policies (see Box 3 and Appendix 2 in de la Maisonneuve and Oliveira Martins, 2013). It is not feasible at this stage to project these drivers individually. Nonetheless, in an attempt to interpret this residual growth, an econometric equation has

been estimated, incorporating explicitly the effects of prices and a proxy for quality. Accordingly, the income elasticity (ϵ), as well as the price (β) and quality (γ) elasticities, were derived from a panel regression (in log terms and expressed in quality-adjusted health volumes per capita), augmented with the drift factors related to demography and exogenous policy factors captured by a time trend (e.g. institutional reforms increasing the coverage of the health coverage system):

$$\log\left(\frac{he}{N}\right) = \alpha_c + \theta \cdot \log(Age) + \beta \cdot \log\left(\frac{P}{P_Y}\right) + \gamma \cdot \log(Q) + \epsilon \cdot \log\left(\frac{Y}{N}\right) + \tau \cdot T + u$$

where he denotes health care demand volumes (deflated for price and quality), N is total population, Age is the average age of the population, P are health care prices, P_Y is the GDP factor, Q is a quality index for health care products and Y is real GDP; α_c correspond to country fixed-effects; T to a time trend; and u a randomly distributed residual. The equation was estimated both in levels and in first differences. The demographic effect is captured by the average age of the population. The quality effect (Q) was loosely proxied by a ratio involving the relative pace of patent creation in the health sector compared with the rest of the economy for the average of OECD countries. This ratio is assumed to represent the change in the technology frontier. To capture country-specific effects related to the diffusion and adoption of technology, the ratio is multiplied by the share of R&D expenditures in GDP.

The results of the econometric estimation are relatively robust, with the value of the income elasticity fluctuating from around 0.5 to close to 1, depending on the specifications. The preferred econometric estimate (column 8 of Table 2) gives an estimated elasticity of 0.42 for relative prices (β) and 0.91 for the technology/quality effect (γ). Translating these into expenditure elasticities, $(\beta + 1)$ and $(1 - \gamma)$, and the observed growth rates of (P/P_Y) and (Q) the overall effect of technology and relative prices is estimated to have increased health-care expenditures by 0.8% per year. In other words, the econometric model is able to explain more than one-third of the past expenditure residual of 2% per year for the OECD average, displayed in Table 1. Estimates suggest that the residual expenditure is also driven by other factors, such as changes in policy and institutions which can be loosely captured by a time trend. These other factors account for 0.9% of the increase in health expenditure per year. Overall, these estimates suggest that the residual expenditure has increased health-care expenditures by around 1.7% (i.e. 0.8% + 0.9%) per year.⁶

As stated previously, a common residual growth is assumed for all countries in order not to extrapolate country-specific idiosyncrasies over a long period. Based on the estimates of Table 2, the projections assume a uniform residual expenditure growth of 1.7% per year.

2.4. Summing up

To summarise graphically the way in which demographic and non-demographic drivers contribute to health care spending (Figure 6), the mechanical effect of population ageing on expenditures can be displayed as moving up along the expenditure curve, assuming that the age profile of expenditures remains constant over time (Panel 1). This age factor is then adjusted by incorporating the healthy ageing hypothesis, corresponding to a rightward shift of the curve (Panel 2). This shift implies that older people still cost more than the young, but at progressively older ages. Finally, the expenditure curve may shift upwards⁷ (Panel 3) due to non-demographic drivers (income, the quality-adjusted relative health prices and other unexplained factors, such as policies and institutions).

Table 2. **Real public health care expenditure per capita: Econometric estimates**

	Pooled	Fixed effects with time dummies	Fixed effects with time trend	Fixed effects with time dummies	Fixed effects with time trend	Fixed effects with time dummies	Fixed effects with time dummies	Fixed effects with time trend	Income elasticity = 0.8	<i>Memo item:</i> First differences estimates
log(GDPV per capita)	0.914*** (0.02)	0.394** (0.18)	0.495*** (0.13)	0.775*** (0.14)	0.634*** (0.14)	0.964*** (0.14)	0.749*** (0.12)	0.532*** (0.12)		
log(Average age of population)	2.603*** (0.56)	3.007*** (0.59)	1.399*** (0.45)	1.396*** (0.42)	2.611*** (0.46)	2.606*** (0.42)	1.342*** (0.41)	1.471*** (0.44)	0.962** (0.43)	
time trend			0.009*** (0.00)		0.009*** (0.00)			0.009*** (0.00)	0.001 (0.00)	
log(GDP deflator)			0.453*** (0.08)	0.482*** (0.08)	0.929*** (0.06)	0.777*** (0.06)				
log(Health prices)			-0.415*** (0.07)	-0.505*** (0.06)						
lagged log(Technology)			0.908*** (0.03)	0.918*** (0.04)			0.919*** (0.04)	0.912*** (0.03)	0.930*** (0.03)	
lagged dlog(Technology)										1.027*** (0.05)
lagged log(Quality adjusted health prices)					-0.788*** (0.03)	-0.755*** (0.03)				
Log[Relative prices (Health prices/GDP deflator)]							-0.503*** (0.06)	-0.415*** (0.07)	-0.492*** (0.06)	
dlog(GDPV per capita)										0.535*** (0.12)
dlog(Average age of population)										1.509 (1.45)
Dlog[Relative prices (Health prices/PGDP)]										-0.626*** (0.06)
_cons	-11.591*** (2.08)	-7.824*** (2.19)	-3.513** (1.78)	-5.889*** (1.85)	-7.535*** (1.86)	-10.936*** (1.90)	-5.414*** (1.54)	-4.179*** (1.54)	-6.646*** (1.38)	0.006 (0.02)
N	474	474	463	463	453	453	463	463	463	447
r2	0.825	0.863	0.924	0.941	0.913	0.934	0.941	0.924	0.847	0.777
r2_a	0.824	0.845	0.918	0.933	0.906	0.925	0.933	0.918	0.837	0.746

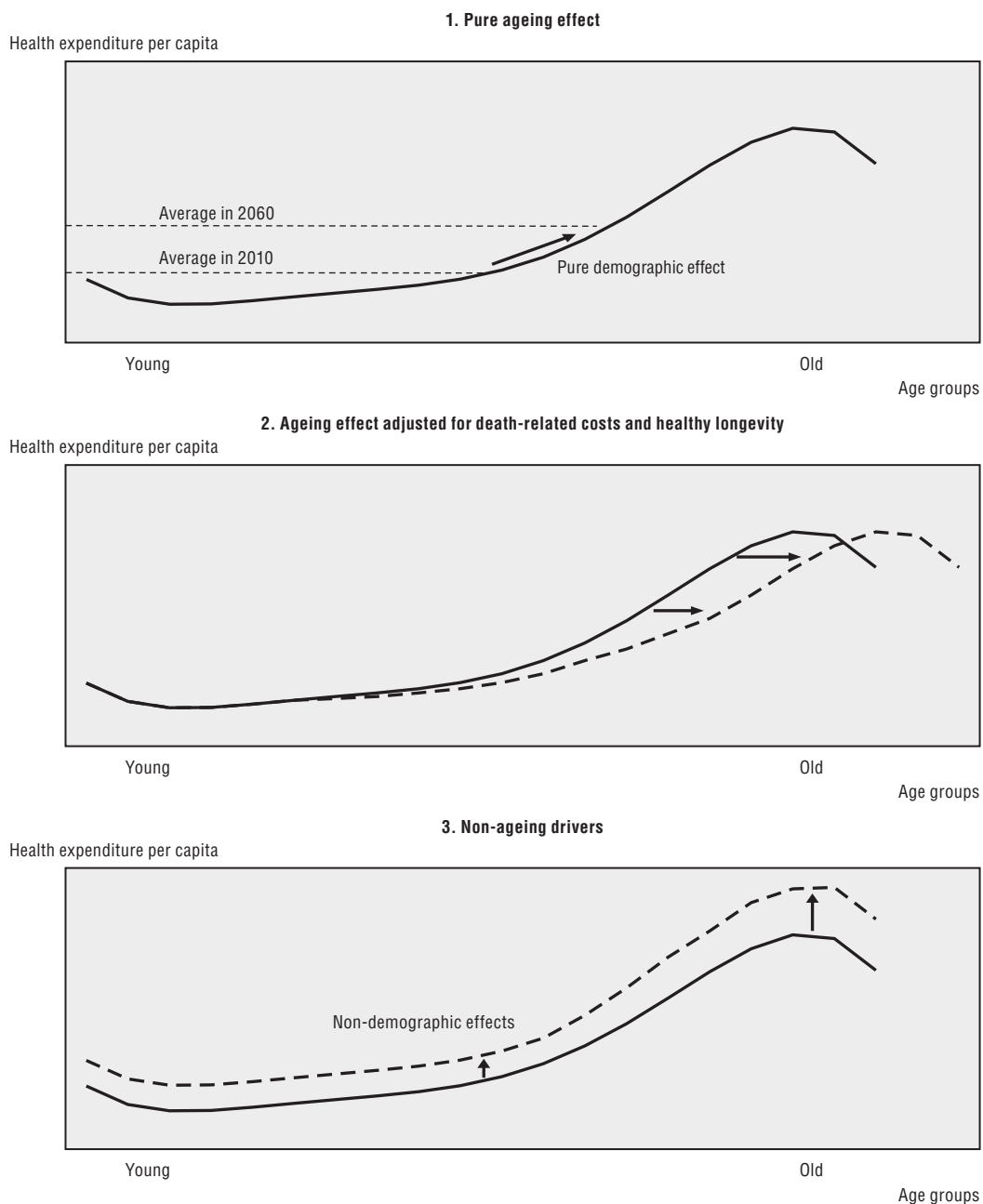
Standard errors in parentheses.

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

Note: Projections are based on the estimates highlighted within the box.

2.5. Projection results: Public health expenditures, 2010-2060

This section applies the analytical framework to project public health care expenditures in different scenarios over 2010-60. The main assumptions underlying each scenario are listed in Table 3. The section looks first at the demographic and income effects on expenditures and, subsequently, two scenarios regarding the evolution of drivers other than income and demographics (i.e. the residual growth) are analysed. In the “cost pressure scenario” no policy action is undertaken to curb pressures on expenditure whereas the “cost-containment scenario” assumes some policy action to rein in these pressures on expenditures. Even though the nature of such policies is not made explicit, they can be thought of as actions to limit the pressures arising from excessive growth of quality-adjusted relative health prices, e.g. by monitoring more closely the adoption of new technologies or modifying incentives via changes in the governance of health institutions.

Figure 6. **Shifts in expenditure profiles, ageing and non-ageing effects**

All these scenarios assume healthy ageing and an income elasticity of 0.8. Finally, sensitivity analysis is carried out in the context of the cost-containment scenario. More specifically, sensitivity analysis is carried out regarding the value of the income elasticity, the evolution of residual spending growth (i.e. unrelated to income and demographics) and the “healthy ageing” hypothesis. Concerning the latter, results are obtained also in the case of no healthy ageing (i.e. morbidity expands with gains in longevity) and for a more optimistic hypothesis under which each year increase in life expectancy is translated into two years spent in good health.

Table 3. **Assumptions underlying the alternative projection scenarios: Public health care**

Scenarios	Health status	Income elasticity	Non-demographic, non-income drivers
Demographic effect	Healthy ageing: Longevity gains are translated into equivalent additional years in good health.	n.a.	n.a.
Cost-pressure scenario	Healthy ageing: Longevity gains are translated into equivalent additional years in good health.	Income elasticity = 0.8	The residual is assumed to grow at 1.7% over the projection period.
Cost-containment scenario	Healthy ageing: Longevity gains are translated into equivalent additional years in good health.	Income elasticity = 0.8	The residual growth is 1.7% in the starting period and then converges to 0 in 2060.
Sensitivity analysis based on the cost-containment scenario			
Compression of morbidity	Longevity gains are doubled into additional years in good health.	Income elasticity = 0.8	The residual growth is 1.7% in the starting period and then converges to 0 in 2060.
Expansion of morbidity	No healthy ageing adjustment, i.e. longevity gains do not translate into additional years in good health.	Income elasticity = 0.8	The residual growth is 1.7% in the starting period and then converges to 0 in 2060.
Country-specific quality adjusted relative prices	Healthy ageing: Longevity gains are translated into equivalent additional years in good health.	Income elasticity = 0.8	The residual growth is country specific and converges to 0 in 2060.
Income elasticity = 0.6	Healthy ageing: Longevity gains are translated into equivalent additional years in good health.	Income elasticity = 0.6	The residual growth is 1.7% in the starting period and then converges to 0 in 2060.
Income elasticity = 1	Healthy ageing: Longevity gains are translated into equivalent additional years in good health.	Income elasticity = 1	The residual growth is 1.7% in the starting period and then converges to 0 in 2060.

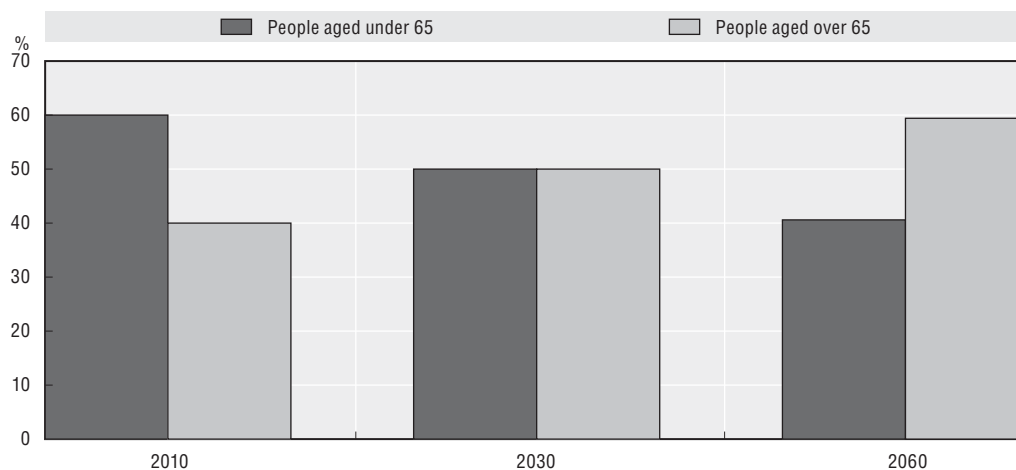
Note: The key assumption changed in each scenario is in bold.

2.6. Demographic and income effects

As discussed above, demographic effects on public health care expenditures can be decomposed into health care costs for survivors, the adjustment for “healthy ageing” and death-related costs. The pure ageing effect can be quite large in some countries, but tends to be compensated by better health status. Whereas on a per capita basis death-related costs account for the largest part of lifetime expenditures, for the population as a whole they account for only a small fraction of the increase in expenditures as a share of GDP since they concern only the non-survivors.

On average, the demographic effect only accounts for a small increase in expenditure. In OECD countries and on its own, it pushes spending from 5.5% of GDP on average to 6.2% in 2060. In non-OECD countries, the demographic effect on its own increases spending from 2.4% of GDP to 3.4% in 2060. While the “healthy ageing” assumption may render the simulation of demographic effects relatively optimistic, this is in line with observed patterns of health status regimes in many OECD countries. It is assumed that the same pattern will apply to non-OECD countries. The effect is projected to apply evenly before and after 2030 for both OECD and non-OECD countries.

It is noteworthy that while demography is not a main driver of health care spending growth, it will induce a significant change in the structure of spending over time. In 2010, 60% of health care expenditures were directed to people below 65 years old. In 2060, roughly the same percentage of expenditures will be directed to people aged above 65, reflecting an increase from 15% to 30% of their share in total population (Figure 7). This evolution is only the consequence of demographic effects as non-demographic effects are

Figure 7. **Shares of health care expenditure by age in total health care expenditure**

assumed to affect all age groups in the same proportion. In other words, while demography is not a main driver of health spending growth, it will induce a significant change in the structure of the spending.

The total effect of demographics and income on the increase in public health care expenditures by 2060 is very different across countries (Table 4). In OECD countries, the demographic effect ranges from 0.2 percentage points of GDP in Belgium, Iceland and the United Kingdom to more than 1.8 points in Chile, Korea and Turkey. For non-OECD countries it varies from zero in South Africa to 1.7 percentage points of GDP in Brazil and China, reflecting more rapid ageing in the latter two countries.

With an assumed less-than-unitary real income elasticity, the underlying increase in income would imply *ceteris paribus* a decline in the health expenditure to GDP ratio. As the GDP projections embody a degree of convergence towards living standards of high-income countries, this downward income effect on spending ratios is more important for rapidly catching-up low-income countries. Indeed, on average for OECD countries, the decline in expenditure ratios due to income growth will amount to -0.8 percentage points, while in non-OECD countries it will be -1.4 percentage points.

2.7. The cost pressure scenario

As mentioned earlier, in the “cost pressure scenario”, it is assumed that on top of the demographic and income effects, health expenditure will grow by a residual 1.7% per year over the whole projection period. In this scenario, the OECD average health expenditure to GDP ratio is projected to increase by more than 6 percentage points, reaching close to 12% in 2060 from a starting value of 5.5% in 2006-10. The largest increases (above 7 percentage points of GDP) are found in Chile, Korea and Turkey, while the lowest (below 6 percentage points of GDP) are found in Belgium, Denmark, Estonia, Finland, Hungary, Iceland, Sweden and the United Kingdom. Starting from a much lower ratio of health expenditure to GDP at 2.4%, the BRIICS will experience a larger increase as the ratio is set to more than triple, to reach 8.3%. The most important part of these increases is due to the residual component. For both OECD and non-OECD countries, the increase, on average, will be slightly higher in the second part of the projections period.

Table 4. **Projection scenarios for public health care expenditure**

	Average 2006-10	Percentage point deviations from starting period in 2030				Percentage point deviations from starting period in 2060			
		Contributions of demographic and income effects		Total increase in spending ratio		Contributions of demographic and income effects		Total increase in spending ratio	
		Demographic effect	Income effect	Cost- pressure	Cost- containment	Demographic effect	Income effect	Cost- pressure	Cost- containment
Australia	5.6	0.4	-0.4	2.0	1.5	0.8	-0.8	6.3	2.5
Austria	6.6	0.4	-0.2	2.1	1.7	0.6	-0.6	6.3	2.4
Belgium	5.8	0.2	-0.3	1.8	1.4	0.2	-0.7	5.8	1.9
Canada	5.8	0.6	-0.3	2.2	1.8	0.7	-0.7	6.3	2.5
Chile	3.1	0.9	-0.6	2.3	1.8	1.8	-1.1	7.1	3.2
Czech Republic	5.5	0.4	-0.5	1.9	1.4	0.7	-0.9	6.1	2.2
Denmark	6.3	0.4	-0.2	2.1	1.6	0.3	-0.7	5.9	2.0
Estonia	4.3	0.2	-0.5	1.6	1.1	0.6	-1.1	5.8	2.0
Finland	5.2	0.4	-0.3	2.0	1.5	0.3	-0.7	5.9	2.0
France	7.4	0.3	-0.3	1.9	1.4	0.3	-0.6	6.1	2.2
Germany	7.3	0.5	-0.3	2.1	1.6	0.6	-0.7	6.2	2.3
Greece	5.4	0.3	-0.3	1.9	1.4	0.7	-0.6	6.4	2.5
Hungary	4.8	0.2	-0.5	1.6	1.2	0.4	-1.0	5.8	1.9
Iceland	5.8	0.3	-0.2	2.0	1.5	0.2	-0.7	5.9	2.0
Israel	4.0	0.3	-0.2	2.0	1.5	0.7	-0.7	6.4	2.5
Ireland	5.5	0.3	-0.2	2.1	1.6	0.5	-0.5	6.4	2.5
Italy	6.1	0.3	-0.2	2.1	1.6	0.7	-0.6	6.4	2.6
Japan	6.1	0.6	-0.3	2.3	1.8	0.8	-0.8	6.3	2.5
Korea	3.3	1.1	-0.5	2.6	2.1	2.0	-0.8	7.6	3.7
Luxembourg	4.4	0.4	-0.1	2.2	1.7	0.7	-0.2	6.9	3.0
Mexico	2.5	0.6	-0.5	2.1	1.6	1.6	-1.1	6.9	3.0
Netherlands	6.4	0.6	-0.3	2.2	1.8	0.7	-0.7	6.3	2.4
New Zealand	6.4	0.5	-0.3	2.1	1.6	0.8	-0.9	6.3	2.4
Norway	5.1	0.3	-0.4	1.9	1.4	0.5	-0.7	6.1	2.2
Poland	4.1	0.4	-0.5	1.9	1.4	1.0	-0.9	6.4	2.6
Portugal	6.5	0.4	-0.2	2.0	1.6	0.8	-0.7	6.5	2.6
Slovak Republic	5.4	0.5	-0.5	1.9	1.5	1.1	-0.9	6.5	2.6
Slovenia	5.2	0.5	-0.3	2.1	1.7	1.0	-0.8	6.6	2.7
Spain	5.6	0.4	-0.3	2.0	1.6	1.0	-0.6	6.7	2.8
Sweden	6.6	0.2	-0.3	1.8	1.3	0.3	-0.7	5.9	2.0
Switzerland	5.7	0.4	-0.3	2.1	1.6	0.9	-0.8	6.5	2.6
Turkey	3.8	0.7	-0.6	2.0	1.5	1.8	-1.1	7.0	3.1
United Kingdom	6.5	0.2	-0.2	1.9	1.4	0.2	-0.7	5.9	2.0
United States	7.1	0.4	-0.3	2.0	1.5	0.5	-0.7	6.1	2.2
OECD average¹	5.5	0.4	-0.3	2.0	1.6	0.8	-0.8	6.3	2.5
Brazil	3.7	0.8	-0.6	2.1	1.6	1.7	-1.1	7.0	3.1
China	1.9	0.9	-1.1	1.7	1.3	1.7	-1.7	6.4	2.5
India	1.2	0.3	-1.0	1.3	0.8	0.9	-1.7	5.5	1.6
Indonesia	1.2	0.6	-0.8	1.7	1.3	1.3	-1.5	6.2	2.3
Russian Federation	3.1	0.1	-0.6	1.5	1.0	0.2	-1.0	5.5	1.6
South Africa	3.2	0.0	-0.6	1.4	0.9	0.0	-1.2	5.2	1.3
Non-OECD average¹	2.4	0.5	-0.8	1.6	1.2	1.0	-1.4	5.9	2.1
Total average¹	5.0	0.4	-0.4	2.0	1.5	0.8	-0.8	6.3	2.4

1. Unweighted average.

Source: Authors' calculations.

2.8. The cost-containment scenario

While the cost-pressure projections may be useful as a benchmark, they do not seem very plausible. It is unlikely that public health-care expenditures to GDP could continue to grow at such rates, without limit. For instance, regarding drugs, the following policies can help containing the costs: a more intense use of generics, a revision of reimbursement practices for drugs with low medical service (and more stringent restrictions on exemptions to this rule), a more frequent re-evaluation of prices of new drugs (in general a more transparent and dynamic pricing of new drugs), centralised bargaining for the purchase of drugs and medical goods in public hospitals, reduction in excessive consumption of specific drugs. Regarding in-patient care, shorter length of stay, development of ambulatory hospitalisation or more user choice among health providers could help reining in the expenditure growth. Regarding out-patient care, a more stringent health monitoring of patients affected by long-term diseases in order to reduce costly complications; financial incentives set on general practitioners to reduce prescription inflation and enhance prevention could also help in maintaining the health expenditure on a sustainable path.

Accordingly, a long-run convergence condition is considered in the cost-containment scenario. Specifically, the growth contribution of the spending residual, whose past growth was partly attributable to technology and relative price effects, is assumed to converge to zero in 2060, implicitly representing the assumption that policies are more effective than in the past in controlling the expenditure growth driven by non-demographic non-income related factors.

In this scenario (Table 4), the health expenditure to GDP ratio for OECD countries would increase by 2.5 percentage points, to reach close to 8% on average by 2060. The largest increases (above 3 percentage points of GDP) are found in Chile, Korea, Luxembourg, Mexico and Turkey. The lowest increases (below 2 percentage points of GDP) are found in Estonia, Hungary, Sweden and the United Kingdom. The increase in non-OECD countries averages around 2 percentage points, to reach 4.5% of GDP by 2060. The effect of non-demographic non-income factors is higher before 2030 than after for all countries, as the residual growth is assumed to converge to zero over the projection period. The difference between cost-pressure and cost-containment scenarios increases over time due to compounding effect of the residual.

2.9. Sensitivity analysis

In the sensitivity analysis, a number of parameters were changed in the context of the “cost-containment scenario”: the income elasticity, the residual spending growth, as well as factors underlying health status at old age (Table 5). Overall, the previous results seem relatively robust. Specifically, these alternative simulations do not change qualitatively the picture emerging from the comparison of the “cost-pressure” and “cost-containment” scenarios discussed above.

First, given that assumptions on the future evolution of the residual are necessarily arbitrary, their impact on projections was checked by computing a cost-containment scenario in line with the assumptions behind such a scenario in the previous 2006 OECD projections. In this variant (not shown in the table), on top of the demographic and income effects, the residual component is assumed to contribute 1% per year, tending towards 0% in 2060. On average, the health expenditure to GDP ratio across OECD countries reaches 6.8% instead of 7.9% in this new exercise. For non-OECD countries, the ratio will reach 3.3% instead of 4.4%. As the assumption underlying this scenario is common for all countries,

Table 5. **Sensitivity analysis on public health care projections**

Performed on the cost-containment scenario

	Income elasticity = 0.6	Income elasticity = 1	Country specific residual growth (non-income, non-demographic effect)	Compression of morbidity ¹	Expansion of morbidity ²
Percentage point deviations from cost-containment scenario in 2060					
Australia	-0.7	0.8	9.9	-0.6	0.6
Austria	-0.5	0.6	6.4	-0.7	0.9
Belgium	-0.6	0.7	0.2	-0.6	0.8
Canada	-0.6	0.7	1.0	-0.5	0.7
Chile	-0.8	1.1	0.0	-0.6	0.7
Czech Republic	-0.8	0.9	-4.0	-0.8	1.0
Denmark	-0.6	0.7	5.4	-0.7	0.8
Estonia	-0.9	1.1	0.0	-0.7	0.8
Finland	-0.6	0.7	4.0	-0.7	0.7
France	-0.5	0.6	-2.4	-0.6	0.8
Germany	-0.6	0.7	-0.7	-0.7	0.9
Greece	-0.5	0.6	2.9	-0.6	0.8
Hungary	-0.8	1.0	-1.9	-0.9	1.3
Iceland	-0.6	0.7	15.7	-0.5	0.6
Israel	-0.6	0.7	0.0	-0.5	0.7
Ireland	-0.4	0.5	0.0	-0.7	0.8
Italy	-0.5	0.6	-1.8	-0.6	0.7
Japan	-0.6	0.8	0.0	-0.6	0.7
Korea	-0.7	0.8	0.5	-0.7	0.8
Luxembourg	-0.2	0.2	-0.9	-0.7	0.9
Mexico	-0.8	1.1	0.0	-0.7	0.9
Netherlands	-0.6	0.7	-2.7	-0.6	0.8
New Zealand	-0.7	0.9	2.0	-0.6	0.6
Norway	-0.6	0.7	-1.6	-0.6	0.7
Poland	-0.7	0.9	-5.3	-1.0	1.3
Portugal	-0.6	0.7	8.7	-0.8	0.9
Slovak Republic	-0.8	0.9	-6.5	-1.0	1.4
Slovenia	-0.6	0.8	0.0	-0.7	0.7
Spain	-0.5	0.6	5.8	-0.7	0.8
Sweden	-0.6	0.7	-0.4	-0.6	0.7
Switzerland	-0.6	0.8	1.1	-0.6	0.8
Turkey	-0.8	1.1	0.0	-0.8	1.0
United Kingdom	-0.6	0.7	-2.3	-0.6	0.7
United States	-0.6	0.7	-1.6	-0.6	0.6
OECD average³	-0.6	0.8	0.9	-0.7	0.8
Brazil	-0.9	1.1	0.0	-0.8	1.0
China	-1.1	1.7	0.0	-0.8	1.1
India	-1.1	1.7	0.0	-0.8	1.2
Indonesia	-1.0	1.5	0.0	-0.9	1.3
Russian Federation	-0.8	1.0	0.0	-0.7	1.0
South Africa	-0.9	1.2	0.0	-0.6	1.4
Non-OECD average³	-1.0	1.4	0.0	-0.8	1.2
Total average³	-0.7	0.8	0.8	-0.7	0.9

1. Longevity gains are doubled into additional years in good health.

2. No healthy ageing adjustment, i.e. longevity gains do not translate into additional years in good health.

3. Unweighted average.

Source: Authors' calculations.

the percentage point deviation from cost-containment scenario is the same for all countries (-1.1). However, the assumptions behind this scenario may be overly optimistic: the lower value of the residual growth contribution in the 2006 exercise reflected in part a higher income elasticity which arithmetically left less of historical spending growth to be explained by the residual. Assuming both lower income elasticity than in the 2006 exercise and a similar residual spending component may be too much to hope for.

As regards the residual growth, we also tested the effect of assuming country-specific residuals. This test is equivalent to the IMF assumption (IMF, 2012) except that in our scenario the residual is assumed to converge to zero in 2060. On average, the health expenditure deviation from the base case will be 0.8 percentage points of GDP. But this moderate average change represents in fact very different evolutions across countries. This assumption would lead to rather extreme scenarios for some countries (e.g. Iceland, Australia, Portugal, Austria, Spain and Denmark). Such large increases in the health expenditure ratios are due to country-specific conditions in the base period, which are unlikely to persist in the future.

There is considerable uncertainty concerning the income elasticity of health spending. The sensitivity analysis is carried out taking plus and minus two standard deviations from the value estimated in the benchmark regressions, corresponding to an income elasticity equal to 0.6 and 1, respectively. When the income elasticity is set to 1, on average for OECD countries the health expenditure deviation from the base case will be 0.8 percentage points of GDP to reach 8.7%. It is slightly higher for non-OECD countries at 1.4 percentage points of GDP. In both groups of countries, those with the largest projected GDP per capita growth are obviously the most affected by changes in the income elasticity. When the income elasticity is set to 0.6, health expenditure deviates from the base case by -0.6 percentage points of GDP on average in OECD countries, to reach 7.3%. This deviation is much higher in non-OECD countries, amounting to -1 percentage point of GDP.

Sensitivity to alternative health status was also explored. In an “expansion of morbidity” scenario, the share of life spent in ill health would increase as life expectancy increases (e.g. an increase in the survival rates of sick people), while a “compression of morbidity” scenario would mean the opposite (e.g. an improvement in the health status and health behaviour of the new cohorts of old people). Under these scenarios, by 2060, average health expenditures in OECD countries range from 7.2 to 8.7% of GDP. In non-OECD countries, they range from 3.7% to 5.6%. Thus, while alternative health regimes matter for projecting future expenditure trends, their impact on projections is smaller than that of non-demographic effects.

3. Long-term care

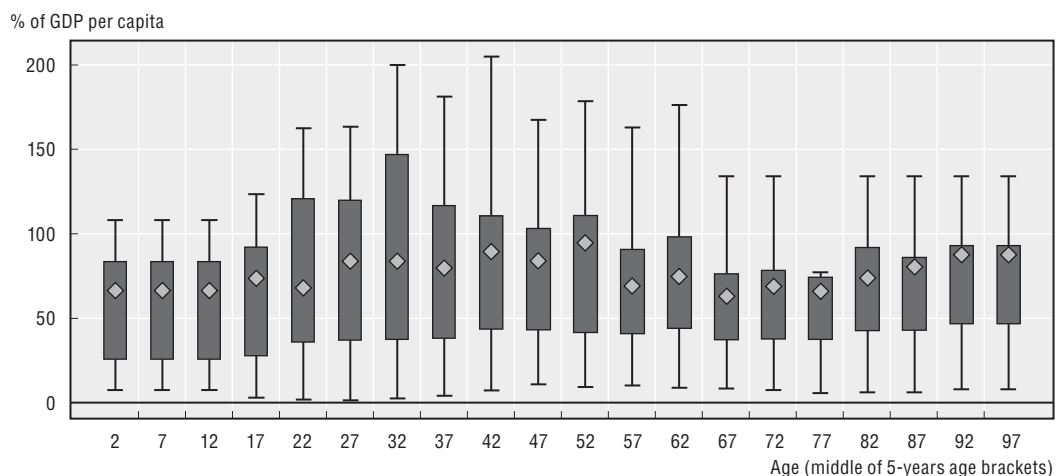
Long-term care differs radically from health care. While health care services aim at changing the health condition (from unwell to well), LTC merely aims at making the current condition (unwell) more bearable. Individuals need LTC due to disability, chronic condition, trauma, or illness, which limit their ability to carry out basic self-care or personal tasks that must be performed each day. Such activities are defined as activities of daily living (ADLs), which embrace eating, dressing, bathing, getting in and out of bed, toileting and continence or instrumental activities of daily living (IADLs), which include preparing own meals, cleaning, laundry, taking medication, getting to places beyond

walking distance, shopping, managing money affairs and using the telephone/internet. A person is dependent if he or she has limitations in ADLs or IADLs.

It should be noted that LTC also includes a small part of health services that are thus not accounted for in health care. Indeed, total LTC spending is calculated as the sum of health care and social services for those in LTC (Colombo et al., 2011). Health-related LTC spending includes palliative care, long-term nursing care, personal care services, and health services in support of family care. Social services provided for LTC include home help (e.g. domestic services) and care assistance, residential care services, and other social services. In other words, the health component of LTC spending includes episodes of care where the main need is either medical or personal care services (ADL support), while services whose dominant feature is help with IADL are considered outside the health-spending boundaries.

A striking difference between spending on health and LTC is that the cost of LTC per beneficiary is roughly independent of age (Figure 8). Indeed, the cost of helping one person in ADLs or IADLs could be more or less the same, irrespective of their age. Moreover, while potentially the entire population may benefit from health care, only dependent persons will benefit from LTC. Therefore, while the age-specific cost curve for health care was expressed per capita for each age group that for LTC is expressed per dependant.

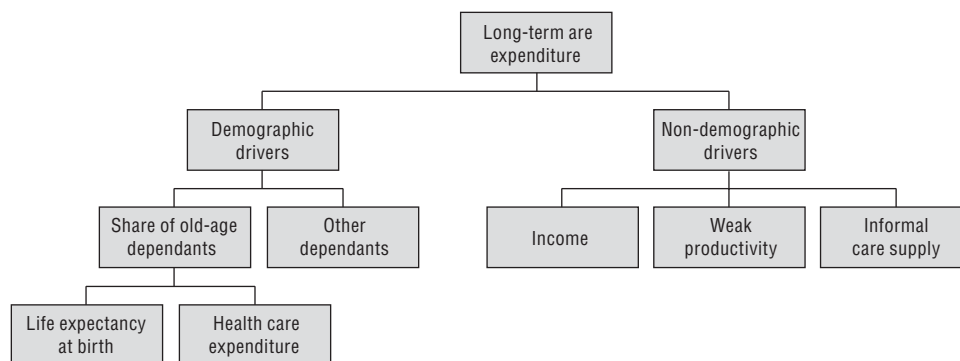
Figure 8. **Public long-term care expenditure per beneficiary as a % of GDP per capita¹**



1. The graph shows the dispersion of long-term care expenditure across countries by age groups. The diamonds represent the median. The boxes are the 2nd and 3rd quartiles of the distribution of expenditure across countries. The whiskers are the 1st and 4th quartiles.

Source: European Commission, 2009 Ageing Report: Economic and Budgetary Projections for the EU27 Member States (2008-60).

As for health care, two kinds of determinants drive LTC expenditure: demographic and non-demographic (Figure 9). The demographic driver is related to the number of dependent people in the population. The evolution of this factor depends on the evolution of life expectancy and health expenditure. The non-demographic drivers are related to income developments and changes in the demand for public-financed LTC services. Income is assumed to have a direct effect via increases in living standards (GDP per capita) and an indirect effect via cost-disease (relative productivity or Baumol) effects. Given the importance of home production of LTC services, the demand for public spending on LTC is assumed to depend on developments in formal labour force participation.

Figure 9. **The determinants of long-term care expenditure**

3.1. Projection framework

The framework used to project public LTC expenditure distinguishes between demographic and non-demographic drivers. The evolution of each driver is projected separately. They are then combined to compute the future growth of total LTC expenditure. The basic projection equation can be interpreted as a reduced form combining demand and supply factors that determine equilibrium LTC expenditures. More details on these assumptions and the corresponding variable definitions are provided below and in Annex A. They combine a demographic effect which in the case of LTC stands for the number of dependent people; a Baumol effect captured by the labour productivity differential between the LTC sector and the rest of the economy with an elasticity that can take the value of 1 or 0.5, depending on the scenario (0.5 or 0.25 for the BRIICS); an income effect with an elasticity of LTC expenditures that can take the value of 1 or 2; and the changes in the provision of informal care proxied by the evolution of labour force participation of women aged 50-64 with an elasticity of 0.7.

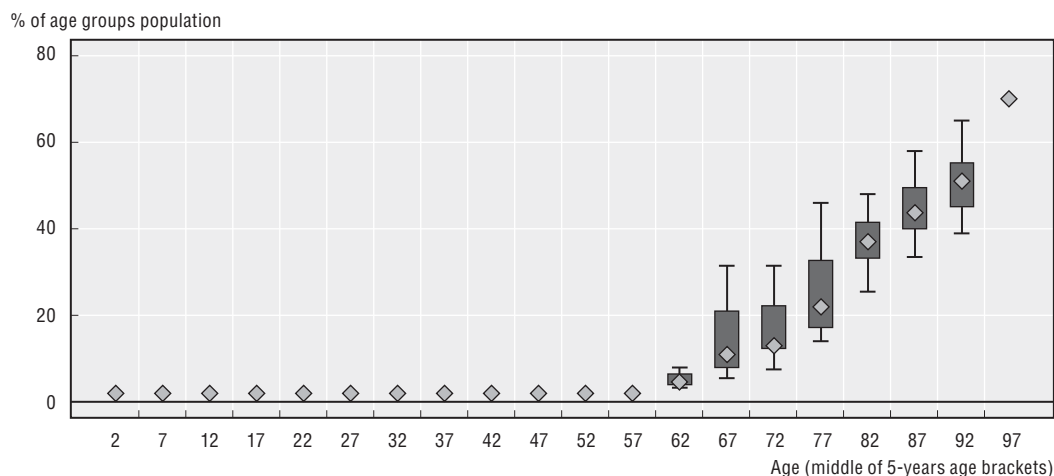
The growth rates are first projected for each country. Applying the same methodology as for health care, those growth rates are then adjusted to allow for a certain convergence of LTC expenditure to GDP ratio across countries. This allows taking into account a faster catching up of countries starting from a low basis. For instance, emerging economies such as China or Indonesia which will face a fast growing ageing population and as a consequence increasing demand for LTC might introduce LTC policies.

In order to smooth out the impact of the recent crisis on expenditure, the starting point for each country is computed as the average share of public LTC spending in GDP over 2006-10. It has to be noticed that this assumption may be too conservative. Indeed, contrary to health care, spending has increased recently in countries where public LTC is provided.⁸ Public LTC is expected to continue growing or to be introduced where it is not yet available.

3.2. Demographic drivers of expenditure

Whereas health care projections distinguished between survivors and non-survivors, the LTC projections split each age group into dependants and non-dependants. The age-specific dependency ratios have been estimated by the European Commission for 22 European countries (Ageing Report 2009). Even if the shape of the dependency ratios by age differs across countries, the dispersion remains relatively limited (Figure 10). Thus, as a first approximation, dependency ratios (number of dependants by age group) are assumed to be broadly uniform across countries. For the purpose of projecting

Figure 10. **Dependency ratios by age**¹
 Number of dependants as a % of population by age groups



1. The graph shows the dispersion of the dependency ratio across countries by age groups. The diamonds represent the median. The boxes are the 2nd and 3rd quartiles of the distribution of expenditure across countries. The whiskers are the 1st and 4th quartiles.

Source: European Commission, 2009 *Ageing Report: Economic and Budgetary Projections for the EU27 Member States (2008-60)*.

expenditure, this also has the advantage of eliminating current differences in the prevalence of dependency across countries as a possible cause for future differential increases in LTC expenditures. In other words, the projections become less sensitive to initial conditions. Noteworthy, this assumption therefore abstracts from the influence of differences in policy settings on these initial conditions.

For the calculation of the pure demographic effect, it is assumed that the LTC spending per dependant remains constant. Thus, the variation in LTC expenditure resulting from the pure demographic effect is only driven by the increase in the number of dependants by age group. The latter is derived from the average age-specific dependency ratio (see above) multiplied by the population by age group.

The number of dependent people in the population depends on the evolution of longevity and spending on health care. As the age-specific dependency ratio rises sharply after the age of 75, any increase in life expectancy above that threshold can increase significantly the number of old-age dependants, thereby putting pressure on LTC spending. Moreover, as health care spending improves the probability of survival at old-age, it can also push up LTC spending. This will be the case, for instance, if survival at older ages translates into an increase in the prevalence of chronic diseases. However, if improvements in life expectancy at birth translate into additional years in good health, the increase in dependency occurs later in life. Therefore the initial spending pressures from higher longevity are mitigated by such healthy ageing.

In order to project the evolution of dependency until 2060, taking into account the possible link between dependency and health-care expenditures, the age-specific dependency ratios have been estimated based on historical data as a function of age, age-specific per capita health-care expenditures and life expectancy at birth (see Box 2). Consistent with the healthy-ageing hypothesis, dependency at old age is found to decline over time. But the decline of the dependency ratio by age group depends in turn on the evolution of life expectancy at birth and per capita public health care expenditures (as projected in the cost-containment scenario for health care).

Box 2. Dependency ratio estimates

In order to gauge the evolution of the dependency ratio, its past determinants have been investigated by means of panel regression techniques. Defining $Depr_{i,a}$ as the dependency ratio (number of dependent people for country i and age a), Age_i as the central point in each age bracket (2, 7, 12, ..., 97), $he_{i,a}$ the real public health care expenditure per capita for country i and age a and LE_i life expectancy at birth for country i , the following equation was estimated:

$$\log(Depr_{i,a}) = \alpha + \varphi \cdot \log(Age_i) + \beta \cdot \log(he_{i,a}) + \gamma \cdot \log(LE_i) + u_{i,a}$$

The equation was estimated for the population aged 52 and above, as the dependency ratio for people below 52 is small and roughly constant over time. As expected, the age variable has a highly significant, positive impact (see de la Maisonneuve and Oliveira Martins, 2013). Public health expenditure per capita also has a significant positive effect, though much smaller. Conversely, increased life expectancy at birth delays the prevalence of dependency.

Including life expectancy at birth minimises the possibility that this effect could be driven by health expenditures themselves, thereby avoiding multicollinearity problems between health expenditure and life expectancy.

Accounting for the influence of health-care expenditure on the dependency ratio generates a link between the health-care and LTC projections even though the two spending items are projected separately. Clearly, this link may materialise with a lag, which is however difficult to ascertain empirically, especially in the context of the mostly cross-section data used in this paper. Therefore, a contemporaneous link is assumed for simplicity.

3.3. Non-demographic drivers of expenditure

Apart from the evolution of the number of dependants in the population, non-demographic drivers also have an impact on LTC expenditure growth. The projections account for three of them: changes in the relative price of LTC, income effects and changes in the demand for public-financed LTC, which in turn depend on the availability of informal care.

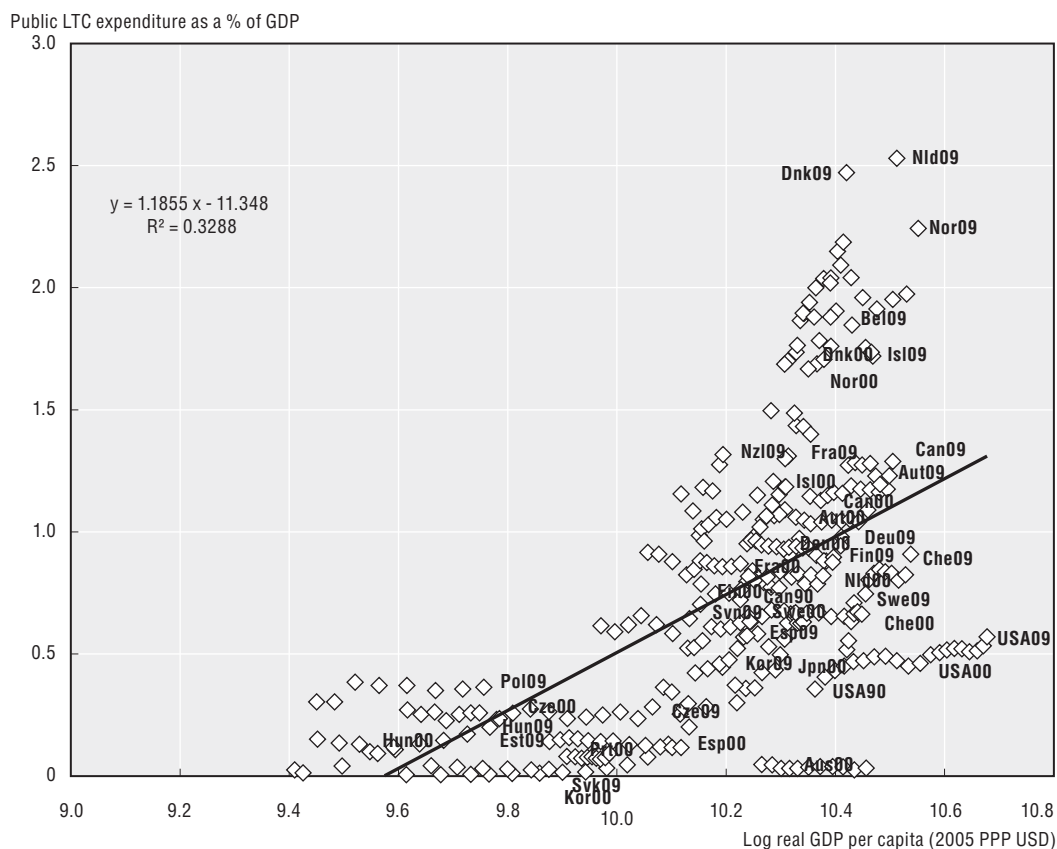
One of the main non-demographic drivers of public LTC expenditure is the relative share of informal and formal care.^{9, 10} Most informal care is provided by family and friends (Colombo et al., 2011). Even using a narrow definition of the family care “workforce”, its size is at least twice that of the formal care workforce (e.g. in Denmark), and in some cases it is estimated to be more than ten times the size of the formal-care workforce (e.g. Canada, New Zealand, the United States, the Netherlands). On average, around 70% to 90% of those who provide care are family carers (Fujisawa and Colombo, 2009). Changing societal models – such as declining family size, changes in residential patterns of people with disabilities and rising female participation in the formal labour market – are likely to contribute to a decline in the availability of informal care-givers, leading to an increase in the need for paid care (Colombo et al., 2011). Since there is evidence that informal elderly care is associated with lower female labour force participation (Viitanen, 2005), informal carers have been proxied by the labour force participation of women aged 50-64 to project the future evolution of LTC spending (a sensitivity check has also been carried out using their exit rate from the labour force). As participation rates by age and gender are not readily available for non-OECD countries, informal carers have been proxied by the overall participation rates in these countries.

Another important non-demographic driver of public LTC expenditure is a “cost disease” or Baumol effect (Baumol, 1967; 1993). The LTC sector is highly labour-intensive

and the room for productivity gains is often seen as more limited than elsewhere in the economy. With equalisation of wages across sectors, this implies that the relative price of LTC vis-à-vis other goods and services in the economy, tends to rise as aggregate productivity and GDP per capita increase. With a price-inelastic demand, the share of LTC expenditure in GDP would, therefore, tend to increase over time. A possible way to capture this effect is to assume that unit costs rise in line with average earnings of care staff or a measure of wage inflation in the economy (Comas-Herrera et al., 2003). In this paper, productivity growth in the total economy has been used as a proxy for the Baumol effect on public LTC expenditures.¹¹

Finally, income is also likely to play a role in LTC expenditure. Indeed, as real incomes rise, demand may be directed at higher quality services (Colombo et al., 2011), thereby increasing expenditures. Figure 11 shows a positive relationship between income and public LTC expenditure with an elasticity above one. It is difficult however to distinguish the income from the cost-disease effect, proxied by aggregate productivity, since they are too highly correlated. Given the multicollinearity between productivity and income, the two effects were estimated separately. The estimated influence of each of the non-demographic drivers of LTC expenditures has been used to parameterise the projection equation in Box 3.

Figure 11. **Long-term care expenditure and GDP per capita**



Box 3. Estimations to calibrate the LTC framework

In order to estimate the LTC spending elasticities to productivity (Baumol effect) and to the participation rate, which are used to parameterise the projections model, the following equation was estimated over the period 1990-2009:

$$\log\left(\frac{LTC}{P_Y \cdot Y}\right) = \alpha + \beta \cdot \log(OAdep) + \gamma \cdot \log(Prod) + \delta \cdot \log(PR) + u$$

where the share of long-term care expenditure in GDP [$LTC/(P_Y \cdot Y)$] is explained by *Prod*, the total economy productivity capturing the Baumol effect (or alternatively the income per capita variable to estimate the income elasticity) and *PR*, the female labour force participation rate used as a proxy for the provision of informal care (or alternatively their exit rate from employment was also used). In this equation *OAdep*, the ratio of people aged 80 and above to total population, is a control variable that plays no role in the projections, as a demographic effect already covers the effects of ageing.

As expected, the old age dependency ratio (people above 80) is a significant determinant of LTC expenditure (see de la Maisonneuve and Oliveira Martins, 2013). The relative price effect (proxied by total economy productivity) emerges with an elasticity of around 2. Alternatively, when the income variable is introduced, its elasticity amounts to around 2.7. While regression estimates point to Baumol or Income elasticities higher than unity, a conservative choice has been made of a unitary elasticity for both the relative price and the income variable. Sensitivity analysis has been carried out to test for the impact of these assumptions with the income elasticity alternatively fixed at two (see below). Finally all the proxies for informal care (female participation or exit rate from the labour force) also turned out to be significant.

3.4. Projection results: Long-term care, 2010-60

As for health care, a cost-pressure scenario and a cost-containment scenario were computed. Both scenarios are based on a unitary income elasticity assumption and the “healthy ageing” hypothesis. However, in the cost-pressure scenario, for OECD countries, a full Baumol effect is assumed, meaning that LTC unit labour costs increase fully in line with aggregate labour productivity; for non-OECD countries, excess labour supply especially in the non-tradeable sector suggests weaker wage pressures than in the OECD countries, and therefore the cost-pressure scenario only incorporates half of the Baumol effect. In the cost-containment scenario, the elasticity of LTC spending to productivity increases is set at half the value of the cost-pressure scenario (0.5 for OECD countries and 0.25 for non-OECD countries), possibly reflecting policy action aimed at mitigating relative wage increases of LTC providers. For example, action to curb expenditure could be aimed at facilitating access to LTC provision by low-skilled migrants or at providing incentives to balance institutional and home-based LTC.

Sensitivity analysis has been carried out in the context of the cost-containment scenario, with respect to the Baumol effect, the income elasticity and the implications of healthy ageing for the number of dependants. As for health care, the starting year of the projections is an average over the period 2006-10 so as to smooth out the impact of the recent crisis. The main assumptions underlying each projection scenario are listed in Table 6.

As with projections of public health-care spending, non-demographic drivers account for the lion’s share of future expenditure increases (Table 7), although with an assumed

Table 6. **Assumptions underlying the alternative projection scenarios: Public long-term care**

Scenarios	Health status	Income effect	Baumol effect	Availability of informal care
Demographic effect	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy and health expenditure.	n.a.	n.a.	n.a.
Cost-pressure scenario	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy and health expenditure (cost-containment scenario).	Income elasticity = 1	OECD countries: Full Baumol effect (elasticity = 1). BRIICS: Half Baumol effect.	OECD countries: Baseline scenario for participation rates of women aged 50-64. BRIICS: Total participation rates.
Cost-containment scenario	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy and health expenditure (cost-containment scenario).	Income elasticity = 1	OECD countries: Half Baumol effect (elasticity = 0.5). BRIICS: 1/4 of Baumol effect.	OECD countries: Baseline scenario for participation rates of women aged 50-64. BRIICS: Total participation rates
Sensitivity analysis based on the cost-containment scenario				
Income elasticity = 2	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy and health expenditure.	Income elasticity = 2	OECD countries: Half Baumol effect (elasticity = 0.5). BRIICS: 1/4 of Baumol effect.	OECD countries: Baseline scenario for participation rates of women aged 50-64. BRIICS: Total participation rates.
Life expectancy <i>plus</i> 2 standard deviation	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy plus 2 standard deviation and health expenditure.	Income elasticity = 1	OECD countries: Half Baumol effect (elasticity = 0.5). BRIICS: 1/4 of Baumol effect.	OECD countries: Baseline scenario for participation rates of women aged 50-64. BRIICS: Total participation rates.
Life expectancy <i>minus</i> 2 standard deviation	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy minus 2 standard deviation and health expenditure.	Income elasticity = 1	OECD countries: Half Baumol effect (elasticity = 0.5). BRIICS: 1/4 of Baumol effect.	OECD countries: Baseline scenario for participation rates of women aged 50-64. BRIICS: Total participation rates.
Cost-pressure health scenario	Healthy ageing: The prevalence of dependency per age is shifted according to the evolution of life expectancy and health expenditure (cost pressure scenario).	Income elasticity = 1	OECD countries: Half Baumol effect (elasticity = 0.5). BRIICS: 1/4 of Baumol effect.	OECD countries: Baseline scenario for participation rates of women aged 50-64. BRIICS: Total participation rates.

Note: The key assumption changed in each scenario is in bold.

elasticity of unity, the income effect is not creating additional pressures in terms of expenditure shares to GDP. Demographic changes are projected to exert a relatively minor influence on future public LTC expenditures. Indeed, as the LTC cost is independent of age, the pure age effect has only a moderate impact on spending except where ageing is fast enough to increase significantly the share of very-old people in the population (e.g. Korea, China, Indonesia). Moreover, this effect is mitigated by the “healthy ageing” assumption.

3.5. The cost-pressure scenario

In the cost-pressure scenario, starting from 0.8% of GDP on average for OECD countries, the ratio of public LTC expenditure to GDP is projected to increase by 1.4 percentage points to reach 2.1% of GDP. However, some countries, like Chile, Estonia, Korea, Mexico and Turkey, will experience higher increases (equal to or above 2 percentage points). These countries are

Table 7. **Projection scenarios for public long-term care expenditure**

	Average 2006-10	Percentage point deviations from starting period in 2030			Percentage point deviations from starting period in 2060		
		Demographic effect	Cost- pressure	Cost- containment	Demographic effect	Cost- pressure	Cost- containment
Australia	0.0	0.1	0.6	0.3	0.3	1.3	0.8
Austria	1.1	0.1	0.5	0.3	0.2	1.0	0.7
Belgium	1.7	0.1	0.5	0.3	0.1	1.2	0.7
Canada	1.2	0.2	0.5	0.3	0.2	1.2	0.7
Chile	0.0	0.2	0.9	0.6	0.6	2.3	1.5
Czech Republic	0.3	0.2	0.7	0.3	0.4	1.6	0.9
Denmark	2.2	0.1	0.4	0.2	0.2	1.0	0.6
Estonia	0.2	0.1	0.7	0.3	0.3	2.0	0.9
Finland	0.8	0.2	0.5	0.3	0.2	1.0	0.5
France	1.1	0.1	0.5	0.2	0.2	1.0	0.6
Germany	0.9	0.2	0.5	0.3	0.3	1.1	0.7
Greece	0.5	0.1	0.6	0.4	0.3	1.4	0.9
Hungary	0.3	0.1	0.6	0.3	0.3	1.8	1.0
Iceland	1.7	0.2	0.3	0.1	0.2	1.0	0.5
Israel	0.5	0.2	0.4	0.3	0.3	0.9	0.6
Ireland	0.4	0.1	0.5	0.4	0.2	1.0	0.7
Italy	0.7	0.1	0.5	0.3	0.2	1.2	0.8
Japan	0.7	0.2	0.5	0.3	0.3	1.4	0.8
Korea	0.3	0.4	0.8	0.5	0.8	2.0	1.3
Luxembourg	0.9	0.1	0.5	0.4	0.3	0.8	0.7
Mexico	0.0	0.2	0.7	0.5	0.6	2.2	1.5
Netherlands	2.3	0.2	0.6	0.4	0.3	1.4	0.8
New Zealand	1.3	0.2	0.5	0.3	0.3	1.4	0.8
Norway	2.1	0.1	0.4	0.2	0.2	1.0	0.5
Poland	0.4	0.2	0.7	0.3	0.5	1.8	1.0
Portugal	0.1	0.1	0.4	0.2	0.3	1.3	0.8
Slovak Republic	0.0	0.2	0.8	0.4	0.6	1.9	1.1
Slovenia	0.7	0.2	0.6	0.4	0.4	1.5	0.9
Spain	0.5	0.1	0.6	0.5	0.3	1.5	1.0
Sweden	0.7	0.1	0.4	0.1	0.1	0.9	0.5
Switzerland	1.2	0.1	0.5	0.3	0.3	1.3	0.7
Turkey	0.0	0.2	0.7	0.3	0.6	2.3	1.6
United Kingdom	0.9	0.1	0.3	0.2	0.1	0.9	0.5
United States	0.6	0.1	0.4	0.2	0.2	0.8	0.4
OECD average¹	0.8	0.1	0.5	0.3	0.3	1.4	0.8
Brazil	0.0	0.2	0.3	0.3	0.7	1.2	0.9
China	0.1	0.3	0.4	0.3	0.7	2.0	1.1
India	0.1	0.1	0.3	0.2	0.4	1.5	0.8
Indonesia	0.1	0.2	0.3	0.2	0.7	1.6	1.0
Russian Federation	0.2	0.1	0.2	0.1	0.2	0.8	0.4
South Africa	0.2	0.1	0.2	0.2	0.3	0.7	0.5
Non-OECD average¹	0.1	0.2	0.3	0.2	0.5	1.3	0.8
Average¹	0.7	0.2	0.5	0.3	0.3	1.4	0.8

1. Unweighted average.

starting from a very low level of LTC expenditure and are in the process of catching up to higher levels. Conversely, Israel, Luxembourg, Sweden, the United Kingdom and the United States are projected to experience only small increases in their LTC expenditure to GDP ratios (less than 1 percentage point). Differences in the initial level of female participation rates as well as policies in place also explain the differences across countries.

For non-OECD countries, the increase in LTC spending is projected to be, on average, broadly the same as that experienced by OECD countries but starting from a much lower level (0.1%).

3.6. The cost-containment scenario

In the cost-containment scenario, on average for OECD countries, the public LTC expenditure to GDP ratio increases by 0.8 percentage point from 0.8% to 1.6% of GDP. Chile, Korea, Mexico and Turkey experience the most important increases (above 1.3 percentage points), while the increase in Finland, Iceland, Norway, Sweden, the United Kingdom and the United States does not exceed 0.5 percentage points. Non-OECD countries will experience on average the same increase as OECD countries (0.8 percentage points).

3.7. Sensitivity analysis

Given the uncertainty about income effects, projections were performed also assuming an income elasticity of 2 for the cost-containment scenario for OECD countries. For non-OECD countries, the sensitivity analysis assumed an income elasticity of 1.5. On average, for OECD countries LTC expenditure would reach in this case 2.5% of GDP in 2060. In this scenario, Chile, Estonia, Mexico and Turkey would experience a deviation from base case of around 1.7 percentage points of GDP. Again, the assumptions underlying the GDP projections assume some catch-up for these low income countries, explaining why their LTC expenditure patterns are more affected by the higher income elasticity assumption. For non-OECD countries, the share of LTC in GDP would reach 1.8% of GDP with the higher income elasticity. China, India and Indonesia would experience a deviation from the base case of more than 1 percentage point (see Table 8).

A sensitivity analysis has also been carried out for the evolution of the number of dependants. The two drivers of the dependency ratio, life expectancy and health care spending, have been changed separately relative to baseline. Taking life expectancy *plus* (minus) two standard deviations would induce an average deviation from the base case for OECD countries of 0.3 (-0.2) percentage points in the LTC expenditure to GDP ratio, which would then reach, on average, 1.9% (1.4%) of GDP. For each scenario, the expenditure for non-OECD countries will deviate from the base case by the same amount as for OECD countries. Due to the very limited impact of health expenditure, alternative assumptions in this area have only a slight impact. Indeed, even under the cost-pressure health care spending scenario, the LTC expenditure to GDP ratio would only deviate from the base case by 0.1 percentage points on average for both OECD and non-OECD countries.

4. Prospects for total public spending on health

Both health and LTC expenditures are projected to put sizeable pressure on public finance over the next 50 years. Overall, on average across OECD countries, the total health and LTC expenditure to GDP ratio is projected to more than double in the cost-pressure scenario, increasing from 6.2% in the starting period to 13.9% of GDP in 2060. In the cost-containment scenario the ratio would still increase by more than half, to reach 9.5% (Figure 12). For the BRIICS, starting from a much lower level, public health and LTC expenditure will increase from 2.5% in the starting period to about 10% and above 5% in 2060 in the cost-pressure and the cost-containment scenarios, respectively.

There are, however, striking differences across countries (Figure 13). In the cost-containment scenario, one group of countries experiences a 4 to 5 percentage point increase in total spending to GDP ratios over the period 2010-60. It includes countries that are ageing very rapidly (Korea and Chile), as well as those with currently low labour force

Table 8. **Sensitivity analysis on public long-term care projections**

Performed on the cost-containment scenario

	Income elasticity = 2 ¹	Life expectancy <i>plus</i> 2 standard deviation	Life expectancy <i>minus</i> 2 standard deviation	Cost-pressure health expenditure
Percentage point deviations from cost-containment scenario in 2060				
Australia	1.0	0.2	-0.2	0.1
Austria	0.6	0.2	-0.2	0.1
Belgium	0.8	0.2	-0.2	0.1
Canada	0.8	0.2	-0.2	0.1
Chile	1.7	0.3	-0.2	0.1
Czech Republic	1.3	0.3	-0.2	0.1
Denmark	0.8	0.2	-0.2	0.1
Estonia	1.8	0.3	-0.2	0.1
Finland	0.8	0.2	-0.2	0.1
France	0.6	0.2	-0.2	0.1
Germany	0.8	0.3	-0.2	0.1
Greece	0.6	0.3	-0.2	0.1
Hungary	1.4	0.3	-0.2	0.1
Iceland	0.8	0.2	-0.2	0.1
Israel	0.7	0.2	-0.2	0.1
Ireland	0.5	0.2	-0.2	0.1
Italy	0.6	0.2	-0.2	0.1
Japan	0.9	0.3	-0.2	0.1
Korea	1.0	0.4	-0.3	0.2
Luxembourg	0.1	0.3	-0.2	0.1
Mexico	1.7	0.3	-0.3	0.1
Netherlands	0.8	0.3	-0.2	0.1
New Zealand	1.2	0.2	-0.2	0.1
Norway	0.8	0.2	-0.2	0.1
Poland	1.1	0.3	-0.2	0.1
Portugal	0.7	0.3	-0.2	0.1
Slovak Republic	1.3	0.3	-0.3	0.1
Slovenia	0.9	0.3	-0.2	0.1
Spain	0.7	0.3	-0.2	0.1
Sweden	0.9	0.2	-0.2	0.1
Switzerland	0.9	0.3	-0.2	0.1
Turkey	1.7	0.3	-0.2	0.1
United Kingdom	0.8	0.2	-0.2	0.1
United States	0.7	0.2	-0.2	0.1
OECD average²	0.9	0.3	-0.2	0.1
Brazil	0.6	0.4	-0.3	0.1
China	1.2	0.4	-0.3	0.1
India	1.3	0.2	-0.2	0.1
Indonesia	1.0	0.3	-0.3	0.1
Russian Federation	0.5	0.2	-0.2	0.1
South Africa	0.7	0.2	-0.1	0.1
Non-OECD average²	0.9	0.3	-0.2	0.1
Average²	0.9	0.3	-0.2	0.1

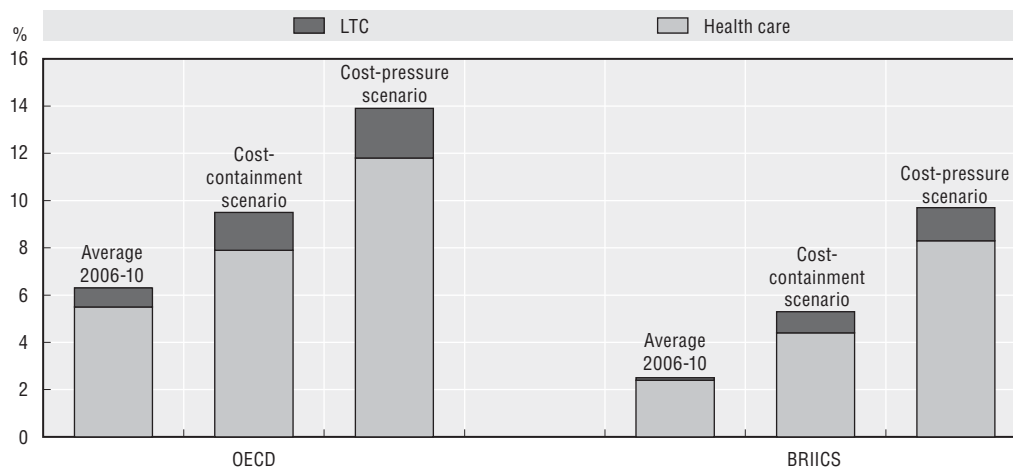
1. Income elasticity = 1.5 for the BRIICS.

2. Unweighted average.

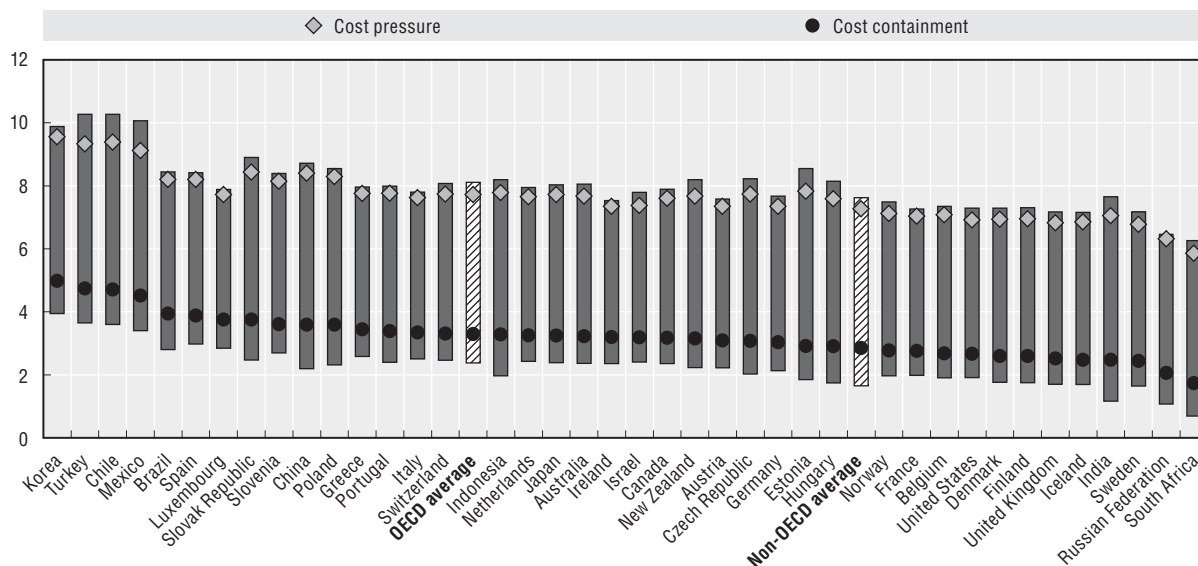
participation especially of women, which may face a substantial increase in the demand for formal long-term care (Spain and Mexico) as labour force participation increases and care is no longer provided by family. By contrast, in the United States, the United Kingdom and the Nordic countries the projected increase is below 3 percentage points of GDP. In

Figure 12. **Projected public health and long-term care expenditure**

As a % of GDP, in 2060



these countries demographic trends will be more stable and the share of formal long-term care is already relatively high. On average, spending ratios in the BRICS will remain lower than in OECD countries. However, they will rise significantly under both scenarios. For instance, even in the cost-containment scenario, ratios will more than double by 2060 – close to levels currently observed in OECD countries (Table 9).

Figure 13. **Percentage point increase in the total public health and long-term care spending ratio to GDP, 2010-60**Range of estimates across sensitivity analyses¹

1. The vertical bars correspond to the range of the alternative scenarios, including sensitivity analysis. Countries are ranked by the increase of expenditures between 2010 and 2060 in the cost-containment scenario.

These new projections broadly deliver the same message as those in Oliveira Martins and de la Maisonneuve (2006); even in the cost-containment scenario, public health care and LTC spending will put an important burden on future public finances. The new set of

Table 9. **Projection scenarios for public health and long-term care spending**
In % of GDP

	Health care			Long-term care			Total		
	Average 2006-10	2060 cost-pressure	2060 cost-containment	Average 2006-10	2060 cost-pressure	2060 cost-containment	Average 2006-10	2060 cost-pressure	2060 cost-containment
Australia	5.6	12.0	8.1	0.0	1.4	0.8	5.7	13.3	8.9
Austria	6.6	12.9	9.1	1.1	2.2	1.8	7.8	15.1	10.8
Belgium	5.8	11.6	7.7	1.7	3.0	2.5	7.5	14.6	10.2
Canada	5.8	12.2	8.3	1.2	2.5	1.9	7.1	14.6	10.2
Chile	3.1	10.2	6.3	0.0	2.3	1.5	3.1	12.4	7.8
Czech Republic	5.5	11.6	7.7	0.3	1.9	1.1	5.8	13.5	8.9
Denmark	6.3	12.2	8.3	2.2	3.3	2.8	8.5	15.4	11.1
Estonia	4.3	10.1	6.2	0.2	2.2	1.1	4.5	12.3	7.4
Finland	5.2	11.2	7.3	0.8	1.8	1.3	6.0	13.0	8.6
France	7.4	13.5	9.6	1.1	2.1	1.7	8.5	15.5	11.3
Germany	7.3	13.5	9.6	0.9	2.1	1.6	8.2	15.5	11.2
Greece	5.4	11.8	7.9	0.5	1.9	1.5	5.9	13.7	9.3
Hungary	4.8	10.6	6.7	0.3	2.0	1.2	5.0	12.6	7.9
Iceland	5.8	11.7	7.8	1.7	2.7	2.2	7.6	14.4	10.0
Israel	4.0	10.4	6.6	0.5	1.5	1.2	4.5	11.9	7.7
Ireland	5.5	11.9	8.0	0.4	1.3	1.1	5.9	13.3	9.1
Italy	6.1	12.6	8.7	0.7	1.9	1.5	6.9	14.5	10.2
Japan	6.1	12.5	8.6	0.7	2.0	1.4	6.8	14.5	10.0
Korea	3.3	10.9	7.0	0.3	2.3	1.6	3.6	13.2	8.6
Luxembourg	4.4	11.3	7.4	0.9	1.7	1.7	5.3	13.0	9.1
Mexico	2.5	9.4	5.5	0.0	2.2	1.5	2.5	11.6	7.0
Netherlands	6.4	12.7	8.8	2.3	3.7	3.1	8.7	16.4	12.0
New Zealand	6.4	12.7	8.8	1.3	2.6	2.0	7.7	15.3	10.8
Norway	5.1	11.2	7.3	2.1	3.1	2.7	7.2	14.3	10.0
Poland	4.1	10.5	6.7	0.4	2.2	1.4	4.5	12.7	8.0
Portugal	6.5	13.0	9.1	0.1	1.4	0.9	6.6	14.3	9.9
Slovak Republic	5.4	11.9	8.0	0.0	1.9	1.1	5.4	13.9	9.2
Slovenia	5.2	11.8	7.9	0.7	2.2	1.5	5.8	14.0	9.4
Spain	5.6	12.3	8.5	0.5	2.0	1.6	6.1	14.3	10.0
Sweden	6.6	12.4	8.6	0.7	1.6	1.1	7.2	14.0	9.7
Switzerland	5.7	12.2	8.3	1.2	2.5	1.9	6.9	14.6	10.2
Turkey	3.8	10.9	7.0	0.0	2.3	1.6	3.8	13.2	8.6
United Kingdom	6.5	12.4	8.5	0.9	1.8	1.4	7.4	14.2	9.9
United States	7.1	13.2	9.3	0.6	1.3	1.0	7.6	14.5	10.3
OECD average¹	5.5	11.8	7.9	0.8	2.1	1.6	6.2	13.9	9.5
Brazil	3.7	10.7	6.8	0.0	1.3	0.9	3.7	11.9	7.7
China	1.9	8.3	4.4	0.1	2.1	1.2	2.0	10.4	5.6
India	1.2	6.7	2.8	0.1	1.6	0.9	1.3	8.3	3.7
Indonesia	1.2	7.3	3.5	0.1	1.7	1.0	1.2	9.0	4.5
Russian Federation	3.1	8.6	4.7	0.2	1.0	0.7	3.4	9.7	5.4
South Africa	3.2	8.4	4.5	0.2	0.9	0.6	3.4	9.2	5.1
Non-OECD average¹	2.4	8.3	4.4	0.1	1.4	0.9	2.5	9.8	5.3
Total average¹	5.0	11.3	7.4	0.7	2.0	1.5	5.7	13.3	8.9

1. Unweighted average.

Source: Authors' calculations.

projections suggests a more important increase in health care expenditure, both in the cost-pressure and the cost-containment scenarios while, regarding public LTC funding, the increase in expenditure is lower in the new set of projections than in the previous one.

Overall, on average, the current projections suggest only a slightly lower increase in total health and LTC expenditure confirming the message delivered by the previous set of projections. A comparison between these new projections and the ones from the European Commission, the IMF and the Congressional Budget Office in the United States suggests that the differences arise mainly as a consequence of varying assumptions concerning the future evolution of the main drivers of spending (see de la Maisonneuve and Oliveira Martins, 2013 for details). The EC projected increases in health-care expenditure are the lowest due to the nearly unitary income elasticity associated with the absence of a residual growth. The IMF projections, although difficult to compare as they end in 2050, seem to range between the OECD cost-containment and cost-pressure scenarios. The assumption of low income elasticity is broadly offset by the country specific residual. Nonetheless when the estimated residual is higher than the average like in the United States, the projected increase in health care expenditure is much higher. This high increase is corroborated by the CBO projections.

5. Concluding remarks

Long-term spending projections are inherently uncertain and subject to upside or downside risks. While a more moderate evolution of spending than in these projections cannot be excluded (for instance if cost-saving technologies were to spread out, or if more aggressive cost-containment policies were to be implemented), there are also clear upside risks on spending. For instance, higher health spending could arise due to an extension of the pre-death period of ill health as longevity increases, or because of higher than expected costs induced by technical progress. Regarding LTC, higher spending could arise from increased dependency due to obesity trends or dementia. Indeed, according to recent calculations, some 12% of those aged between 80 and 84 years, and almost one in four of those aged over 85 years, suffer from dementia (Alzheimer Europe, 2006). With ageing populations, strong increases in the prevalence of dementia may be expected (see Box 1.1 in Colombo et al., 2011), though prevention and treatment may also improve in the future.

Even if these upside risks do not materialise, the spending projections point to important policy challenges. These challenges are reinforced by the evidence that macroeconomic cost-containment policies, which had some success in repressing spending trends over the 1980s and 1990s, have their limits. For instance, it is difficult to contain wages and at the same time, attract young and skilled workers into the health-care system. Similarly, controlling prices is not easy when technical progress is permanently creating new products and treatments, while overall constraints on supply result in unpopular waiting lists for these treatments. More generally, it is difficult to determine the appropriate supply of health and LTC services without market signals – but at the same time, health and LTC are areas where market failure is rife.

Notes

1. To focus on the structural factors and eliminate cyclical effects, the ratio is computed using trend instead of actual GDP (from the *OECD Economic Outlook*, No. 91).
2. Tax expenditure can be of different kinds, e.g. exclusion from workers' taxable income of employers' health insurance contributions in the United States, tax credits for medical expenses in Canada or income tax deductions for health care services in Portugal.
3. For more details, see Johansson et al. (2012).

4. While the probability of dying increases with age, the costs of death tend to decline steadily after young and prime age (Aprile, 2004). Other estimates of DRC are coherent with this order of magnitude (e.g. Yang et al., 2003).
5. See, for example Fuchs (1984); Zweifel et al. (1999); Jacobzone (2003); and Gray (2004).
6. While he/N and Y/N are found to be $I(1)$ and co-integrated, the complete VECM with all explanatory variables did not provide good results. Thus a regression on growth rates addresses the problem of a possible spurious correlation between the two variables. The results presented in the last column of Table 2 confirm the level estimates.
7. It has to be noted that depending on the type of expenditure this upward shift may be non-homothetic across ages. For example, expenditure at older ages may be more affected by such an upward shift than at younger ones.
8. For instance, Korea introduced the public LTC insurance in 2008 and it has increased since then. The same occurred in Estonia, Japan and Portugal.
9. Formal care is provided by care assistants who are paid for providing care under some form of employment contract. It includes both care provided in institutions and care provided at home. To be considered informal, the provision of care cannot be paid for as if purchasing a service. However, an informal care-giver may still receive social transfers conditional on his/her provision of informal care and possibly, in practice, some informal payment from the person receiving care.
10. The projections do not distinguish between formal care delivered within institutions and that delivered to the patient's home. There are fundamental differences between countries in the way they organise their formal LTC. Institutional LTC is particularly widespread in the Nordic countries (OECD, 2005). Whether this form of organisation is adopted by other countries or a (cheaper) ambulatory help-at-home strategy is pursued could have important consequences for public expenditures.
11. The relative price effect may be limited by a growing share of immigrants among LTC workers. According to Colombo et al., (2011), foreign-born workers play a significant and growing role in LTC in some countries. The average wage of the immigrant work force is lower than that of native workers and their bargaining power is weaker. The process of equalising the wages of foreign-born and native work forces will take time, but will certainly materialise over the long run.

References

- Acemoglu, D., A. Finkelstein and M. Notowidigdo (2009), "Income and Health Spending: Evidence from Oil Shocks", CEPR Discussion Papers, No. 7255.
- Alzheimer Europe (2006), *Dementia in Europe Yearbook 2006*, Luxembourg.
- Aprile, R. (2004), "How to Take Account of Death-Related Costs in Projecting Health Care Expenditure – Updated Version", *Ragioneria Generale Dello Stato*.
- Baumol, W.J. (1967), "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis", *American Economic Review*, 57, pp. 415-426.
- Baumol, W.J. (1993), "Health Care, Education and the Cost of Disease: A Looming Crisis for Public Choice," *Public Choice*, 77, pp. 17-28.
- Bessen, J. and T. Grid (2013), "Which Patent Systems Are Better For Inventors?"
- Breyer, F. and S. Felder (2006), "Life Expectancy and Health Care Expenditures: A New Calculation for Germany Using the Costs of Dying", *Health Policy*, 75, pp. 178-186.
- Breyer, F., J. Costa-i-Font and S. Felder (2011), "Does Ageing Really Affect Health Expenditures? If So, Why?", *Vox EU*, May.
- Burniaux, J.M., R. Duval and F. Jaumotte (2003), "Coping with Ageing: A Dynamic Approach to Quantify the Impact of Alternative Policy Options on Future Labour Supply in OECD Countries", *OECD Economics Department Working Papers*, No. 371, OECD Publishing.
- Colombo, F. et al., (2011), "Help Wanted? Providing and Paying for Long-Term Care", *OECD Health Policy Studies*, OECD Publishing.
- Comas-Herrera and R. Wittenberg (eds.) (2003), *European Study of Long-Term Care Expenditure: Investigating the Sensitivity of Projections of Future Long-Term Care Expenditure in Germany, Spain, Italy and the United Kingdom to Changes in Assumptions About Demography, Dependency, Informal Care, Formal Care and Unit Costs*, PSSRU, LSE Health and Social Care, London School of Economics.

- Congressional Budget Office (2012), *The 2012 Long-Term Budget Outlook*.
- De la Maisonneuve, C. and J. Oliveira Martins (2013), "A Projection Method for Public Health and Long-Term Care Expenditures", *OECD Economics Department Working Papers*, No. 1048.
- European Commission, 2009 Ageing Report: Economic and Budgetary Projections for the EU27 Member States (2008-60).
- Felder, S., M. Meier and H. Schmitt (2000), "Health Care Expenditure in the Last Months of Life", *Journal of Health Economics*, 19, pp. 679-95.
- Fuchs, V. (1984), "'Though Much is Taken' – Reflections on Ageing, Health and Medical Care", *NBER Working Paper*, No. 1269.
- Fujisawa, R. and F. Colombo (2009), "The Long-Term Care Workforce: Overview and Strategies to Adapt Supply to a Growing Demand", *OECD Health Working Paper*, No. 44, OECD Publishing.
- Getzen, T. (2000), "Health Care is an Individual Necessity and a National Luxury: Applying Multilevel Decision Models to the Analysis of Health Care Expenditure", *Journal of Health Economics*, 19, pp. 259-270.
- Gray, A. (2004), "Estimating the Impact of Ageing Populations on Future Health Expenditures", Public lecture to the National Institute of Economics and Business and the National Institute of Health and Human Science, 4 November, Canberra.
- Holly, A., X. Ke and P. Saksena (2011), "The Determinants of Health Expenditure: A Country-Level Panel Data Analysis", *World Health Organisation Working Paper*.
- International Monetary Fund (2012), "The Economics of Public Health Care Reform in Advanced and Emerging Economies".
- Jacobzone, S. (2003), "Ageing and the Challenges of New Technologies: Can OECD Social and Health Care Systems Provide for the Future?", *The Geneva Papers on Risk and Insurance*, Vol. 28, No. 2, April, pp. 254-74.
- Johansson, A. et al. (2012), "Long-Term Growth Scenarios", *OECD Economics Department Working Papers*, No. 1000, OECD Publishing.
- OECD (2005), *Long-Term Care for Older People*, OECD Publishing.
- OECD (2006), "Projecting OECD Health and Long-Term Care Expenditures: What are the Main Drivers?", *OECD Economics Department Working Papers*, No. 477, OECD Publishing.
- OECD (2010), *Health Care Systems: Efficiency and Policy Settings*, OECD Publishing.
- OECD (2011), *Health Data*, OECD Publishing.
- OECD (2012), *OECD Economic Outlook*, Vol. 2012/1, OECD Publishing.
- OECD (2013), *Health Data*, OECD Publishing.
- Okunade, A.A. and V.N.R. Murthy (2002), "Technology as a Major Driver of Health Costs: A Cointegration Analysis of the Newhouse Conjecture", *Journal of Health Economics*, 21, pp. 147-159.
- Oliveira Martins, J. and C. de la Maisonneuve (2006), "The Drivers of Public Expenditure on Health and Long-Term Care: An Integrated Approach", *OECD Economic Studies*, No. 43, 2006/2.
- Seshamani, M. and A. Gray (2004), "A Longitudinal Study of the Effects of Age and Time to Death on Hospital Costs", *Journal of Health Economics*, 23, pp. 217-35.
- Viitanen, T.K. (2005), "Informal Elderly Care and Female Labour Force Participation Across Europe", Center for European Policy Studies, ENEPRI Research Reports, No. 13, 1, July.
- Werblow, A., S. Felder and P. Zweifel (2007), "Population Ageing and Health Care Expenditure: A School of 'Red Herrings'?", *Health Economics*, 16, pp. 1109-26.
- Yang, Z., E.C. Norton and S.C. Stearns (2003), "Longevity and Health Care Expenditures: The Real Reasons Older People Spend More", *Journal of Gerontology*, Vol. 58B, No. 1, pp. S2-S10.
- Zweifel, P., S. Felder and M. Meiers (1999), "Ageing Of Population And Health Care Expenditure: A Red Herring?", *Health Economics*, 8, pp. 485-496.

ANNEX A

Data sources and method

Health and LTC data

Health and LTC expenditures

The historical public current health and LTC expenditure to GDP ratios come from the *OECD Health Database* (2011) that provides comparable health statistics across OECD countries. Total public current expenditure on health is the sum of total personal and total collective services but excludes investment. The LTC spending corresponds to the “Services of long-term nursing care” item. The latter is subtracted from the Total public current expenditure to get the health spending.

Total expenditure on health is defined as the sum of expenditure on activities that – through application of medical, paramedical, and nursing knowledge and technology – are aimed at: promoting health and preventing disease; curing illness and reducing premature mortality; caring for persons affected by chronic illness who require nursing care; caring for persons with health-related impairments, disability, and handicaps who require nursing care; assisting patients to die with dignity; providing and administering public health; providing and administering health programmes, health insurance and other funding arrangements. Public expenditure on health care is health expenditure incurred by public funds. Public funds are state, regional and local Government bodies and social security schemes.

Long-term health care comprises ongoing health and nursing care given to in-patients who need assistance on a continuing basis due to chronic impairments and a reduced degree of independence in activities of daily living. In-patient long-term care is provided in institutions or community facilities. Long-term care is typically a mix of medical (including nursing care) and social services.

Cost curves by age groups

The public health expenditure by age groups have been provided by the European Commission for the European countries (2009 *Ageing Report*) and by the Swiss Federal Statistical Office, Health Canada and the Australian Institute for Health and Welfare for Switzerland, Canada and Australia, respectively. These cost curves are split into survivors and non-survivors as explained in the text.

An average expenditure curve for survivors was estimated econometrically in a panel of 23 countries and 20 age groups, using a spline function as follows:

$$\frac{he_{age\ group}}{N_{age\ group}} = -199.1.age + 14.5.age^2 - 0.41.age^3 + 0.005.age^4 - 0.00002.age^5$$

where he is the real health expenditure by age group, N population by age group and age is the central point in each bracket (e.g. 2, 7, 12, ..., 97).

For each country, this expenditure curve by age group is multiplied by country-specific population by age group in order to obtain a total country-specific health expenditure. This expenditure is then calibrated to fit with the OECD database total expenditure in the base year.

LTC Dependency ratios

The dependency ratios (number of dependants by age group) were derived from the European Commission 2009 *Ageing Report*. Even if the shape of the dependency ratios by age differs across countries, the dispersion remains relatively limited (Figure 10 in the main text). Accordingly, dependency ratios are assumed to be broadly uniform across countries and an average dependency ratio by age group was calculated and used in the projections. The dependency ratios are then projected using the evolution of life expectancy at birth and health expenditure (cost-containment scenario). The life expectancies at birth are those underlying official population projections.

Macro data

GDP

GDP projections are published in the *OECD Economic Outlook*, No. 91. More details regarding the sources and the methodology can be found in Johansson et al. (2012). In short, GDP projections are based on a conditional convergence model.

Population

Population projections come from Eurostat for European Countries (EUROPOP2010) and from the United Nations for non-European Countries (*World Population Prospects: The 2010 Revision*).

Health prices and technology

The health price measure used in this paper is the value-added deflator in the health and social work sectors, taken from the *OECD STAN Database*. Even if this measure may not be as accurate as needed, it is nonetheless more comparable across countries than more precise price indicators available from other sources. For instance, CPI series for the health sector can be gathered from Eurostat (for European countries) and national sources (for Australia, Canada, Korea, New Zealand and the United States). The CPI series cannot, however, be used because they are available for too short a period (the HICP from Eurostat is only available from 1996 onwards) and there is also a potential lack of comparability across countries.

The quality effect combines an indicator of patents with an indicator of R&D statistics. The patents granted by the European Patent Office (EPO) have been chosen as they seem to generate substantially greater market power than their US equivalents (Bessen, 2013) and may represent the OECD-wide technology frontier (see de la Maisonneuve and

Oliveira Martins, 2013). They come from the OECD database on Patents. To benefit from advances in this frontier, countries need to innovate and absorb foreign technology (via technology pass-through and catching up effects). While not all R&D expenditures are medical, they may advance health care technology because of externalities (Okunade and Murthy, 2002); moreover, high R&D spending is needed to enable the adoption of foreign technologies (the so-called absorption potential). Hence, the ratio of total R&D expenditure to GDP has been used as a proxy for the ability of a country to reach the frontier. The data come from the OECD R&D database.

Baumol effect

The “cost-disease” Baumol effect is proxied by the productivity growth in total economy. The productivity data come from the *OECD Economic Outlook*, No. 91, and Johansson et al. (2012). It is calculated as real GDP per worker in 2005 constant PPP USD.

Participation rate

Participation rates are those underlying GDP projections and come from the *OECD Economic Outlook*, No. 91, and Johansson et al. (2012). For OECD countries, the participation rate projections are based on the so-called “cohort approach” (Burniaux et al., 2003). The cohort approach assumes that the observed participation behaviour of individuals belonging to the most recent cohorts, such as the lower exit rates of current old-age workers relative to previous cohorts, or the higher entry rates of current young women relative to previous cohorts, will continue to apply to future cohorts as well. Therefore, future participation rates are determined by the participation behaviour of the most recent cohorts and the evolution of the relative weight of different cohorts, which is driven by demographic developments (see Johansson et al., 2012, for more details).

Some policy reforms are taken into account in the labour force participation projections. First, the long-term trend expansion in education – and the associated increase in average years of schooling – is assumed to continue in all countries. Second, longer life expectancy and health improvements raise the scope for policies that encourage higher labour market participation at older age. And finally, a number of countries have already implemented or plan to implement reforms aiming to extend working lives, including by increasing the legal age to get a full pension. Recently-legislated pension reforms that involve an increase in the normal retirement age by 2020 are assumed to be implemented as planned* (see Johansson et al., 2012, for more details).

* The countries for which an adjustment on current exit rates of older workers are made include Australia, Belgium, Canada, Czech Republic, Germany, Spain, Estonia, France, the United Kingdom, Greece, Hungary, Ireland, Israel, Italy, Japan, New Zealand, Slovak Republic, Slovenia, Turkey and the United States.

Japan's challenging debt dynamics

by

Yvan Guillemette and Jan Stráský*

A small simulation model is used to evaluate the contribution that the three arrows of the government's strategy – bold monetary policy to achieve higher inflation, flexible fiscal policy and growth-boosting structural reforms – could make to reversing the rise in Japan's public debt ratio, currently about 230% of GDP. The findings indicate that with fiscal consolidation amounting to around 7½ percentage points of GDP by 2020, modestly higher growth coming from increased female labour force participation and higher productivity growth, as well as inflation gradually rising to 2% thanks to unconventional monetary policy measures, the debt ratio could be put on a downward trajectory by the end of this decade, although it is likely to remain above 200% of GDP in 2035. Among the many uncertainties surrounding this scenario, the risk of a larger-than-projected increase in interest rates stands prominently and could prevent the turnaround in debt dynamics.

JEL classification codes: E63; H68.

Keywords: Japan; debt; deficit; fiscal; budget; projection; simulation; arrow; consolidation; growth; inflation; reform.

* Both authors are with the OECD Economics Department (yvan.guillemette@oecd.org and jan.strasky@oecd.org). The views expressed in this paper are those of the authors and not necessarily those of the OECD or its member countries. The authors wish to thank Sebastian Barnes, Sveinbjorn Blöndal, Balázs Égert, Jørgen Elmeskov, Richard Herd, Randall Jones, Vincent Koen, Tomasz Koźluk, Jean-Luc Schneider and Piritta Sorsa for comments on earlier drafts; Ane Kathrine Christensen for statistical assistance; and Lyn Urmston for assistance in preparing the final document.

1. Introduction

With very high public debt, a large primary deficit, low growth and a history of persistent deflation, Japan is vulnerable to a loss of market confidence in the sustainability of its public finances. A sovereign debt crisis in Japan would be an important source of instability for the world economy given the size of the Japanese economy and its deep international linkages, including those of financial institutions, which hold the bulk of government liabilities. The challenge for Japan is thus to reduce the structural budget deficit and boost nominal GDP growth to prevent the risk of such a scenario. After setting out the scope of the fiscal challenge, this paper uses a small simulation model to evaluate the contribution that the three-arrow strategy announced by the new government in January 2013 – flexible fiscal policy, structural reforms to boost trend growth and a higher inflation target – could make to lowering the debt ratio over the medium term. The main finding is that the government's strategy, if fully implemented, would put the gross debt ratio on a mild downward track, but it could remain above 200% of GDP in 2035, implying that more radical strategies may be necessary if the debt burden is to be reduced more quickly.

2. The scale of the challenge

The sustainability issues facing Japan can be illustrated in a very simple framework. The evolution of debt as a share of GDP (d) depends on the primary balance as a share of GDP (pb), on the difference between the real interest rate (r) and the growth rate of real GDP (g), and approximately follows:

$$\Delta d_t = -pb_t + (r_t - g_t)d_{t-1}$$

Gross public debt has risen from 70% of GDP in 1992 to more than 200% today, and the May 2014 OECD Economic Outlook short-term projections put it at just over 230% of GDP in 2015. Despite financial assets amounting to some 85% of GDP, net public debt would remain about 145% of GDP, the highest in the OECD. The increase in the debt burden over the past two decades is due to a combination of high primary deficits and high real interest rates relative to real GDP growth.

Japan has run a primary deficit for 20 years and it is projected to be over 7% of GDP in 2014. One fundamental factor behind this structural deficit has been demographics, as population ageing tends to increase social spending and reduce tax revenue. These demographic headwinds on public finances are projected to continue: the old-age dependency ratio is expected to keep rising and peak only in the 2050s. Accordingly, public spending on health and long-term care is expected to increase by between 2.1% and 2.8% of GDP by 2030, of which 0.8 percentage points is due purely to demographics (de la Maisonneuve and Oliveira Martins, 2013). Without the reforms planned by the government which imply a falling replacement rate, upward pressure on public pension expenditure could also be expected.

Trend real economic growth has averaged less than 1% since 1995, also weighed down by population ageing, and is set to average less than 1½ per cent through 2060 in the

long-term baseline scenario produced by the OECD (see Chapter 4 in OECD, 2014). Over the past 15 years, the economy has also generally been operating below capacity, entrenching deflation, which in turn has worsened debt dynamics by increasing the real interest rate. Indeed, despite low nominal interest rates – the policy rate has been close to the zero bound for a decade and a half – deflation has kept real long-term interest rates hovering between 2½ and 3%, noticeably higher than average real GDP growth, and not stimulative enough to jump-start the economy. According to accumulating empirical evidence, this significant wedge between real interest rates and real growth rates may be partly due to the high public debt burden itself (Turner and Spinelli, 2012; Kumar and Woo, 2010; Caner et al., 2010; Elmeskov and Sutherland, 2012; Égert, 2012). With the fiscal and monetary stimulus introduced from early 2013 as part of the government's three-arrow strategy, growth has picked up, the output gap has closed and deflation has ended. But it is unclear how durable these effects will be. As noted above, growth is projected to be weak for decades to come. Slipping back into deflation remains an important risk. And the high public debt problem remains.

Even during the deflationary period, real interest rates remained lower than if they had included the normal fiscal risk premium observed in other highly indebted countries. One reason why the risk premium may be low in Japan is the high proportion of government debt which is financed from domestic sources, about 92%.¹ This has been possible thanks to significant home bias, a high private saving rate, and a current account that has been in surplus since the early 1980s, so that for the past three decades Japan has not had to rely on external sources to finance its government deficits. The current account could move into deficit for structural reasons sometime over the next decade if population ageing led to a decline in the private saving rate, though this is by no means certain and also depends on the evolution of government saving. Japanese investors could also decide to diversify their portfolios by investing more overseas. If either scenario, or both, occurs and the government needs to seek foreign sources of financing, foreign investors could ask for a more “normal” fiscal risk premium.

3. How much can the government's three-arrow strategy lower the debt burden?

3.1. Flexible fiscal policy

Using the May 2014 OECD *Economic Outlook* short-term projections to 2015 as a starting point, a first scenario is constructed using a simulation model that is anchored on long-term trend output projections for the 2016-to-2060 simulation period (Box 1). The scenario includes the fiscal policy aspect of the government's three-arrow strategy, including the fiscal stimulus packages announced in 2013 and 2014, the consumption tax increases of April 2014 and October 2015, and the large amount of fiscal consolidation planned for 2015. Thereafter, it adheres to the government's commitment to keep to the medium-term fiscal objective first announced in the 2010 Fiscal Management Strategy of eliminating the primary budget deficit by 2020.² A substantial amount of fiscal consolidation is necessary, otherwise the debt ratio would be on an exploding path. To isolate the impact of the three arrows one by one, this first scenario does not take the recent re-orientation of monetary policy into account, nor does it include the new growth strategy that was announced in June 2013 and revised in June 2014. A fiscal multiplier of 0.5 is used. On this basis, gross debt reaches close to 245% of GDP in 2020, and thereafter fiscal policy continues to tighten slowly so as to eventually stabilise debt at 230% of GDP, close to its current level (Table 1).

Box 1. Main features of the simulation model

The simulation model is based on Rawdanowicz (2014) and Johansson et al. (2013). Its main features are:

- Initial potential output projections come from the long-term conditional convergence model described in Chapter 4 of OECD (2014) and in Johansson et al. (2013). They are consistent with the supply side underlying the May 2014 *OECD Economic Outlook* short-term projections. However, because there is no investment in the simulation model, potential output is not affected via the impact of interest rates on capital accumulation as in Johansson et al. (2013). On the other hand, potential output is endogenised via a hysteresis effect, so that a 1% output gap for one year permanently affects the level of potential output by 0.1%, for both positive and negative gaps, an effect consistent with estimates in DeLong and Summers (2012) and Guichard and Rusticelli (2010). Through the effect of real interest rates on the demand side (see below), the supply side is thus indirectly, if only slightly, affected.
- The cyclical component of real GDP growth is driven by closure of the output gap with an elasticity of -0.3, so that an output gap roughly halves in two years. In addition, real output growth is affected by fiscal policy via a short-term fiscal multiplier, assumed to be 0.5 in most scenarios. Short-term real output growth is also affected by changes in real long-term interest rates: a 1 percentage point increase in the real interest rate reduces growth by 0.3 percentage points.
- Fiscal policy is implemented through assumptions on the evolution of the underlying primary balance. This measure is cyclically-adjusted via an assumption that a 1% output gap lowers the actual primary balance by 0.3% of GDP, an estimate taken from Girouard and André (2005). Projected increases in social security expenditure due to population ageing are not included in the model, so such costs represent additional fiscal consolidation needs.
- Inflation, defined in term of the GDP deflator, is modelled using an expectations-augmented Phillips curve, with expectations set as a weighted average of past inflation (weight of 0.4), future inflation (assuming perfect foresight, weight of 0.2) and the inflation target, set at 1% in scenario 1 (weight of 0.4). A 1% negative output gap for one year is assumed to lower inflation by 0.4 percentage points.
- The monetary policy stance is driven by interest-rate smoothing toward a standard Taylor rate that cannot fall below an assumed bound of 0.1%. The Taylor rate is set in response to the output gap and to deviations of actual inflation from the target. The natural nominal short-term interest rate is assumed to be equal to a 10-year average of real trend output growth, plus the inflation target, plus a constant of 0.4 percentage points. The monetary authorities are assumed to keep the short-term interest rate at the zero bound despite the effect on inflation of the 2014/15 indirect tax increases. Remittances from the Bank of Japan to the treasury are not modelled.
- The long-term (10-year) interest rate is modelled as a 10-year average of future short-term policy rates (under perfect foresight), a term premium (fixed at 0.7%) and a fiscal-risk premium which depends on the gross debt ratio. This fiscal risk premium is equal to $\frac{1}{2}$ a basis point for each percentage point of gross debt in excess of 75% of GDP, and an additional $\frac{1}{2}$ basis point for each percentage point of gross debt in excess of 125% of GDP. This premium is low, only one-quarter of that typically used in OECD fiscal simulations of other countries.
- The cost of debt servicing depends on the maturity structure of debt, as well as past and projected interest rates. The initial maturity structure of debt at the start of the simulations is calibrated on the maturity distribution of Japanese government bonds (JGBs) outstanding in December 2013. The maturity structure of new debt issues to cover projected gross financing needs is calibrated on the JGB issuance plan for fiscal year 2014 as reported in December 2013. Debt is assumed to be issued at 1, 2, 5, 10, 20 and 30-year maturities. The 20-year (30-year) bond interest rate is assumed to have a term premium of 90 (110) basis points over the 10-year bond rate, corresponding approximately to observed term premiums in April 2014. One-year debt is assumed to be financed at the policy rate, and the interest rates on two-year and five-year debt are assumed to be weighted averages of the short and long-term rates.

Table 1. **Scenario 1: Fiscal stimulus followed by consolidation**

	2014	2015	2016	2017	2018	2019	2020	2025	2030	2035
Potential real GDP growth (%)	0.8	0.8	0.9	0.8	0.8	0.8	0.8	1.0	1.2	1.2
Actual real GDP growth (%)	1.2	1.2	-0.4	0.1	0.3	0.7	0.9	1.1	1.2	1.2
Output gap (%)	0.5	1.0	-0.3	-1.0	-1.4	-1.5	-1.4	-0.7	-0.3	-0.2
Output price inflation (%)	1.3	1.5	0.8	0.3	-0.1	-0.3	-0.3	0.3	0.6	0.8
Short-term interest rate (%)	0.1	0.2	0.1	0.3	0.1	0.1	0.1	0.7	1.5	2.0
Long-term interest rate (%)	0.9	1.7	1.9	2.1	2.3	2.5	2.6	3.5	4.0	4.3
Net lending (% of GDP)	-8.4	-6.7	-5.2	-3.6	-2.3	-1.5	-1.0	-1.1	-1.9	-2.5
Underlying primary balance (% of potential GDP)	-7.1	-5.4	-3.9	-2.4	-1.0	-0.1	0.5	1.9	2.6	3.1
Cumulative fiscal consolidation since 2014 (% of potential GDP)	0.0	1.6	3.1	4.6	6.1	7.0	7.6	9.0	9.6	10.1
Gross interest payments (% of GDP)	2.0	2.3	2.6	2.7	2.9	3.2	3.4	5.0	6.6	7.8
Net interest payments (% of GDP)	1.1	1.3	1.0	0.7	0.7	0.9	1.1	2.8	4.4	5.5
Gross debt (% of GDP)	229.6	232.5	237.1	240.2	242.1	243.0	243.1	238.9	234.3	231.8
Net debt (% of GDP)	142.5	145.4	150.1	153.1	155.1	156.0	156.0	151.8	147.3	144.7

Source: OECD Economic Outlook 95 Database and model simulations.

This requires fiscal consolidation – measured using the underlying primary balance – amounting to 7½ percentage points of GDP from 2014 to 2020, and an additional 2½ percentage points of GDP in consolidation between 2020 and 2035. The effort is largely front-loaded, with some 3 percentage points of GDP in consolidation over 2015 and 2016 combined, due in large part to planned indirect tax increases. This large fiscal retrenchment weakens activity, as reflected in the output gap going from 1% to –1.5% between 2015 and 2019, and keeps the monetary policy rate up against the zero bound until the early 2020s. Notwithstanding the effects of the planned indirect tax increases in 2014 and 2015, deflation also returns in 2018 and persists until the early 2020s in this scenario.

3.2. A new growth strategy to lift potential growth

The third arrow of the government's strategy is to boost growth, so a second scenario adds a higher trend growth projection to the first scenario. In line with the government's initial announcement of the strategy in June 2013, as well as the June 2014 revision, Japan is assumed to implement product-market reforms that raise total factor productivity growth as well as childcare and tax reforms that increase the female labour-force participation rate. Specifically, the government is assumed to put in place policies that would raise the female labour-force participation rate to the OECD median over a 50-year period, so that by 2035 the female participation rate is some 3 percentage points higher than in the first scenario. In addition, a faster pace of product-market reform is assumed to raise total factor productivity growth gradually by 0.2 percentage points above its growth rate in the first scenario, a modest but realistic increase given that Japan is already close to the world technology frontier. On account of increased participation and productivity combined, real potential growth quickly rises by up to 0.3 percentage points above the first scenario over the 2016-to-2035 period (Table 2). The government's announced objective is to raise trend growth to 2% over the next decade, but the simulation is more conservative with a 1.6% maximum growth rate reached at the end of the simulation period. The level of potential output is 7% higher than in the first scenario by 2035. Actual growth responds to higher potential growth via the assumed elasticity of real GDP growth to the output gap (see Box 1). For greater realism, in this scenario the budget balance is made endogenous to the change in labour-force participation compared with the first scenario, as higher

participation would be expected to raise fiscal revenue and lower social spending in line with the size of automatic stabilisers. Higher productivity growth, on the other hand, is assumed not to affect the primary budget balance, even though commensurate increases in both revenue and spending would tend to worsen the absolute fiscal position given the large initial deficit, though not necessarily as a share of GDP.

Table 2. **Scenario 2: Scenario 1 + higher potential growth**

	2014	2015	2016	2017	2018	2019	2020	2025	2030	2035
Potential real GDP growth (%)	0.8	0.8	1.0	1.1	1.1	1.1	1.1	1.4	1.5	1.6
Actual real GDP growth (%)	1.2	1.2	-0.4	0.4	0.6	1.0	1.3	1.5	1.6	1.6
Output gap (%)	0.5	1.0	-0.4	-1.0	-1.5	-1.6	-1.4	-0.7	-0.3	-0.2
Output price inflation (%)	1.3	1.5	0.8	0.2	-0.2	-0.3	-0.3	0.3	0.7	0.8
Short-term interest rate (%)	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.9	1.9	2.4
Long-term interest rate (%)	0.9	1.7	2.1	2.3	2.5	2.7	2.8	3.8	4.3	4.6
Net lending (% of GDP)	-8.4	-6.7	-5.2	-3.6	-2.2	-1.4	-0.9	-0.7	-1.4	-1.4
Underlying primary balance (% of potential GDP)	-7.1	-5.4	-3.9	-2.4	-0.9	0.2	0.8	2.4	3.2	3.9
Cumulative fiscal consolidation since 2014 (% of potential GDP)	0.0	1.6	3.1	4.7	6.2	7.2	7.9	9.5	10.3	11.0
Gross interest payments (% of GDP)	2.0	2.3	2.6	2.7	3.0	3.3	3.6	5.2	6.7	7.6
Net interest payments (% of GDP)	1.1	1.3	1.0	0.7	0.8	1.0	1.2	2.9	4.5	5.3
Gross debt (% of GDP)	229.6	232.5	237.1	239.7	241.2	241.5	241.0	232.9	224.2	216.3
Net debt (% of GDP)	142.5	145.4	150.0	152.7	154.1	154.5	153.9	145.9	137.2	129.3

Source: OECD Economic Outlook 95 Database and model simulations.

Faster GDP growth has several positive mutually reinforcing effects on Japan's debt dynamics. It directly increases the denominator of the debt-to-GDP ratio. Higher participation simultaneously reduces the budget deficit and thus the rate of increase of the numerator. By reducing the fiscal risk premium, declining debt as a fraction of GDP also leads to lower market interest rates, thus lowering refinancing costs, deficits and debt in a virtuous circle. The effects build up over time: the gross debt-to-GDP ratio is only 2 percentage points lower than in the previous scenario by 2020, but some 15 percentage points lower by 2035, mainly through the positive effect of higher participation on the primary balance.

3.3. Bold monetary policy to achieve the 2% inflation target

The first arrow of the government's strategy is to end the period of sustained deflation and target 2% inflation using bold monetary policy measures. Its impact is illustrated in a third scenario which adds to the second the increase in the inflation target from 1% to 2% and implicitly assumes determined monetary policy action to reach the new target. The underlying primary balance is assumed to evolve exactly as in scenario 2. While in reality higher inflation would increase tax receipts, it would also increase nominal government spending, and the extent of *de facto* indexation may well be higher on the spending side. Besides, the primary balance would tend to deteriorate given that spending is higher than revenue, so the assumption of an unchanged primary balance may overestimate the impact of higher inflation on debt reduction somewhat.

In this scenario, inflation is higher than in the first two scenarios because inflation expectations gradually adjust upward and Japan does not return to deflation after the indirect tax increases of 2014/15 (Table 3). Meanwhile, higher inflation rapidly works its way into short

Table 3. **Scenario 3: Scenario 2 + 2% inflation target**

	2014	2015	2016	2017	2018	2019	2020	2025	2030	2035
Potential real GDP growth (%)	0.8	0.8	1.0	1.1	1.1	1.1	1.1	1.4	1.5	1.6
Actual real GDP growth (%)	1.2	1.2	-0.2	0.3	0.6	1.0	1.2	1.5	1.6	1.6
Output gap (%)	0.5	1.0	-0.2	-1.0	-1.4	-1.6	-1.5	-0.7	-0.3	-0.2
Output price inflation (%)	1.3	1.5	1.5	1.1	0.8	0.7	0.7	1.3	1.7	1.8
Short-term interest rate (%)	0.1	0.2	0.2	0.2	0.4	0.4	0.5	1.9	2.9	3.4
Long-term interest rate (%)	0.9	1.7	2.5	2.8	3.1	3.3	3.6	4.7	5.2	5.5
Net lending (% of GDP)	-8.4	-6.7	-5.2	-3.6	-2.4	-1.7	-1.3	-1.6	-2.2	-2.0
Underlying primary balance (% of potential GDP)	-7.1	-5.4	-3.9	-2.4	-0.9	0.2	0.8	2.4	3.2	3.9
Cumulative fiscal consolidation since 2014 (% of potential GDP)	0.0	1.6	3.1	4.7	6.2	7.2	7.9	9.5	10.3	11.0
Gross interest payments (% of GDP)	2.0	2.3	2.6	2.8	3.2	3.6	4.0	6.1	7.5	8.2
Net interest payments (% of GDP)	1.1	1.3	1.0	0.8	1.0	1.3	1.6	3.8	5.3	5.8
Gross debt (% of GDP)	229.6	232.5	235.9	237.3	237.6	236.9	235.3	224.4	214.5	205.1
Net debt (% of GDP)	142.5	145.4	148.8	150.3	150.6	149.8	148.3	137.3	127.4	118.1

Source: OECD Economic Outlook 95 Database and model simulations.

and long-term nominal interest rates, which also rise relative to previous scenarios. Although higher market interest rates lead to higher financing costs for the government, the implicit average interest rate paid on all outstanding debt rises more slowly than market rates, and more slowly than inflation, reflecting the existing maturity structure of debt. This effect allows the implicit average real interest rate paid on government liabilities to go down, slowly eroding the debt ratio, which is some 10 percentage points of GDP lower in 2035 than in the second scenario. That the effect is modest and front-loaded is due to the relatively short maturity structure of Japanese government debt: close to 20% of currently outstanding market debt matures within one year and 30% is to be refinanced within two years. By way of comparison, if Japan's debt maturity structure and new issuance plan were the same as the United Kingdom's, where less than 20% of debt matures within the next three years, the debt ratio would decline by an additional 15-to-20 percentage points by 2035. Also, by allowing higher inflation to fully pass through into higher government financing costs, the scenario does not account for two channels through which aggressive monetary policy actions could directly lower such costs. First, quantitative easing via the Japanese Government Bond (JGB) market may imply that interest rates rise less than inflation expectations due to portfolio rebalancing effects. Second, quantitative easing substitutes base money and low-earning reserves for higher-earning JGBs in the private sector's portfolio, which is likely to reduce overall interest costs for the government sector (after taking into account remittances from the Bank of Japan, which are not modelled). Hence, public debt could well fall more than implied by this scenario, the assumption of no impact from higher inflation on the primary balance notwithstanding.

The potential additional debt-reduction effects of the unconventional monetary policy actions necessary to substantially raise inflation are illustrated in a final scenario in which higher inflation does not pass through into higher nominal interest rates, though long-term interest rates are assumed to remain sensitive to the evolution of the debt ratio via the fiscal risk premium (see Box 1). In this scenario, not only can the government finance its deficits more cheaply in real terms as inflation increases without an increase in nominal interest rates, but in addition, the resulting fall in real interest rates boosts activity. As a result, the debt burden falls much more rapidly than in the previous scenario where nominal interest rates adjusted to higher inflation. The extreme assumption underlying this scenario needs to be

emphasised, however. To maintain nominal interest rates unchanged compared with Scenario 3 despite inflation having reached the 2% target, the Bank of Japan would have to engage in continuous quantitative easing. Continued aggressive monetary easing would at some point almost certainly generate destabilising effects on inflation and financial markets. Hence, Scenario 4 could at most be followed for a limited time, after which policy would have to revert to that underlying Scenario 3. During this limited time, it might nonetheless have a durable impact on debt levels compared with Scenario 3 (Table 4). The gross debt-to-GDP ratio in this scenario is more than 15 percentage points lower in 2030 than in Scenario 3.

Table 4. Scenario 4: Scenario 3 with no pass-through of inflation into higher interest rates

	2014	2015	2016	2017	2018	2019	2020	2025	2030	2035
Potential real GDP growth (%)	0.8	0.8	1.0	1.1	1.1	1.1	1.1	<i>1.4</i>	<i>1.5</i>	-
Actual real GDP growth (%)	1.2	1.2	0.0	0.3	0.6	1.0	1.2	<i>1.5</i>	<i>1.6</i>	-
Output gap (%)	0.5	1.0	0.0	-0.8	-1.3	-1.4	-1.4	<i>-0.6</i>	<i>-0.3</i>	-
Output price inflation (%)	1.3	1.5	1.5	1.2	0.9	0.8	0.8	<i>1.3</i>	<i>1.7</i>	-
Short-term interest rate (%)	0.1	0.2	0.2	0.2	0.2	0.2	0.2	<i>0.9</i>	<i>1.9</i>	-
Long-term interest rate (%)	0.9	1.7	2.1	2.2	2.4	2.6	2.8	<i>3.6</i>	<i>4.1</i>	-
Net lending (% of GDP)	-8.4	-6.7	-5.1	-3.4	-2.1	-1.2	-0.6	<i>-0.1</i>	<i>0.0</i>	-
Underlying primary balance (% of potential GDP)	-7.1	-5.4	-3.9	-2.4	-0.9	0.2	0.8	<i>2.4</i>	<i>3.2</i>	-
Cumulative fiscal consolidation since 2014 (% of potential GDP)	0.0	1.6	3.1	4.7	6.2	7.2	7.9	<i>9.5</i>	<i>10.3</i>	-
Gross interest payments (% of GDP)	2.0	2.3	2.6	2.7	2.9	3.1	3.4	<i>4.6</i>	<i>5.4</i>	-
Net interest payments (% of GDP)	1.1	1.3	0.9	0.7	0.7	0.8	1.0	<i>2.3</i>	<i>3.2</i>	-
Gross debt (% of GDP)	229.6	232.5	235.4	236.6	236.4	235.0	232.7	<i>216.1</i>	<i>198.3</i>	-
Net debt (% of GDP)	142.5	145.4	148.4	149.5	149.3	148.0	145.7	<i>129.1</i>	<i>111.2</i>	-

Note: Results appear in italics after 2020 and not at all for 2035 to reflect the fact that policies underlying the scenario could in practice not be maintained for very long (see text).

Source: OECD Economic Outlook 95 Database and model simulations.

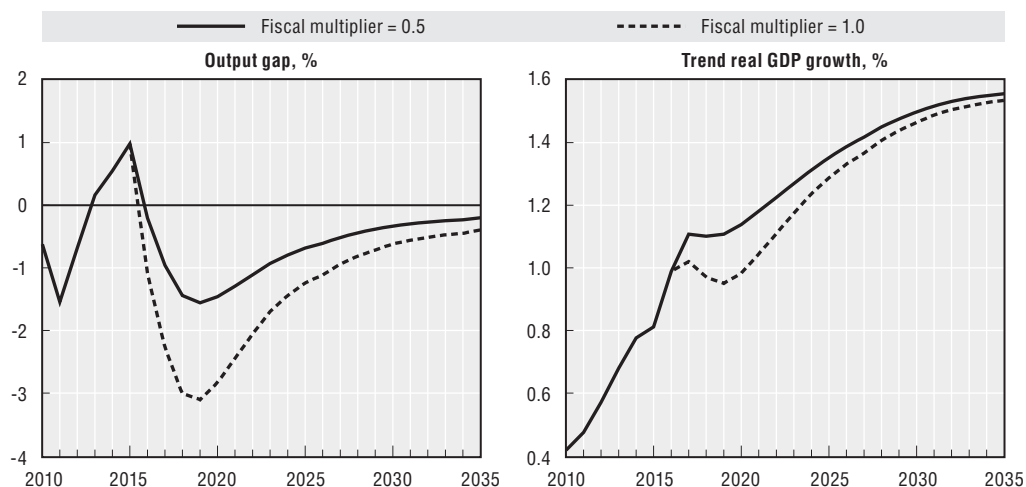
4. What if the fiscal multiplier were higher?

All of the above scenarios have used a fiscal multiplier of 0.5, meaning that budget consolidation equal to 1% of GDP in a given year reduces real GDP growth in that year by 0.5 percentage point and, through the assumed hysteresis effect, reduces potential real GDP growth in that year by 0.05 percentage point (see Box 1). This should be considered a conservative fiscal multiplier, especially for a depressed economy (Auerbach and Gorodnichenko, 2012 and 2013; Owyang, Ramey and Zubairy, 2013; Fazzari, Morley and Panovska, 2012; Mitnik and Semmler, 2012). Recent estimates of Japan-specific multipliers for fiscal contractions when the output gap is negative place the multiplier at 2 for spending cuts and at about 0.6 for revenue increases (Baum, Poplawski-Ribeiro and Weber, 2012). When the output gap is positive, a spending cut is still estimated to have a multiplier of 1.7. An overall multiplier of 0.5 might accordingly be an appropriate assumption for revenue-based fiscal consolidation, but much too low for spending-based consolidation. Given large consolidation needs, a workable strategy is likely to rely on both types of measures, but should rely mostly on increasing revenue given the large revenue space that Japan enjoys when compared with other OECD countries, with total tax revenue as a share of GDP 7½ percentage points lower than the OECD average (OECD, 2013). Thus, a reasonable, but still conservative, multiplier estimate for a consolidation programme that relies mostly, but not exclusively, on revenue measures might be 1.

With a fiscal multiplier of 1, the negative effects of fiscal consolidation on economic activity and prices that occur in the scenarios presented above become more significant.

For instance, in Scenario 3, going from a 0.5 to a 1.0 fiscal multiplier lowers nominal GDP growth by an average of 1.5 percentage points between 2016 and 2020 when fiscal consolidation is most rapid, through both real activity and price effects. At its widest in 2019, the negative output gap would be twice as large as in Scenario 3, pointing also to the potential political difficulties of rapid fiscal consolidation in a high-multiplier context (Figure 1). Moreover, through the assumed hysteresis effect, the larger negative output gap

Figure 1. **Impact of increasing the fiscal multiplier from 0.5 to 1.0 in Scenario 3**



Source: OECD Economic Outlook 95 Database and model simulations.

lowers trend real GDP growth slightly, marking down the level of both actual and potential output permanently. The size of the effect is modest: the level of real potential output is 1.3% lower in 2035 with a fiscal multiplier of 1 than with the 0.5 multiplier. If hysteresis mechanisms were to be stronger than assumed, however, the negative impact of fiscal consolidation on the long-term productive capacity of the economy would be more material. Finally, with a fiscal multiplier of 1, the fiscal consolidation path in Scenario 3 does not succeed in substantially reducing the gross debt ratio by 2035: the ratio then is some 10 percentage points higher than with a fiscal multiplier of 0.5 (Table 5).

Table 5. **Scenario 5: Scenario 3 with a fiscal multiplier of 1.0**

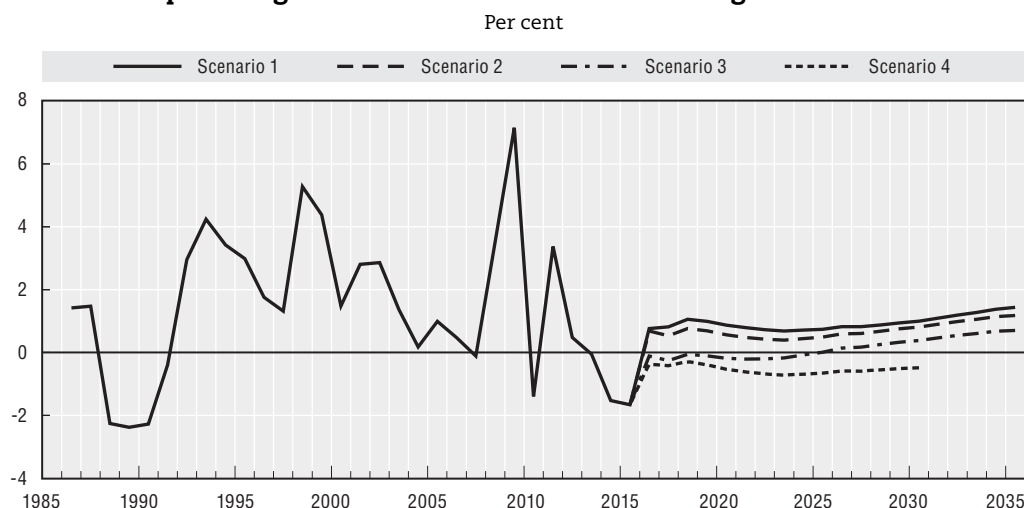
	2014	2015	2016	2017	2018	2019	2020	2025	2030	2035
Potential real GDP growth (%)	0.8	0.8	1.0	1.0	1.0	1.0	1.0	1.3	1.5	1.5
Actual real GDP growth (%)	1.2	1.2	-1.1	-0.2	0.2	0.8	1.3	1.5	1.6	1.6
Output gap (%)	0.5	1.0	-1.1	-2.3	-3.0	-3.1	-2.8	-1.2	-0.6	-0.4
Output price inflation (%)	1.3	1.5	0.9	0.1	-0.5	-0.7	-0.6	0.7	1.4	1.6
Short-term interest rate (%)	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.6	2.1	2.9
Long-term interest rate (%)	0.9	1.7	2.0	2.2	2.4	2.6	2.8	4.0	4.8	5.3
Net lending (% of GDP)	-8.4	-6.7	-5.5	-4.0	-2.7	-1.9	-1.3	-1.0	-1.8	-2.1
Underlying primary balance (% of potential GDP)	-7.1	-5.4	-3.9	-2.4	-0.9	0.2	0.8	2.4	3.2	3.9
Cumulative fiscal consolidation since 2014 (% of potential GDP)	0.0	1.6	3.1	4.7	6.2	7.2	7.9	9.5	10.3	11.0
Gross interest payments (% of GDP)	2.0	2.3	2.6	2.8	3.0	3.3	3.6	5.3	7.0	8.2
Net interest payments (% of GDP)	1.1	1.3	1.0	0.7	0.8	1.0	1.3	3.0	4.8	5.9
Gross debt (% of GDP)	229.6	232.5	238.3	242.4	245.6	247.2	247.6	239.1	227.2	216.9
Net debt (% of GDP)	142.5	145.4	151.2	155.3	158.5	160.2	160.5	152.0	140.1	129.9

Source: OECD Economic Outlook 95 Database and model simulations.

5. Uncertainties

The uncertainties around the scenarios presented above are large. As pertains to the evolution of the debt ratio, uncertainties can be thought about using the debt dynamics equation presented in Section 2. While the evolution of the primary budget deficit is uncertain, it is to a large extent under the control of government via its tax and spending decisions. The other crucial parameter in the evolution of the debt burden, the differential between the real interest rate paid on government debt (r) and the real growth rate of GDP (g), is less directly controlled by the authorities and its evolution is more uncertain. Any positive differential pushes up the debt ratio unless it is offset by a primary budget surplus of corresponding size. In the scenarios presented above, the $r - g$ differential is generally projected to increase over the projection period as interest rates normalise and rise somewhat more than the real GDP growth rate (Figure 2). But it remains relatively low and always below two percentage points. This is roughly in line with the historical experience: from 1986 to 2013, the $r - g$ differential has averaged 1.6%. It has, however, often been higher than that for extended periods of time, and sometimes much higher. That this differential could increase more than projected, probably mainly through a faster or sharper increase in interest rates than the gradual normalisation modelled here, is the main risk around the scenarios. The high level of government debt itself could cause this risk to materialise, particularly if future deficits must be partially financed externally. For non-euro area countries with high debt, one recent study estimates that each percentage point increase in government debt as a share of GDP raises the $r - g$ differential by 2½ basis points if it is financed entirely domestically, and by 3½ to 5 basis points if it is financed externally (Turner and Spinelli, 2013).

Figure 2. **Historical and projected differential between the real interest rate paid on government debt and the real GDP growth rate**



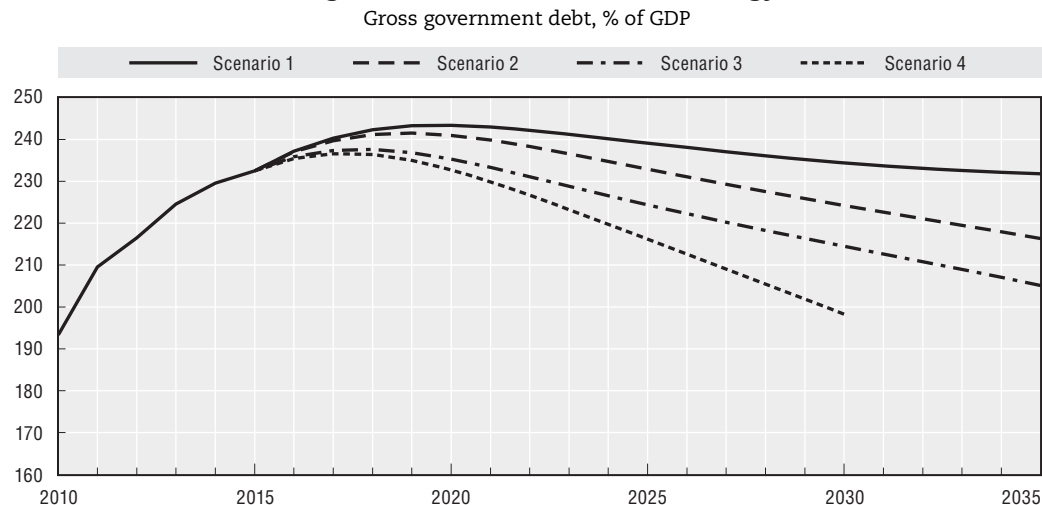
Note: The real interest rate paid on government debt is computed as the implicit nominal interest rate paid on gross government liabilities minus output price inflation.

Source: OECD Economic Outlook 95 Database and model simulations.

6. Conclusion and way forward

The analysis has shown that, together, the three arrows of the government's strategy could be successful in arresting the rise of the public debt ratio around 2020 and putting it on a downward track (Figure 3). Nevertheless, the debt burden would likely still be above 200%

Figure 3. **Evolution of Japan's public debt over the medium term under the government's three-arrow strategy**



Source: OECD Economic Outlook 95 Database and model simulations.

of GDP in 2035, particularly if the fiscal multiplier is 1 or more. High refinancing and debt turnover needs would keep the fiscal situation tense and vulnerable to a crisis of confidence. Maintaining market trust is thus paramount in the months and years to come, and this means fleshing out the monetary, fiscal and structural policy measures to be taken, their timing, and avoiding any slippage in their implementation to maintain credibility.

Notes

1. At the end of 2013, about 40% of government bonds were held by Japanese commercial banks, about 20% by the Bank of Japan, about 30% by Japanese insurance and pension funds (including the national pension fund) and about 3% by Japanese households.
2. The government's commitment to eliminate the primary deficit refers only to the central and local government budget. The social security budget, which has been in deficit at about 1% of GDP from 2008 to 2012, is not included in the target. In the simulation model used here, the budget concept used is that of the general government, which includes social security. The amount of fiscal consolidation necessary to eliminate the general government deficit is correspondingly higher than to strictly meet the government's commitment.

References

- Auerbach, A.J. and Y. Gorodnichenko (2012), "Measuring the Output Responses to Fiscal Policy", *American Economic Journal: Economic Policy*, Vol. 4, No. 2, pp. 1-27.
- Auerbach, A.J. and Y. Gorodnichenko (2013), "Fiscal Multipliers in Recession and Expansion" in A. Alesina and F. Giavazzi (eds), *Fiscal Policy After the Financial Crisis*, University of Chicago Press.
- Baum, A., M. Poplawski-Ribeiro and A. Weber (2012), "Fiscal Multipliers and the State of the Economy", *IMF Working Paper*, No. 12/286.
- Caner, M., T. Grennes and F. Koehler-Geib (2010), "Finding the Tipping Point – When Sovereign Debt Turns Bad", *World Bank Policy Research Working Paper*, No. 5391.
- de la Maisonneuve, C. and J. Oliveira Martins (2013), "Public Spending on Health and Long-Term Care: A New Set of Projections", *OECD Economic Policy Papers*, No. 6, OECD Publishing.
- DeLong, J.B. and L.H. Summers (2012), "Fiscal Policy in a Depressed Economy", *Brookings Papers on Economic Activity*, Spring, pp. 233-274.

- Égert, B. (2012), "Public Debt, Economic Growth and Nonlinear Effects: Myth or Reality?", *OECD Economics Department Working Papers*, No. 993, OECD Publishing.
- Elmeskov, J. and D. Sutherland (2012), "Post-Crisis Debt Overhang: Growth and Implications across Countries", *OECD Economics Department mimeo*, www.oecd.org/dataoecd/7/2/49541000.pdf.
- Fazzari, S.M., J. Morley and I. Panovska (2012), "State-Dependent Effects of Fiscal Policy", *UNSW Australian School of Business Research Paper*, No. 2012-27.
- Girouard, N. and C. André (2005), "Measuring Cyclically-adjusted Budget Balances for OECD Countries", *OECD Economics Department Working Papers*, No. 434, OECD Publishing.
- Guichard, S. and E. Rusticelli (2010), "Assessing the Impact of the Financial Crisis on Structural Unemployment in OECD Countries", *OECD Economics Department Working Papers*, No. 767, OECD Publishing.
- Johansson, Å., Y. Guillemette, F. Murtin, D. Turner, G. Nicoletti, C. de la Maisonneuve, P. Bagnoli, G. Bousquet and F. Spinelli (2013), "Long-term Growth Scenarios", *OECD Economics Department Working Papers*, No. 1000, OECD Publishing.
- Kumar, M.S. and J. Woo (2010), "Public Debt and Growth", *IMF Working Paper*, No. 10/174.
- Mauro, P., R. Romeu and A. Binder (2013), "A Modern History of Fiscal Prudence and Profligacy", *IMF Working Paper*, No. 13/5.
- Mitnik, S. and W. Semmler (2012), "Regime Dependence of the Fiscal Multiplier", *Center for Quantitative Risk Analysis Working Paper*, No. 5.
- OECD (2013), *Revenue Statistics 2013*, OECD Publishing.
- OECD (2014), *OECD Economic Outlook*, Vol. 2014/1, OECD Publishing.
- Owyang, M.T., V.A. Ramey and S. Zubairy (2013), "Are Government Spending Multipliers Greater During Periods of Slack? Evidence from Twentieth-Century Historical Data", *American Economic Review*, Vol. 103, No. 3, pp. 129-134.
- Rawdanowicz, L. (2014), "Choosing the Pace of Fiscal Consolidation", *OECD Journal: Economic Studies*, Vol. 2013/1, pp. 91-119.
- Turner, D. and F. Spinelli (2012), "Interest-rate-Growth Differentials and Government Debt Dynamics", *OECD Journal: Economic Studies*, Vol. 2012/1, OECD Publishing.
- Turner, D. and F. Spinelli (2013), "The Effect of Government Debt, External Debt and their Interaction on OECD Interest Rates", *OECD Economics Department Working Papers*, No. 1103, OECD Publishing.

Rescuing the Phillips curve: Making use of long-term unemployment in the measurement of the NAIRU

by
Elena Rusticelli*

Despite the increased importance of cyclically-adjusted measures of labour market slack for policymaking, estimates of the NAIRU have become increasingly fragile. Particularly for euro area countries, NAIRU estimates represent a crucial input to compute cyclically-adjusted budget balances adopted to formulate medium-term fiscal objectives under the EU fiscal surveillance framework. However, the apparent reduced sensitivity of inflation to labour market dynamics and unemployment gaps seriously undermines the use of Phillips curve equations in estimating the NAIRU. Estimates of the NAIRU are particularly problematic when changes in unemployment are both very large and rapid as in the aftermath of the global crisis. This paper proposes a refinement to the standard OECD approach of using a Kalman filter to estimate the NAIRU in the context of the Phillips curve. The proposed refinement strengthens the relationship between inflation and labour market developments by considering the risk of hysteresis effects associated with changes in long-term unemployment. Testing the revised methodology on a broad selection of OECD countries gives mixed results. For a group of countries in the euro area periphery (Greece, Ireland, Italy, Portugal and Spain) there is an increase in the magnitude and statistical significance of the unemployment gap, with the NAIRU revised upward by on average 1¾ percentage points. However, the revised methodology provides less improvement to the standard OECD methodology for a second set of countries considered, namely the G7 excluding Italy. The United States is an interesting intermediate case as the statistical evidence for the proposed methodology is marginal, but the policy implications of the revised point estimate of the NAIRU are major.

JEL classification: C32, E24, E31, E32, J64.

Keywords: Long-term unemployment, flattening Phillips curve, NAIRU, euro area periphery, Kalman filter.

* Elena Rusticelli works in the Economics Department (elena.rusticelli@oecd.org). The opinions expressed and arguments employed in this paper are the responsibility of the author and do not necessarily reflect the official views of the Organisation or of the governments of its member countries. The author would like to thank Dave Turner for fundamental guidance and discussions, Fabrice Orlandi and Alex Hijzen for helpful comments and suggestions on a previous version and Ines Gomez Palacio for help with documentation preparation.

1. Introduction

Unemployment in many OECD countries has soared and estimates of the NAIRU (Non-Accelerating Inflation Rate of Unemployment) appear increasingly fragile, while at the same time such estimates have become more important to policymaking: most notably, because of their formal incorporation in fiscal frameworks, but also because they provide a basis for assessing future deflation risks. The fragility of NAIRU estimates stems from the reduced ability of such estimates (specified as the gap with actual unemployment) in explaining inflation dynamics in Phillips curves equations. NAIRU estimates have critical implications for fiscal adjustment strategies which, in many countries, are set in terms of the cyclically-adjusted budget balances, with structural unemployment being an input to the cyclical adjustment. In case of the euro area countries, the cyclically-adjusted balance (CAB) has become the key indicator in the EU fiscal surveillance framework as it is adopted to assess country-specific medium-term fiscal objectives (MTO) as well as the corrective fiscal adjustment required for member states experiencing excessive deficit situations. Moreover, since the *European Fiscal Compact* entered into force in March 2012, EU member states are obliged to transfer their MTO into national binding laws and to pay financial sanctions in case of a repeated failure to comply with the set fiscal objectives.

At the same time, inflation seems to have become less sensitive to labour market dynamics and unemployment gaps, with such variables losing explanatory power in Phillips curves equations. Different factors have been recognised as playing a crucial role in this “flattening” of the Phillips curve, although their impact varies across different OECD countries. Rising unemployment levels and duration have contributed to exacerbate nominal wage stickiness, which associated with rigid labour market institutions, tend to curb inflation adjustments to macroeconomic dynamics. At the same time, the recent increase in labour mobility and migration, especially in Europe, contributes to reducing inflation due to existing differences in wage settings as well as consumption propensities between natives and immigrants. In addition, inflation expectations which are well anchored to central banks’ inflation targets have weakened the relationship between inflation and labour market imbalances. However, previous attempts to include central bank inflation targets in the Phillips curve estimation did not produce satisfactory results (Guichard and Rusticelli, 2011). Finally, greater economic integration may have shifted the drivers of the inflation process away from domestic pressure towards more global drivers.

The purpose of this paper is to strengthen the relationship between inflation and the unemployment gap by considering the impact that hysteresis effects have on the NAIRU. The estimation here focuses on a selection of OECD countries: those of the European periphery, where risks of recent hysteresis dynamics are the highest, and the G7 which form a control group to verify whether the proposed methodology can be applied more generally. More specifically, the inclusion of long-term unemployment to model the NAIRU allows an upward shift in the structural component of actual unemployment associated with a tighter relationship between inflation and the unemployment gap. More specifically,

this modification helps increasing the size and the statistical significance of the unemployment gap coefficient in the Phillips curve estimated for the European peripheral countries, while it leaves the results for the G7 economies mostly little changed.

The next section briefly reviews the standard OECD method of estimating the NAIRU by means of a Kalman filter within a Phillips curve framework. Section 3 provides evidence of the reduced sensitivity of inflation to unemployment. The following section describes recent developments in aggregate and long-term unemployment, discussing the role of the latter in hysteresis dynamics and as a predictor of the NAIRU. Section 5 presents the proposed innovation to the existing standard estimation methodology which considers the inclusion of the change in long-term unemployment in modelling the NAIRU. The final section analyses the results and conclusions are drawn.

2. The current OECD approach to estimating the NAIRU

The standard state-space framework introduced in Gianella et al. (2008) and adopted to regularly update the NAIRU estimates for all OECD member states (most recently in Guichard and Rusticelli, 2011) can be summarised as:

$$\Delta\pi_t = \beta(u_t - u_t^*) + \sum_{j=1}^A \alpha_j \Delta\pi_{t-j} + \sum_{j=1}^L \lambda_j \Delta\pi_{t-j}^{imp} + \sum_{j=1}^G \gamma_j \Delta\pi_{t-j}^{oil} + \varepsilon_t \quad [1]$$

$$u_t = u_t^* + u_t^{gap} \quad [2]$$

$$u_t^{gap} = \phi_1 u_{t-1}^{gap} + \phi_2 u_{t-2}^{gap} + \eta_t \quad [3]$$

$$u_t^* = u_{t-1}^* + \nu_t \quad [4a]$$

The first equation represents the measurement equation of the state space model, i.e. the Phillips curve which relates changes in core consumer price inflation ($\Delta\pi_t$) to the unemployment gap (u_t^{gap}), lags of the change in inflation and control variables.¹ The control variables consist of two types of short-term supply shocks: changes in real import price inflation weighted by import penetration ($\Delta\pi^{imp}$) and changes in real oil price inflation ($\Delta\pi^{oil}$) weighted by oil intensity of production. The number of lags is determined on the basis of their statistical significance and to yield a parsimonious model. The second equation is an identity and specifies actual unemployment u_t as the sum of a structural component, i.e. the NAIRU u_t^* , and a cyclical one, i.e. the unemployment gap u_t^{gap} . The third and fourth equations are the transition equations of the state space describing the law of motion of the two unobserved components of unemployment and are consistent with the idea that both variables can be inferred on the basis of the inflationary pressure in the economy. For this purpose, equation [3] specifies the unemployment gap as a stationary second-order autoregressive process (Jaeger and Parkinson, 1994; Laubach, 2001).² The three error terms of the model $\varepsilon_t, \eta_t, \nu_t$ are all assumed to be i.i.d. normally distributed with mean zero and uncorrelated variances $\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\nu^2$, respectively.³

3. Evidence of reduced sensitivity of inflation to unemployment gaps

The reduced sensitivity of the inflation process to the economic activity has become a stylised fact which has generated a plethora of studies trying to assess whether the Phillips curve still remains a reliable tool to express the link between the evolution of prices and wages and labour market dynamics (Smets and Wouters, 2007; Borio and Filardo, 2007; Kuttner and Robinson, 2010; Williams, 2006; Benkovskis et al., 2011). A small or

insignificant slope coefficient β on the unemployment gap term in the Phillips curve in equation [1] implies high unemployment in excess of the NAIRU has little or no disinflationary effect. However, the source of this “flattening” of the Phillips curve is still a widely debated issue.

Many of these studies focus on the change that the reduced volatility implies for the properties of the inflation process, as well as the parameters of the Phillips curve (Roberts, 2006; Sargent, Williams and Zha, 2006; Smets and Wouters, 2007; Kuttner and Robinson, 2010). These latter conclude that changes in monetary policy affect the volatility of inflation and are the main reason for large declines in the Phillips curve slope when all other parameters underlying the inflation behaviour are held fixed. Changes in monetary policy, such as those associated with the introduction of inflation targeting or with euro area monetary integration, create both a mean and a persistence break in the inflation series which can undermine the stability of its relationship with labour market indicators (Bataa et al., 2014). Relatedly, the decline in inflation persistence observed in the last decade with respect to the two preceding ones is also the result of better-anchored inflation expectations around a clear long-run inflation objective which contributes to reduce the uncertainty inherent in estimated econometric models. As a consequence, both lower inflation persistence and a flatter slope in the Phillips curve imply a larger sacrifice ratio, i.e. a larger unemployment or negative output gap that the economy has to bear in order to achieve a given disinflation.

Alternatively, rather than being the result of effective monetary policies, some authors argue that the flattening of the Phillips curve is due to structural changes in the labour market that, when specifically accounted for, would leave the link between inflation and unemployment unchanged. Razin and Binyamini (2007) and Borio and Filardo (2007) explore the effect that an increased openness in labour markets has on domestic marginal costs and inflationary mechanisms and conclude that global factors exert a higher explanatory power on domestic inflation than those from domestic measures of economic slack, such as output or unemployment gaps. Blanchard and Galí (2010) and Elsy et al. (2010) show that the degree of labour market tightness, captured by increasing hiring costs and real wage rigidities, increases marginal costs, with “sclerotic” labour markets, such as those in Europe, displaying higher inflation persistency to productivity shocks than more flexible labour markets.

The reduced sensitivity of inflation to unemployment gaps is confirmed by re-estimating the OECD Phillips curves over sub-samples, comparing the most recent period 2000Q1-2012Q4 with a sub sample of similar length 1987Q1-1999Q4 (Table 1). For all countries examined, the coefficient on the unemployment gap estimated over the most recent period is smaller and less statistically significant when compared with the results obtained over the entire sample. In particular, when the estimates are carried out using two datasets of similar size, the relationship between unemployment and inflation seems to have especially weakened in the cases of Canada, Germany, Greece, and the United Kingdom.

The loss of responsiveness of inflation dynamics to cyclical unemployment is not a recent phenomenon generated by the global financial crisis, nor is it completely unrelated to future expectations on prices and wages developments. Since the late 1990s, the success of central banks in holding inflation stable and low, while anchoring inflation expectations to explicit target levels, contributes to explain the quite modest declines in inflation in the

Table 1. **Estimated coefficients of the unemployment gap in the Phillips curve**

	Whole data sample	1987Q1-1999Q4	2000Q1-2012Q4
Canada	-0.103***	-0.129***	0.037
France	-0.112***	-0.065*	-0.065
Germany	-0.062***	-0.106*	-0.007
Greece	-0.046*	-0.167***	-0.006
Ireland	-0.035**	-0.051*	-0.005
Italy	-0.048*	-0.034	-0.034
Japan	-0.232***	-0.005	-0.004
Portugal	-0.092**	-0.052	-0.017
Spain	-0.028**	-0.015	-0.001
United Kingdom	-0.149***	-0.131**	-0.039
United States	-0.036***	-0.044**	-0.006

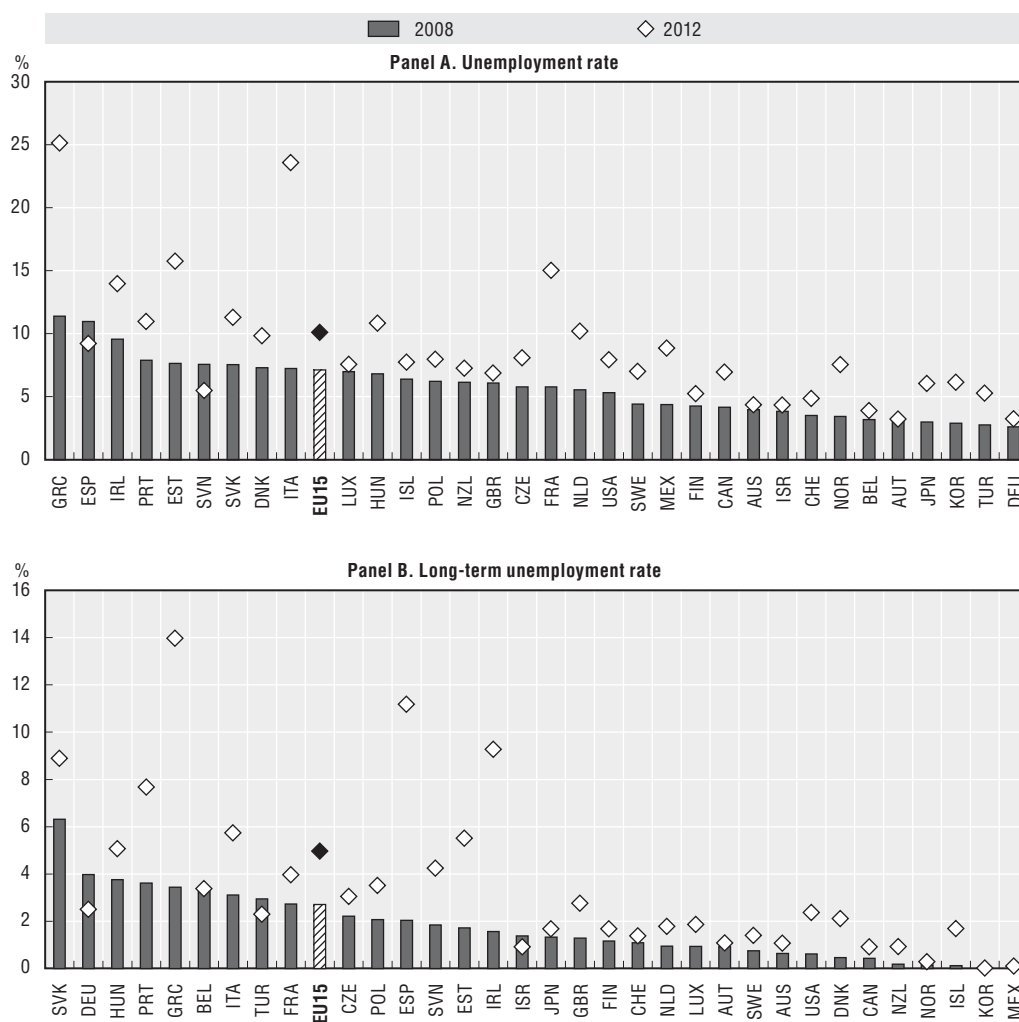
Notes: The Phillips curve specification and the exact period to which the estimation over “the whole data sample” refers are reported in Table 3. *, **, *** indicate statistically significant at 10%, 5% and 1%, respectively.

aftermath of the global crisis, despite the existence of a great economic slack (IMF, 2013). As empirical evidence, the OECD Phillips curves estimated over the sub-sample covering the pre-crisis period, i.e. 2000Q1-2007Q4, return a not statistically significant coefficient of the unemployment gap for all the countries considered in the present exercise. These results suggest that the flattening of the Phillips curve began much earlier than the global crisis, and most likely around the time when inflation-targeting regimes were adopted across advanced economies (Rusticelli et al., 2014).

The observed flattening of the Phillips curve at low levels of inflation is also consistent with the existence of downward nominal wage rigidities (Yellen, 2012). The relevance of nominal downward wage rigidities has risen considerably since the beginning of the crisis, accompanied by an increased incidence of nominal wage freezes across the OECD as a whole (OECD *Employment Outlook*, 2014). As proof, recent regression-based estimates of wage-Phillips curves, which control for price developments, indicate that the slope of the Phillips curve has flattened during the course of the global financial crisis as nominal wage growth has slowed (OECD *Employment Outlook*, 2014). Moreover, the adjustment of nominal wages to the persistent labour market slack tends to be higher in good times rather than in bad times, introducing therefore an asymmetric responsiveness of nominal wage growth over the cycle.⁴

4. Recent developments in long-term unemployment and risks of hysteresis

The global financial crisis has had alarming repercussions on OECD unemployment rates, with particularly severe increases being experienced by the European peripheral countries (Figure 1, panel A). In the five-year period following the crisis, unemployment has increased on average by 3¾ percentage points in the euro area as a whole and by 12 percentage points on average in the peripheral countries of Greece, Ireland, Spain and Portugal. In these countries, unemployment continues to rise with average unemployment rates equal to about 20% in 2013. Beyond Europe, the effects of the economic recession on labour markets have been less extreme in the United States and Japan, where the unemployment rate has increased by about 2½ and ½ percentage points, respectively, since the beginning of the financial crisis.

Figure 1. **Developments in unemployment rates**

Note: The long-term unemployment rate refers to the unemployment rate for those unemployed for more than 12 months.

Source: Eurostat, OECD Labour Force Statistics Database.

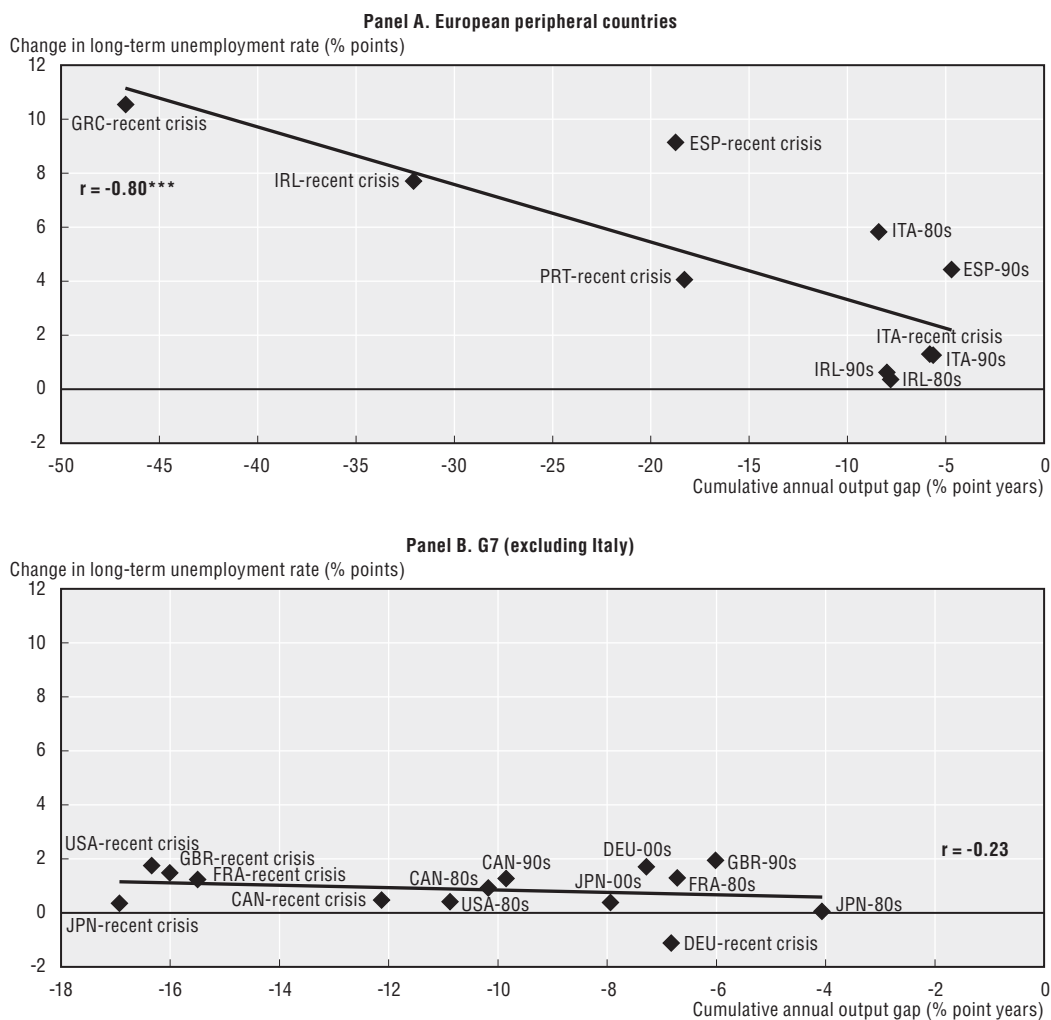
From the mid-1980s until the beginning of the sovereign debt turmoil in 2010, euro area long-term unemployment remained fairly stable at around 4%. Indeed, the boom in the real estate sector in both Spain and Ireland contributed to reduce long-term unemployment in the European peripheral countries, which reached the average lowest rate of 2.7% at the onset of the global financial crisis in 2007. But, as unemployment started to rise, as a consequence of the economic recession generated from both the global financial and the sovereign debt crisis, the share of long-term unemployment (the percentage of the active population having been unemployed for 12 months or more) first declined due to the increase in inflows with short duration, then it increased given the decline in both the exit probabilities and the outflows from unemployment. The latter are more extreme in Europe than the United States and Japan (OECD *Employment Outlook*, 2011) and in part explain the quite heterogeneous behaviour of unemployment indicators across the OECD. It follows that the

average duration of unemployment is much longer in the euro area than elsewhere, and as consequence long-term unemployment in the euro area is at present more than double that of the United States (Figure 1, panel B), at 5% and 2.4%, respectively. In the European peripheral countries (notably Greece, Ireland, Spain and Portugal), the rate of long-term unemployment in 2012 is on average twice as large as that of the entire euro area, i.e. 10%, with an average rise by about 8 percentage points since the beginning of the crisis.

The rise in aggregate unemployment caused by economic downturns is more likely to degenerate into an increase in long-term unemployment depending on the setting of labour and product-market structural policies (Llaudes 2005; Guichard and Rusticelli, 2010). More precisely, in those countries where product-market regulations (PMR) are tighter both the level of long-term unemployment and its response to a shock on the actual unemployment rate are higher (by about 0.1 to 0.25 percentage points per unit of PMR). Likewise, where long-term unemployment benefits are generous, the response of long-term unemployment to a shock on aggregate unemployment will tend to increase by 0.1 percentage points for each percentage point shock. Conversely, active labour market policies and the level of protection on regular contracts contribute to reducing both the level of long-term unemployment and its response to aggregate unemployment.

Moreover, this increase in long-term unemployment seems to be proportional to the severity of recessions in European peripheral countries (Figure 2, panel A). In these countries, prolonged downturns, represented here by episodes in which negative output gaps are at least 2% and last for at least eight quarters, are significantly associated with a rise in long-term unemployment, and it is particularly evident for the recent economic crisis.⁵ Conversely, this relationship does not seem to hold for the other large economies considered here, where against the backdrop of adverse economic conditions the risk of unemployment persistence is lower (Figure 2, panel B). At the same time, Figure 2 also highlights that the large dispersion of unemployment rates across both groups of countries in the aftermath of economic downturns is due not only to different institutional settings, but also to the different magnitude of the shock experienced.

A major risk associated with protracted unemployment spells is that they lead to long-term unemployment, which can cause a deterioration of human capital, with the long-term unemployed detaching progressively from the labour market. A consequent reduction in wage bargaining pressure exerted by the long-term unemployed prevents real wages from adjusting to the adverse shock, with a consequent decline of total labour demand (Phelps, 1972; Blanchard and Summer, 1986; Layard and Nickell, 1987; Lindbeck and Snower, 1988). Under the effect of these dynamics, the NAIRU can rise permanently through a so-called hysteresis effect. In this respect, hysteresis can be regarded as an extreme case of high persistence of unemployment which does not return to any specific value and depends on the history of shocks (Blanchard, 2006).

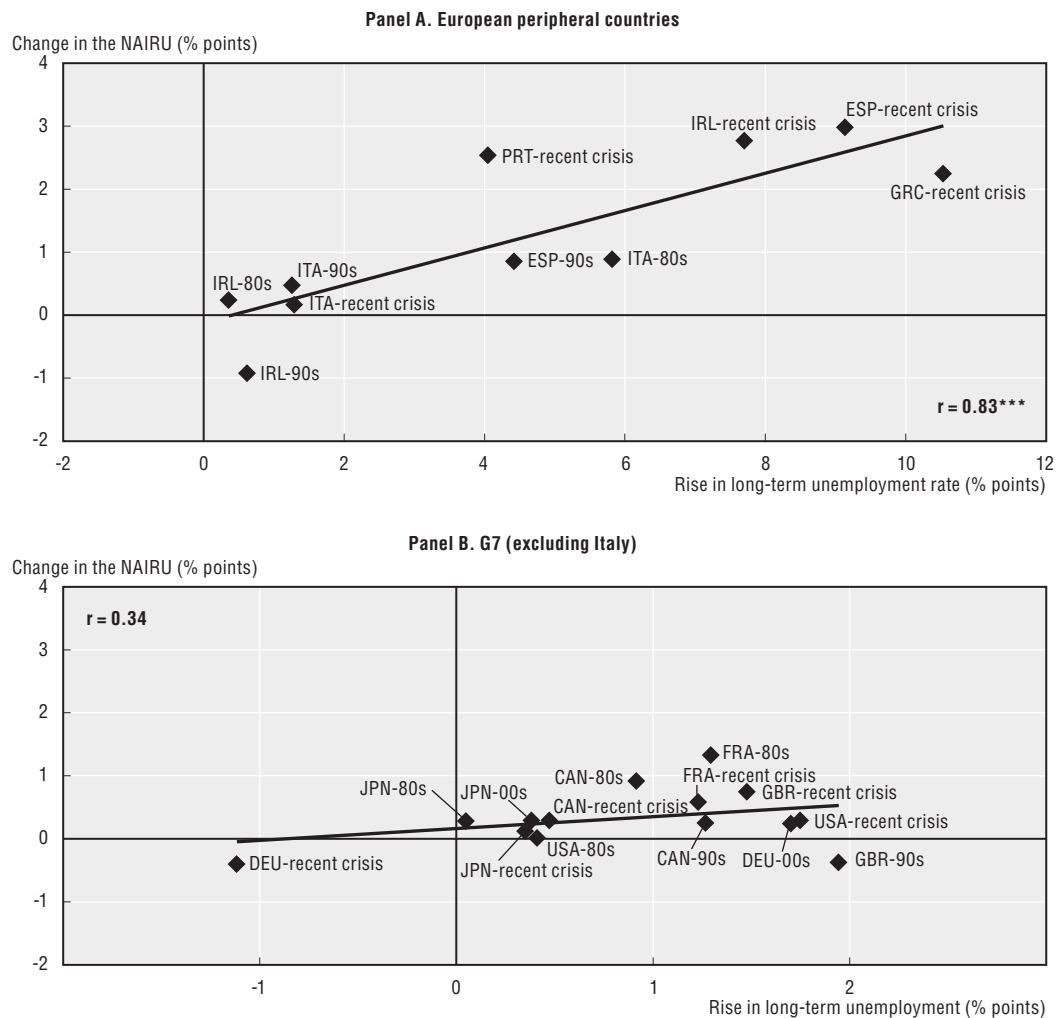
Figure 2. **Changes to long-term unemployment following economic downturns**

Notes: The scatterplot illustrates the rise in the long-term unemployment rate between the year prior to an episode of severe economic downturn to the year after the end of the episode. The episodes selected correspond to periods of negative annual output gap exceeding 2% and lasting for at least two consecutive years. *** indicates that the correlation index r is statistically significant at 1%.

Source: OECD Economic Outlook 90 Database, Eurostat, OECD Labour Force Statistics Database, OECD calculations.

The risk of hysteresis associated with protracted unemployment spells increases proportionally with the rise in long-term unemployment rates. In particular, during severe recessions as labour-market conditions worsen, the NAIRU increases significantly in line with the increase in long-term unemployment in the European periphery (Figure 3, panel A). Following the recent economic crisis, the average rise in the long-term unemployment rate was about 8 percentage points in Ireland, Greece, Spain and Portugal, while the NAIRU increased on average by 2¾ percentage points at the end of 2012 (OECD Economic Outlook, 2012/2). Among G7 countries (excluding Italy) this relationship is weaker and not statistically significant. More specifically, excluding Germany, the average rise in long-term unemployment following the recent economic downturn equals 1 percentage point and it is associated with an average increase in the NAIRU by only ½ percentage point (Figure 3, panel B).⁶

Figure 3. Relationship between long-term unemployment and the NAIRU



Notes: The scatterplot shows the change in the NAIRU associated with a rise in the long-term unemployment rate in correspondence of severe recession episodes. The episodes selected correspond to periods of negative annual output gap exceeding 2% and lasting for at least two consecutive years. *** indicates that the correlation index r is statistically significant at 1%.

Source: OECD Economic Outlook 92 Database, Eurostat, OECD Labour Force Statistics Database, OECD calculations.

5. A modified methodology for estimating the OECD NAIRU using long-term unemployment

In order to take into account the most recent developments in OECD labour markets and the risk of hysteresis,⁷ long-term unemployment has been included in the NAIRU modelling equation. This may be particularly important when changes in unemployment are large and rapid because the smoothness parameters underlying filtering methods, which may be appropriate in “normal” times, may be too restrictive. In the post-crisis period, this could lead to the under-estimation of the NAIRU and contribute to explaining the reduced statistical significance of unemployment gaps in explaining inflation.⁸ This methodological modification helps increasing both the size and the statistical significance of the coefficient on the unemployment gap in the estimated Phillips curve, while improving its goodness-of-fit. However, in line with the very few studies previously considering the inclusion of long-term unemployment to model the unemployment gap

(Box 1), the change in long-term unemployment which translates into a rise of structural unemployment differs across countries and is statistically significant only in the case of the European peripheral countries.

The change in the lagged long-term unemployment rate has been added to the first-order random walk process modelling the NAIRU in equation [4a]. More precisely, the additional variable in equation [4b] takes the form of a stochastic non-stationary time-varying drift:⁹

$$u_t^* = u_{t-1}^* + \rho \Delta u_{t-1}^{ltu} + v_t \quad [4b]$$

The coefficient ρ represents the proportion of the change in long-term unemployment that translates into an increase in structural unemployment, so capturing hysteresis effects. It is estimated by applying the Kalman filter to the same state-space model presented in Section 2 and using the same Phillips curve specification which is regularly estimated to derive the NAIRU series used by the OECD.

Box 1. Previous studies considering long-term unemployment in Phillips curve estimation

To the author's knowledge, only a few empirical studies have considered the inclusion of long-term unemployment in modelling the NAIRU. They mainly focus on European labour markets, which are traditionally more prone to hysteresis phenomena. A frequent limitation to the use of long-term unemployment in such applications is the short time span over which time series for long-term unemployment is available.

Llaudes (2005) constructs an unemployment index in which the weight of both the short-term and the long-term unemployment rate on the unemployment gap is estimated within a Phillips curve cast into a state-space framework by applying Kalman filter techniques. The results confirm that the long-term unemployed have a smaller weight than the short-term unemployed in the determination of prices and wages. The new index is employed to estimate the NAIRU and the unemployment gap entering the Phillips curve. The author finds that the relative impact of long-term unemployment on inflation varies across countries, with the difference between short- and long-term unemployment being less marked in non-European economies than European ones, suggesting that the share of long-term unemployment which is translated into structural unemployment might be higher in Europe than elsewhere.

Following Llaudes's findings, Guichard and Rusticelli (2010) assume that one-third of the increase in long-term unemployment translates into increases in the NAIRU in the United States, the United Kingdom and other non-European OECD countries, whereas a share of two-thirds is assumed for the rest of Europe. This calibration was proposed as a means of accounting for hysteresis effects in the immediate aftermath of the financial crisis. It is consistent with the rationale that the long-term unemployed exert significantly less pressure on wages than the short-term unemployed and that the relative impact of the long-term unemployed on wages and prices varies across countries.

Applying Kalman filter techniques to a similar modelling formulation adopted by the OECD, Logeay and Tober (2006) focus on the euro area as a whole and find that the lagged unemployment rate for the euro area plays a significant role in supporting the importance of hysteresis effects. More precisely, the coefficient of the lagged unemployment rate in the NAIRU estimating equation is significant and equals 0.26, suggesting the presence of hysteresis at the aggregate euro area level (similar results were found in Jaeger and Parkinson, 1994). Due to the limited availability of long series on long-term unemployment of individual European countries, Germany is the only country for which hysteresis effects have been investigated through the inclusion of long-term unemployment in the NAIRU equation. The authors estimate a long-term unemployment coefficient of 0.82,

Box 1. Previous studies considering long-term unemployment in Phillips curve estimation (cont.)

which multiplied by one-third, i.e. the share of long-term unemployment in total unemployment in Germany, gives a corresponding coefficient of 0.25 for lagged unemployment and it is therefore consistent with the coefficient of 0.26 previously estimated for the euro area lagged unemployment rate.

Using a similar framework to the standard Phillips curve set-up estimated by the OECD, Kajuth (2010) adds the long-term unemployment rate to the second-order random walk process with time varying drift modelling the NAIRU for Germany. The author estimates a significant coefficient equal to 0.28 capturing the impact of long-term unemployment on the NAIRU, which reaches 8% in 2009, the end of the estimation period. This result is similar to the OECD estimate of 7.6% (Guichard and Rusticelli, 2011) which was obtained applying a similar methodology, but without the adjustment for hysteresis made by including long-term unemployment.

6. Empirical results from applying the modified methodology for the NAIRU estimation

Both the standard and the modified modelling procedures have been applied and compared on a common data sample, which has been shortened given that data on long-term unemployment are not available over the full sample. The results are, however, mixed: the inclusion of long-term unemployment in modelling the NAIRU appears to improve the fit of the Phillips curve equation for some countries, but not all.

More specifically, the estimation results imply two groups of countries, and it is only for the first of these that the new methodology demonstrates an effective improvement to the standard OECD method for estimating the NAIRU (Table 2).

For the first group of countries – Greece, Ireland, Italy, Portugal and Spain (see upper half of Table 2) – where long-term has typically risen the most while capturing hysteresis effects in the labour market, the new methodology demonstrates an effective

Table 2. Estimated coefficients from the standard and modified methodology to NAIRU modelling

	Change in lagged long-term unemployment (ρ)	Standard methodology	Modified methodology	Difference in NAIRU estimates in 2012 (% pts)
		Unemployment gap (β)	Unemployment gap (β)	
Greece	0.43***	-0.035	-0.049**	1.64
Ireland	0.75***	-0.014	-0.034*	3.71
Italy	0.72***	-0.022*	-0.064***	0.68
Portugal	0.78***	-0.059*	-0.065*	0.15
Spain	0.73***	-0.017	-0.020*	-1.07
Canada	1.53	-0.116***	-0.143***	0.21
France	0.74	-0.084***	-0.097***	-0.03
Germany	0.47	-0.054**	-0.067**	-0.42
Japan	1.65	-0.176*	-0.131*	0.34
United Kingdom	0.58	-0.098**	-0.101**	0.23
United States	0.74	-0.031***	-0.045***	1.40

Notes: The NAIRU estimation with both the standard and the modified methodologies has been performed on the same observations sample according to the availability of data on long-term unemployment. *, **, *** indicate statistically significant at 10%, 5% and 1%, respectively.

improvement to the standard OECD method for estimating the NAIRU in a number of respects:

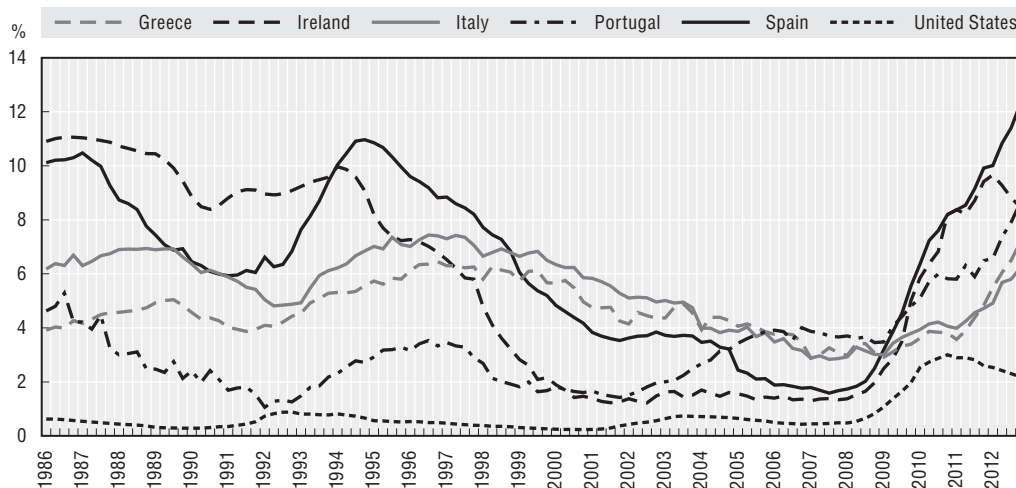
- The estimated coefficient ρ on the change in lagged long-term unemployment is statistically significant and of a plausible magnitude. It varies from 0.43 to 0.78 with an average value of 0.68, this latter in line with Guichard and Rusticelli (2010), who assume that the share of the change in long-term unemployment which is translated into a rise of structural unemployment equals two-thirds for the Continental European countries.
- The fit of the Phillips curve is improved as indicated by the increased magnitude and statistical significance of the slope coefficient β on the unemployment gap; notably for Greece, Ireland and Spain, this coefficient is statistically significant only when long-term unemployment is included in the new estimation technique and becomes much larger in the case of Italy and, to a lesser extent, for Portugal.
- Differences in the NAIRU estimates are meaningfully different from an economic perspective; for 2012 they increase by a minimum of 0.2 percentage points in the case of Portugal to a maximum of 3.7 percentage points for Ireland, while the estimate for Spain is revised downward by 1 percentage point when long-term unemployment is included in the modelling equation.

For the second group of countries – Canada, France, Germany, Japan, the United Kingdom and the United States (see lower half of Table 2) – there is a less compelling case that the new methodology provides an improvement to the standard OECD methodology:

- The coefficient of long-term unemployment is either not significant or takes implausible values (as in case of Canada and Japan) indicating that this latter does not play any relevant role in explaining the NAIRU increase following the financial crisis.
- The new methodology does not improve the fit of the Phillips curve when compared with the standard approach. The Wald test, performed on the unemployment gap coefficients β , does not reject the null hypothesis of equality of the estimated coefficients from the application of the standard and the new approach for any country. These findings seem to cast some doubt on whether the new methodology should be applied mechanically in the same way across all countries. For this reason, more results are shown only for those countries where the new approach is successful.
- NAIRU point estimates for 2012 are revised upward only marginally, by 0.2 percentage points on average for all countries except Germany and the United States, for which the size of the revision is larger. In particular, the NAIRU is estimated to decrease by $\frac{1}{2}$ percentage point for Germany, while for the United States the upwards revision is nearly $1\frac{1}{2}$ percentage points.
- The United States is an interesting intermediate case as the statistical evidence for the proposed methodology is marginal, but the economic relevance of the revised NAIRU is major. The statistical significance of the coefficient associated with the change in long-term unemployment is very low, although the coefficient of the unemployment gap is larger and more significant when hysteresis effects on the NAIRU are taken into consideration in the new specification (Table 2). The revised NAIRU for the final quarter of 2012 is 8%, which is significantly higher than both the point estimate obtained by applying the standard OECD methodology and the latest CBO estimate,¹⁰ by 1.4 and 2.0 percentage points, respectively. The higher estimate would imply that the unemployment gap was already closed by the end of 2012. These results can be considered in light of the recent debate on the role played by the prolonged high rate of

long-term unemployment generated by the recent economic crisis in determining the level of inflation and affecting monetary policy strategies in the United States. More specifically, despite the large and relatively new persistence of long-term unemployment compared with the euro area periphery (Figure 4), most recent studies conclude that long-term unemployment only exerts a limited pressure if any on inflation and has a negligible effect in holding down wages (Gordon, 2013; Watson, 2014; Krueger et al., 2014). Thus to the extent that there has been an increase in US long-term unemployment which is unprecedented in the post-war period, some increase in the NAIRU seems likely but not as large as the one estimated by the new methodology.

Figure 4. **Long-term unemployment developments in the euro area periphery and the United States**

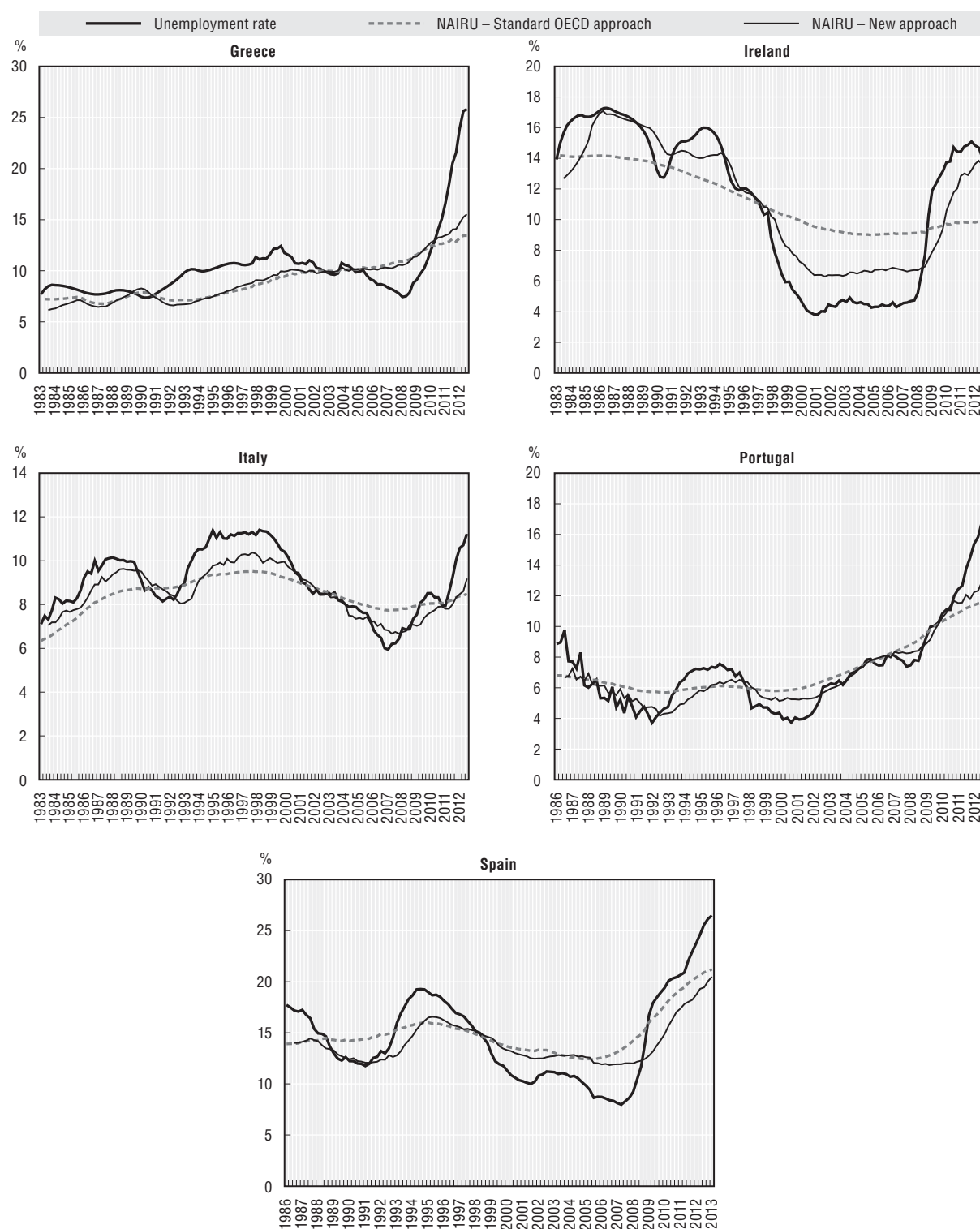


Source: OECD Labour Force Statistics Database.

The increase in the estimated NAIRU since the start of the financial crisis is the largest for Spain, followed by Portugal and Greece (Figure 5). In these countries, the NAIRU has risen by 6 percentage points for Spain and 2½ percentage points for Greece and Portugal when applying the standard OECD approach. The modified estimation technique provides an even larger rise in the NAIRU estimate in the aftermath of the crisis. More precisely, in the period 2008-12, the average increase in the NAIRU equals 5 percentage points, from a minimum rise of 2 percentage points in case of Italy to a maximum of 7 percentage points for Spain. On the other hand, the revised methodology also implies revised NAIRU estimates in the pre-crisis period as a consequence of unusually low long-term unemployment rates. On average, for the first group of countries, the new NAIRU is 0.7 percentage points lower than the estimate obtained through the standard methodology over the period 2005-08, with the largest downward correction by 1¼ percentage points for Ireland and Spain.

The new Phillips curve better explains changes in inflation (Table 3). In the case of Spain, Ireland and Greece, the Phillips curve estimated on a dataset starting from mid-1980s, i.e. shorter than that considered for the regular *Economic Outlook* updates, shows a weaker relationship between inflation and the unemployment gap (as indicated by the unemployment gap coefficient no longer being statistically significant), with unemployment rising and core inflation close to zero (second column in Table 3). However,

Figure 5. **NAIRU estimates following the standard and the new approach**



Source: OECD Economic Outlook 93 Database, OECD estimates.

Table 3. **Phillips curve estimation results**

	Greece			Ireland		
Sample period	1975Q1-2013Q1	1983Q3-2012Q4	1983Q3-2012Q4	1978Q3-2013Q1	1986Q2-2012Q4	1986Q2-2012Q4
Specification	Standard	Standard	New	Standard	Standard	New
<i>constant</i>						
$\Delta\pi_{t-1}$	-0.51***	-0.36***	-0.38***	-0.62***	-0.28***	-0.29***
$\Delta\pi_{t-2}$	-0.35***	-0.39***	-0.39***	-0.25***	-0.24***	-0.25***
$\Delta\pi_{t-3}$	-0.15**	-0.25***	-0.25***	0.30***		
$\Delta\pi_{t-4}$	-0.19***	-0.27***	-0.29***			
$u_t - u_t^*$	-0.05*	-0.04	-0.05**	-0.03**	-0.01	-0.03*
$\Delta\pi_{t-1}^{imp}$				0.20***	0.13***	0.14***
$\Delta\pi_{t-4}^{imp}$						
$\Delta\pi_{t-1}^{oil}$						
Dummy				2009Q1, 2009Q4	2009Q1, 2009Q4	2009Q1, 2009Q4
Sacrifice ratio	2.8	4.0	2.9	3.3	6.6	2.8
Adj. R^2	0.26	0.17	0.18	0.58	0.39	0.39
	Italy			Portugal		
Sample period	1968Q4-2013Q1	1983Q3-2012Q4	1983Q3-2012Q4	1980Q1-2013Q1	1986Q2-2012Q4	1986Q2-2012Q4
Specification	Standard	Standard	New	Standard	Standard	New
<i>constant</i>						
$\Delta\pi_{t-1}$	-0.15**	-0.47***	-0.49***	-0.50***	-0.38***	-0.39***
$\Delta\pi_{t-2}$	-0.16**	-0.28***	-0.36***	-0.26***	-0.34***	-0.35***
$\Delta\pi_{t-3}$						
$\Delta\pi_{t-4}$						
$u_t - u_t^*$	-0.05*	-0.02*	-0.06***	-0.09**	-0.06*	-0.06*
$\Delta\pi_{t-1}^{imp}$	0.18***	0.12***	0.13***	0.29***	0.19***	0.20***
$\Delta\pi_{t-4}^{imp}$						
$\Delta\pi_{t-1}^{oil}$						
Dummy	1975Q1, 1976Q1			1983Q2-Q3		
Sacrifice ratio	1.6	4.9	1.8	1.2	1.8	1.7
Adj. R^2	0.29	0.18	0.22	0.25	0.22	0.20
	Spain					
Sample period	1978Q1-2013Q1	1986Q2-2012Q4	1986Q2-2012Q4			
Specification	Standard	Standard	New			
<i>constant</i>	-0.05*	-0.02				
$\Delta\pi_{t-1}$	-0.67***	-0.71***	-0.71***			
$\Delta\pi_{t-2}$	-0.17***	-0.17***	-0.17***			
$\Delta\pi_{t-3}$						
$\Delta\pi_{t-4}$						
$u_t - u_t^*$	-0.03**	-0.02	-0.02*			
$\Delta\pi_{t-1}^{imp}$						
$\Delta\pi_{t-4}^{imp}$						
$\Delta\pi_{t-1}^{oil}$	0.06*	0.07**	0.07*			
Dummy	1986Q1, 2002Q1	2002Q1	2002Q1			
Sacrifice ratio	3.8	7.0	5.9			
Adj. R^2	0.58	0.64	0.63			

Notes: *, **, *** indicate statistically significant at 10%, 5% and 1%, respectively.

the new approach seems to strengthen this relationship and to improve the goodness-of-fit of the estimated equation for all countries (third column in Table 3). Moreover, the stability of the parameters estimated in the new Phillips curve has been tested against two alternative hypotheses of a structural change in the relationship between inflation and the unemployment gap occurring a) in 1998, assumed as starting period for the explicit inflation-targeting pursued by the European Central bank in the euro area; b) in 2009, at the onset of the global crisis blow to labour markets. The Chow test fails to reject the null hypothesis of absence of a structural break for all European peripheral countries considered, thus confirming the robustness of the estimated Phillips curve specification when long-term unemployment is included to account for hysteresis risks.

7. Conclusions

This paper presents a modification of the existing OECD approach to estimate the NAIRU in the context of a flattening Phillips curve, where the reduced sensitivity of the inflation process to the economic activity is translated into a loss of statistical significance of the unemployment gap. The proposed alternative aims at strengthening the relationship between inflation and labour market developments by considering the impact on aggregate unemployment associated with hysteresis effects, which are particularly important in the latest economic downturn. For this purpose, recent changes in long-term unemployment have been taken as an indication of possible risks of permanent rises in the structural unemployment and included directly in the NAIRU modelling equation. Only a selection of OECD countries has been examined: the European peripheral ones, where risks of recent hysteresis dynamics are the highest, and the G7 economies which form a control group to verify whether the proposed methodology can be applied more generally.

This methodological modification helps increasing both the size and the statistical significance of the coefficient on the unemployment gap in the estimated Phillips curve, while improving its goodness-of-fit and warranting the stability of the estimated parameters. However, the change in long-term unemployment which translates into a rise of structural unemployment differs across countries and is statistically significant only in the case of the European peripheral countries. On average for these countries – namely Greece, Ireland, Italy, Portugal and Spain – the new NAIRU is between 0.2 and 3.7 percentage points higher than the standard OECD NAIRU in 2012, while the estimated coefficient on the change in lagged long-term unemployment is statistically significant and of a plausible magnitude. For the second group of countries – Canada, France, Germany, Japan, the United Kingdom and the United States – there is a less compelling case that the new methodology provides an improvement to the standard OECD methodology.

Following these encouraging results and the need to improve both the macroeconomic and statistical explanatory power of a flattening Phillips curve, these modifications could be extended to a broader set of OECD countries. Similarly, new research could explore whether different lengths of unemployment duration have a different influence on the unemployment gap estimate, by considering an index as in Llaudes (2005) for instance, where the pressure on wage and prices is disentangled by short- and long-term unemployment. An alternative hypothesis for understanding the flattening of the Phillips curve would be to examine the existence of structural breaks directly affecting the inflation process arising from the establishment of explicit inflation targeting and its role in anchoring inflation expectations (Rusticelli et al., 2014).

Notes

1. Due to the longer span of data, headline CPI inflation is used for Canada, Greece, Portugal and the United Kingdom.
2. The sum of the autoregressive parameters is constrained to vary between 0.7 and 0.9 in order to ensure the stationarity of the unemployment gap.
3. Following the approach of Staiger et al. (1997), Laubach (2001) and Llaudes (2005), these parameters are empirically selected on an individual country basis and fixed throughout the estimated state space model in order to preserve good distributive properties for the unobserved components. The estimation of the signal-to-noise ratio can lead to very flat NAIRUs due to the so-called pile-up problem (Stock, 1994): pure maximum likelihood estimation methods of the variance of non-stationary processes yield estimates which are downward biased towards zero when the true unknown parameter is small, as in the case of the standard deviation of the innovation process in the random walk describing the NAIRU model. More details on the initialisation procedure for the state parameters and the selection of the variances of the error terms can be found in Guichard and Rusticelli (2011).
4. It should be noted that the unemployment gap is computed as difference to constant trend unemployment, calculated as historical mean over the estimation period. Therefore, the slope of the Phillips curve and the size of the estimated unemployment gap may be biased by unaccounted movements in the unobserved structural unemployment.
5. The episodes of large negative output gaps associated with the recent financial crisis differ from all other episodes considered since they are not concluded and aggregate unemployment is not back to its natural rate.
6. When the episodes of economic recession associated with the recent crisis are removed from the observation sample, the correlation index r is reduced but still equals 0.72 for the European peripheral countries and -0.10 instead of 0.34 for the G7 economies (excluding Italy).
7. The Perron test of stationarity in the presence of structural breaks has been performed and does not reject the null hypothesis of the existence of a unit root for any of the countries considered when a drift and a trend are also included. This result further suggests that hysteresis is an intrinsic characteristic of the European labour market even when changes in labour market institutions and policies are taken into consideration.
8. This can only be a partial explanation, as the reduced significance of gap terms in explaining inflation is a well-documented effect over the pre-crisis period.
9. The augmented Dickey-Fuller test of stationarity performed on the level of long-term unemployment confirms the presence of a unit root for all countries. As a consequence, the NAIRU model follows instead a second-order random walk process (see Laubach, 2001 or Llaudes, 2005 for references on I(2) models for NAIRU specifications).
10. United States Congressional Budget Office's February 2014 report: *The Budget and Economic Outlook: 2014 to 2024*.

References

- Bataa, E., D.R. Osborn, M. Sensier and D. van Dijk (2014), "Identifying Changes in Mean, Seasonality, Persistence and Volatility for G7 and Euro Area Inflation", *Oxford Bulletin of Economics and Statistics*, Vol. 76, Issue 3, pp. 360-388.
- Benkovskis, K., et al. (2011), "Assessing the Sensitivity of Inflation to Economic Activity", *European Central Bank Working Paper Series*, No. 1357.
- Blanchard, O.J. (2006), "European Unemployment: The Evolution of Facts and Ideas", *Economic Policy*, CEPR, CES, MSH, Vol. 21(45), pp. 5-59.
- Blanchard, O.J. and L.H. Summers (1986), "Hysteresis and the European Unemployment Problem", *NBER Macroeconomic Annual*, MIT Press, Cambridge, MA.
- Blanchard, O.J. and J. Galí (2010), "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment", *American Economic Journal: Macroeconomics*, 2 (2), pp. 1-33.
- Borio, C. and A. Filardo (2007), "Globalisation and Inflation: New Cross-country Evidence on the Global Determinants of Domestic Inflation", *Bank for International Settlements Working Papers*, No. 227.

- Elsby, M., B. Hobijn and A. Sahin (2010), "The Labor Market in the Great Recession", *NBER Working Paper*, No. 15979.
- Gianella, C., et al. (2008), "What Drives the NAIRU? Evidence from a Panel of OECD Countries", *OECD Economics Department Working Papers*, No. 649, OECD Publishing, <http://dx.doi.org/10.1787/231764364351>.
- Gordon, R.J. (2013), "The Phillips Curve is Alive and Well: Inflation and the NAIRU during the Slow Recovery", *NBER Working Paper*, No. 19390.
- Guichard, S. and E. Rusticelli (2010), "Assessing the Impact of the Financial Crisis on Structural Unemployment in OECD Countries", *OECD Economics Department Working Papers*, No. 767, OECD Publishing, <http://dx.doi.org/10.1787/5kmftp8khfjg-en>.
- Guichard, S. and E. Rusticelli (2011), "Reassessing the NAIRUs after the Crisis", *OECD Economics Department Working Papers*, No. 918, OECD Publishing, <http://dx.doi.org/10.1787/5kg0kp712f6l-en>.
- IMF (2013), "The Dog that didn't Bark: Has Inflation been Muzzled or was it Just Sleeping", *World Economic Outlook*, April.
- Jaeger, A. and M. Parkinson (1994), "Some Evidence on Hysteresis in Unemployment Rates", *European Economic Review*, No. 38, pp. 329-342.
- Kajuth, F. (2010), "NAIRU Estimates for Germany: New Evidence on the Inflation-unemployment Trade-off", *Discussion Paper Series 1: Economic Studies*, No. 19, Deutsche Bundesbank, Research Centre.
- Krueger, A.B., J. Cramer and D. Cho (2014), "Are the Long-Term Unemployed on the Margins of the Labor Market?", *Brookings Papers on Economic Activity*, Spring.
- Kuttner, K. and T. Robinson (2010), "Understanding the Flattening Phillips Curve", *The North American Journal of Economics and Finance*, Elsevier, Vol. 21(2), pp. 110-125.
- Laubach, T. (2001), "Measuring the NAIRU: Evidence From Seven Economies", *The Review of Economics and Statistics*, MIT Press, Vol. 83(2), pp. 218-231.
- Layard, R. and S. Nickell (1987), "The Labour Market", in: Dornbusch, Rudiger and Layard, Richard, (eds.), *The Performance of the British Economy*, Clarendon, Oxford, UK.
- Lindbeck, A. and D.J. Snower (1988), *The Insider-Outsider Theory of Employment and Unemployment*, Cambridge: Mass, MIT Press.
- Llaudes, R. (2005), "The Phillips Curve and Long-Term Unemployment", *European Central Bank Working Paper Series*, No. 441.
- Logeay, C. and S. Tober (2006), "Hysteresis and the NAIRU in the Euro Area", *Scottish Journal of Political Economy*, Vol. 53, No. 4, pp. 409-429.
- OECD (2011), *OECD Employment Outlook*, Vol. 2011, OECD Publishing, http://dx.doi.org/10.1787/empl_outlook-2011-en.
- OECD (2012), *OECD Economic Outlook*, Vol. 2012/2, OECD Publishing, http://dx.doi.org/10.1787/eco_outlook-v2012-2-en.
- OECD (2014), *OECD Employment Outlook*, Vol. 2014, OECD Publishing, http://dx.doi.org/10.1787/empl_outlook-2014-en.
- Phelps, E.S. (1972), *Inflation Policy and Unemployment Theory: The Cost-Benefit Approach to Monetary Planning*, Macmillan, London.
- Razin, A. and A. Binyamini (2007), "Flattening of the Short-run Trade-off between Inflation and Domestic Activity: The Analytics of the Effects of Globalization", *Kiel Working Papers*, No. 1363, Kiel Institute for the World Economy.
- Roberts, J. (2006), "Monetary Policy and Inflation Dynamics", *International Journal of Central Banking* 2, pp. 193-230.
- Rusticelli, E., D. Turner and M.C. Cavalleri (2014), "Incorporating Anchored Inflation Expectations in the Phillips Curve and in the Derivation of Measures of OECD Equilibrium Unemployment", *OECD Economics Department Working Papers*.
- Sargent, T., N. Williams and T. Zha (2009), "The Conquest of South American Inflation", *Journal of Political Economy*, University of Chicago Press, Vol. 117(2), pp. 211-256.
- Smets, F. and R. Wouters (2007), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach", *American Economic Review*, Vol. 97(3), pp. 586-606.

- Staiger, D., J. Stock and M. Watson (1997), "How Precise are Estimates of the Natural Rate of Unemployment?", in C. Romer and D. Romer (eds), *Reducing Inflation: Motivation and Strategy*, Chicago, Chicago University Press and NBER.
- Stock, J.H. (1994), "Unit Roots, Structural Breaks and Trends", in R. Engle and D. MacFadden (eds), *Handbook of Econometrics*, Vol. 4, Amsterdam, Elsevier, pp. 2739-2841.
- Watson, M.W. (2014), "Inflation Persistence, the NAIRU, and the Great Recession", *American Economic Review*, Vol. 104, No. 5 (May).
- Williams, J.C. (2006), "Inflation Persistence in an Era of Well-anchored Inflation Expectations", *Federal Reserve Bank of San Francisco Economic Letter*, October 13 issue.
- Yellen, J.L. (2012), "Perspectives on Monetary Policy", speech delivered at the Boston Economic Club dinner, Boston, 6 June.

Foreign direct investment and reverse technology spillovers: The effect on total factor productivity

by

Edmund Amann and Swati Virmani*

The paper analyses the “feedback effect” of Foreign Direct Investment (FDI) on Total Factor Productivity (TFP) growth in emerging economies via technology spillovers across borders. We study the effect of R-D spillovers resulting from outward FDI flows from 18 emerging economies into 34 OECD countries over the 1990-2010 period, comparing the impact with that of spillovers resulting from inward FDI flows. The result confirms that FDI enhances productivity growth; however the impact is much larger when R-D-intensive developed countries invest in the emerging economies than the other way round. Country-specific bilateral elasticities also support this outcome.

JEL classification: F21, F43, F62, O47.

Keywords: Outward FDI, Inward FDI, Reverse technology spillovers, Total factor productivity.

* The authors are Reader in Development Economics and Head of Economics (Teaching and Operations); and PhD Student, Economics, School of Social Sciences, respectively, at the University of Manchester, UK. They would like to thank Bruno Van Pottelsberghe de la Potterie and Giuseppe Nicoletti for comments on a previous version. The opinions expressed and arguments employed in this paper are the responsibility of the authors and do not necessarily reflect the official views of the OECD or of the governments of its member and partner countries.
Email: edmund.amann@manchester.ac.uk; swati.virmani@postgrad.manchester.ac.uk.

1. Introduction

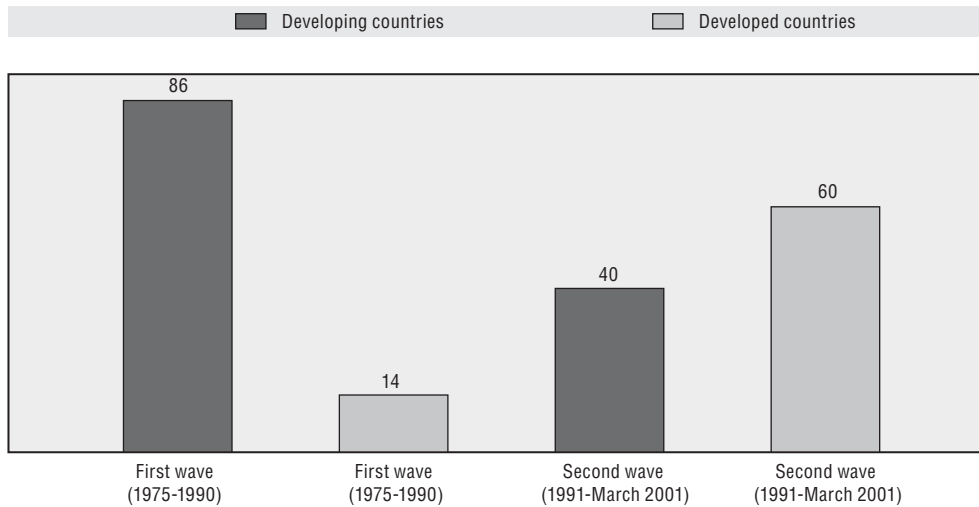
The term “knowledge spillover” refers to the process by which one inventor learns from the research outcomes of others’ projects and is able to enhance their own research productivity without fully compensating the other inventors for the value of such learning (Branstetter, 2006). True knowledge spillovers have the potential to allow for further innovation. In pursuit of such spillovers, it is possible that firms may take investment decisions so as to learn from other firms’ research activities. This is especially true for a developing and emerging country undertaking Outward Foreign Direct Investment (OFDI) in R-D intensive developed countries. While the advanced country firms, such as those of Japan and Korea, have set up their own technological profiles and brands over long periods of time, newly emerging nations often tend to leapfrog by acquiring a well-established corporation, or setting up R-D centres in areas of excellence.

The literature identifies two waves of outward investment: the first during the 1960s and 1970s, and the second from 1980 onwards (UNCTAD, 2007). In the second wave, a diversified pattern of investment was emerging, with an increasing participation of developing nations undertaking OFDI in developed nations. The aim here was to attain strategic assets and markets, acquiring brands, and accessing technologies. Such an OFDI played a mediating role in the “knowledge spillover”, that was traditionally unidirectional, taking place from an innovation-driven developed nation investing in a developing country. In other words, outward investment from a less capital intensive developing nation was now acting as a channel for transferring technology. Dunning recognises this second wave of OFDI as resulting from liberalisation of markets, globalisation of economic activity and dramatic technological advancement within sectors; factors that have affected the structure of the world economy. He also postulates that during this period the countries that came out in front were those that had moved along the Investment Development Path (IDP),¹ and experienced rapid economic growth and restructuring (UNCTAD, 2005b).

In the case of India, for example, the second wave began in the 1990s when enterprises were mostly using Mergers and Acquisitions (M&A) to venture abroad in order to access technology. The lion’s share of such investment was in developed countries, dominated by the United States (24%) and the United Kingdom (27%) (UNCTAD, 2007). Indian companies such as Infosys, Aditya Birla and HCL Technologies invested in the United States. Figure 1 shows the diversity between the first and second phase of India’s investment in terms of the geographical distribution.

Overall, during this period, OFDI from emerging economies rose from USD 149 billion in 1990 to USD 1.4 trillion in 2005, accounting for 13% of world OFDI stock in 2005 (UNCTAD, 2007). More economies became global players, and among the strongest were from Asia. Enterprises from East Asia (China), South-East Asia (Malaysia, Thailand and Singapore), and South Asia (India) were the most active. Following them were the developing regions of Latin America and the Caribbean (Brazil and Argentina). Also,

Figure 1. **India's OFDI flows by geographical distribution in first (1975-90) and second (1991-March 2001) waves (in %)**



Source: UNCTAD, 2007.

enterprises from the transition economies were engaged in OFDI (Russian Federation). Table 1 highlights the OFDI flows from four major host countries for emerging multinationals, the BRIC.

Table 1. **OFDI flows in USD million, 1990-2010**

	Brazil	Russia	India	China	BRIC, % of World
1990	624.6	-	6	830	0.605
1991	1 015	-	-11	913	0.968
1992	136.7	1 566	24	4 000	2.826
1993	492.3	1 022	0.35	4 400	2.438
1994	689.9	281.37	82	2 000	1.064
1995	1 095.64	605.78	119	2 000	1.052
1996	-469.06	922.82	240	2 114	0.706
1997	1 115.56	3 183.91	113	2 562.49	1.461
1998	2 854.01	1 269.75	47	2 633.81	0.987
1999	1 690.41	2 207.62	80	1 774.31	0.529
2000	2 281.59	3 176.78	514.45	915.78	0.562
2001	-2 257.59	2 532.58	1 397.44	6 885.40	1.145
2002	2 482.11	3 532.65	1 678.04	2 518.41	1.932
2003	249.3	9 727.13	1 875.78	2 854.65	2.577
2004	9 806.99	13 782.03	2 175.37	5 497.99	3.377
2005	2 516.70	12 767.47	2 985.49	12 261.17	3.436
2006	28 202.49	23 151.00	14 284.99	2 1160	6.134
2007	7 066.66	45 915.5	19 594.36	22 468.86	4.324
2008	20 457.07	55 593.5	19 256.5	52 150	7.488
2009	-10 084.23	43 665	15 927.1	56 530	9.024
2010	11 587.57	52 523	13 151	68 811	10.064

Notes: FDI includes the three following components: equity capital, reinvested earnings and intra-company loans. Data on FDI flows are presented on net bases (capital transactions' credits less debits between direct investors and their foreign affiliates). Net decreases in assets or net increases in liabilities are recorded as credits, while net increases in assets or net decreases in liabilities are recorded as debits. Hence, FDI flows with a negative sign indicate that at least one of the three components of FDI is negative and not offset by positive amounts of the remaining components. This is called reverse investment or disinvestment.

Source: UNCTAD Stat.

A key driver of such outward investment has been the competitive pressure along with a number of home and host country factors, such as home market growth constraints, lower production costs and availability of natural resources in the host nations. But recently during the current wave of internationalisation, building of brand names, accessing most sophisticated and advanced technologies, and acquiring R-D facilities through OFDI have become more notable features amongst the emerging nations. Moreover, the research-linked investments in the centres of innovation in the developed countries have been an important source of rapid technology diffusion. Good examples here are Lenovo's (China) acquisition of IBM's personal computer division (United States) and Cordlife's (Singapore) acquisition of Cytomatrix (United States) (UNCTAD, 2007). China has considered the Joint Ventures as an important source of new foreign technology and management skills ever since 1979 when the Law of Joint Ventures was promulgated. In the case of Brazil, Embraer linked up with Aeritalia and Aermacchi of Italy in 1981 to manufacture the sophisticated AMX-fighter bomber (Tolentino, 1993). This form of investment is a part of the cumulative process through which developing countries' firms gain access to advanced technology, and further combine and localise it with their indigenous technological capabilities.

Accessing foreign technology may also take the form of establishing R-D centres in developed countries. Looking at some of the Chinese companies, Huawei Technologies and ZTE Corporation established their R-D centres in Sweden; Guangdong Glanz Group Co. set up an R-D centre in Seattle, and Haier has an R-D centre in Germany and a design centre in Boston (United States). Amongst the Indian companies, Ranbaxy Laboratories has set up R-D centres in various countries including the United States (UNCTAD, 2004). The Metal Leve firm of Brazil established a research centre in Ann Arbor, Michigan, to acquire the latest technology for manufacturing auto engine pistons (Tolentino, 1993).

Our argument of strategic asset seeking behaviour is an extension to the literature supporting the "learning by exporting" effect (Bernard and Jensen, 1999; Clerides et al., 1998). OFDI enables firms to not just enter into new markets but also to gain access to foreign technology, and as a result, the entire domestic economy benefits due to increased efficiency of the investing firms and the associated spillovers to the local firms at home (Herzer, 2012). Thus, OFDI benefits the home country at least in terms of the technology transfers that it generates.

In contrast to a plentiful literature on R-D spillovers, there is only a fragmented body of work on the effects of such technology transfers on the home country. Research here is based on the impact of FDI flows from developed countries into developing ones (North-South FDI). Moreover, previous studies have been restricted to individual firms or industries and have not captured the overall macroeconomic effects on the economy as a whole. In addition, those studies are mostly based on firm level data for manufacturing that excludes the FDI in the service sector.

Against this background, this paper's contribution centres on highlighting not only the technology diffusion effect of OFDI, but also the productivity-enhancing effect on emerging economies at the macroeconomic level. The basic objective is to explore whether OFDI serves as a means of generating knowledge flows that affect the growth of TFP. Furthermore, the study entails a comparison of whether FDI transfers technology in both directions, i.e. if a country's productivity is increased not only when R-D-intensive countries invest in it, but also when it invests in R-D-intensive foreign countries. In other

words, the paper tries to distinguish between the impact of R-D spillovers resulting from North-South and South-North FDI flows.²

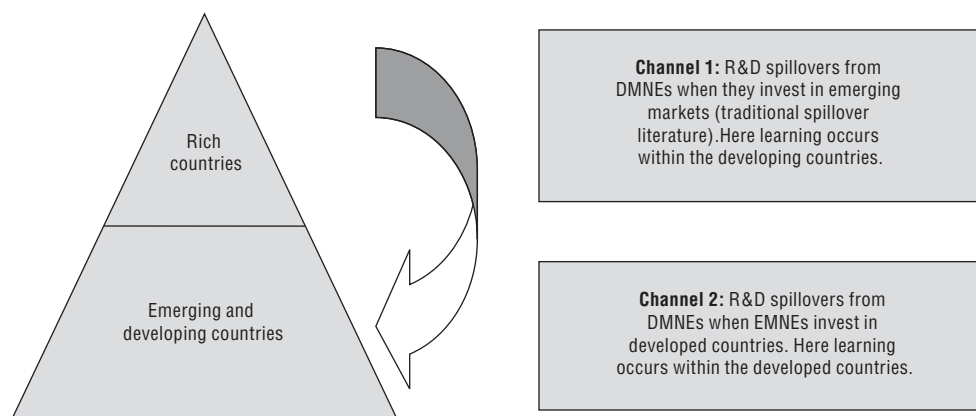
The analysis involves looking at the OFDI flows from 18 emerging economies into 34 OECD countries over the 1990-2010 period and Inward FDI (IFDI) flows into those 18 emerging economies from the OECD nations. We apply the methodology suggested by Van Pottelsberghe de la Potterie and Lichtenberg (2001). However, unlike their study – which looks at FDI flows between 13 industrialised countries (i.e. North-North FDI only) – our paper contributes more extensively by studying both South-North and North-South FDI flows. We also extend the analysis by including catalytic factors that affect TFP, such as human capital, in order to get better estimates of output elasticity of foreign R-D spillovers. According to the UNESCO (1993) *Statistical Year Book*, R-D-expenditures in OECD member countries amounted to 96% of the entire R-D world expenditures in 1990, thus justifying the choice of OECD nations as the investing partners for our analysis.

The paper is structured as follows: Section 2 describes how FDI is linked to technology spillovers. It also discusses the mechanisms that drive the spillovers. Section 3 then provides an overview of the “feedback effect”, i.e. the impact on TFP growth. Section 4 highlights the methodology and data, Section 5 outlines the results, and finally, Section 6 concludes the study with policy implications.

2. Foreign direct investment and R-D spillovers

A vast literature surrounds how firms from developed countries generate technology spillovers for developing country firms. Figure 2 identifies two main channels for such R-D spillovers, both focusing on the trickling down of technology from developed to developing countries. The first channel is based on IFDI into the emerging economies from developed countries’ firms. This is called the Traditional Channel. The second channel entails the OFDI flows from the emerging economies into the developed countries. As the figure illustrates, the traditional channel overlooks the possibility that Emerging Market Multinational Enterprises (EMNEs) could also capture spillovers from Developed Country Multinational Enterprises (DMNEs), when the former invest in the developed countries. In other words, learning could also occur in the developed economy as a result of investment from EMNEs that are motivated by the desire to obtain intangible assets, such as technology. However, in both cases, the DMNEs act as the so-called “teacher”.

Figure 2. **Two channels of R-D spillovers**



Note: EMNEs = Emerging Market Multinational Enterprises, DMNEs = Developed Country Multinational Enterprises.

Source: Author's elaboration and Govindarajan and Ramamurti (2011).

This new spillover channel is based on the recent identification of the investing country learning from local firms in the host country and acquiring knowledge spillovers at the host sites. This could be especially true in the case of an outward investment into a host country that is more capital or R&D-intensive than the home country. The technology-sourcing occurs mainly when firms try to gain access to foreign technology by either acquiring foreign firms or establishing R-D facilities in “Foreign Centers of Excellence” (Herzer, 2012). These firms then acquire new technological know-how and transfer it to the parent company in the home country.

A number of case studies have been carried out, empirically substantiating these knowledge flows, as summarised in Table 2.

Table 2. Summary results of previous studies

Author(s) and year of study	Findings
1. Pradhan and Singh (2009)	Examined OFDI in the Indian Automotive Industry during 1988-2008 and suggested a favourable impact on the R-D intensity. Their study supported that OFDI is an important factor determining domestic R-D performance, more so in the case of a joint-venture.
2. Wei and Ling (2008)	Addressed the case of China's OFDI between 1985 and 2004, and found that inverse technology spillovers exist when Multinational Enterprises (MNEs) invest abroad and transfer technology from overseas subsidiaries to parent companies at home.
3. Deng (2007)	Showed in the case of China that firms used their asset seeking FDI behaviour to obtain strategic resources that were available in more advanced foreign markets but limited in their own country and to access host countries' centres of innovation.
5. Branstetter (2006)	Used “patent citation data” to infer knowledge spillovers at the firm level. The results of the study indicated that with an increase in acquisitions in the United States by Japanese firms, the latter showed a greater tendency to cite US patents as “prior art” in their US patent application. Thus, OFDI was a channel for providing Japanese firms access to foreign technology networks.
6. Pradhan and Abraham (2005)	Indicated that one of the main motivations behind Indian firms' overseas acquisitions was to acquire firm specific intangibles such as technological skills. Because the Indian manufacturing sector is more research intensive and has greater absorptive capacity, it allowed them to integrate acquired foreign capital assets.
7. Makino, Lau and Yeh (2002)	Observed that firms invest more in capital intensive developed nations than in developing countries in order to fulfil their quest for strategic capabilities.
8. Van Pottelsberghe de la Potterie and Lichtenberg (2001)	Investigated econometrically the technology transfers through FDI, and pointed out that such transfers do take place if a country invests in R-D-intensive foreign nations. Countries such as Japan, Germany and France have benefitted from the US R-D capital stock through outward investment.

Source: Authors' compilation.

These studies thus point towards a positive correlation between OFDI and knowledge spillovers, and that the strategic assets acquiring motive could be realised through technology seeking outward investment.

Focusing, in particular, on the OFDI from emerging countries as the channel for R-D spillovers, diverse mechanisms drive such spillovers. One of these is “sharing of the R-D expenditure”; wherein the host and home country firms jointly undertake technological research. The second method is the “feedback mechanism”, in which foreign subsidiaries of the MNCs transfer knowledge to their home base. Then there is the “mechanism of reverse technology transfer”, i.e. acquiring knowledge through direct investment, particularly effective when firms carry out asset seeking FDI through mergers and acquisitions and joint ventures. As a result, firms obtain advanced technologies and enhance their core competitiveness. Finally, there is a fourth way where parent companies “outsource R-D activities” and relocate them overseas (Zhao and Liu, 2008).

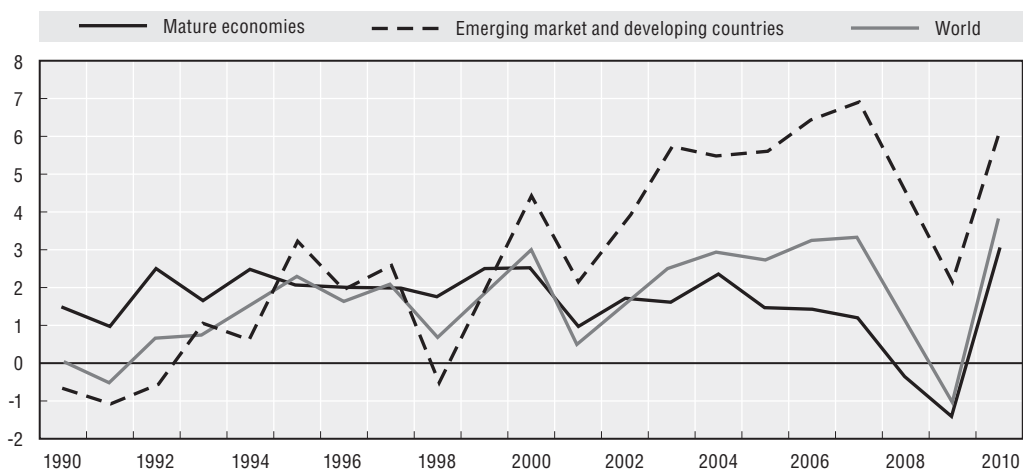
For example, looking at the Indian case, the main mechanism of acquiring technology has been to engage in strategic asset seeking investment, mostly directed towards developed nations, and enhance one's own technological innovation capacity. Pradhan and Singh (2009) explored the Indian firm's OFDI in the auto industry and suggested a significant impact of OFDI on auto firm's R-D intensity. Also, NIIT Ltd., a leading service sector Indian MNE, has been involved in strategic asset seeking investment through brown field investment. In 2002-03, it undertook several acquisitions such as Osprey Systems in March 2002 and Cognitive Arts in February 2003. With these acquisitions, NIIT gained strategic advantage in the US corporate knowledge solution market (Pradhan, 2006).

Given this background about R-D spillovers resulting from FDI, the following section highlights the main objective of our study – the “feedback effect” of such technology spillovers on the growth of domestic productivity of emerging economies.

3. The impact on total factor productivity growth

Recent trends show that the emerging and developing economies have been experiencing rapid TFP growth³ since the early 2000s. Compared with the advanced economies where the growth rate dropped from 0.4% per year between 1995 and 2005 to -0.1% between 2005 and 2008, the emerging economies have witnessed an improvement from 1.0% in 1995-2005 to 2.2% in 2005-2008 (the Conference Board, *Total Economy Database*). However, the level of productivity in developing countries is still much lower than advanced countries, indicating scope for strengthening and catching up. Looking at the labour productivity growth trend (Figure 3), the worldwide long-term trend is towards faster productivity growth with the emerging and developing economies leading the way. It was during this period that outward investment from developing countries rose, indicating a possibility of a positive effect of OFDI on productivity growth.

Figure 3. **Labour productivity growth trend – GDP per person (per cent)**



Source: The Conference Board, *Total Economy Database*, January 2014, www.conference-board.org/data/economydatabase/.

One of the sources of interdependence between OFDI and domestic productivity in the long run is the Strategic Asset Seeking motive (Herzer, 2011). Firms undertake asset seeking investment in order to gain access to the assets that are unavailable in the home country. Such a type of investment also includes the technology sourcing FDI, which transfers knowledge across borders and it is this transfer of technological know-how, attained by

foreign affiliates, to the parent companies, that facilitates the increase in productivity at home. Thus OFDI affects TFP growth through R-D spillovers, as productivity changes are critical signals of technology upgrading.

Moreover, it is important to note that OFDI not only affects the productivity of the investing firm but also of the economy as a whole (Blomstrom and Kokko, 1998). The investing firms could well be a source of advanced technology to local firms, which may benefit by copying or through labour turnover. Also, the benefit could be from the high quality intermediate goods produced by the investing firm that may now be available to the local firms at lower prices. Domestic firms may even take the advantage of economies of scale, as OFDI opens up the opportunity for the investing firm to grow larger than would have been possible with production in just one nation. In addition, it is obvious that the increased competition between the MNCs and the domestic firms would be an impetus for the latter to run more efficiently.

However, there are a few considerations related to the impact of knowledge spillovers resulting from OFDI. First of all, the learning ability of the investing firm matters. If the firms have low technological capacity, as is true for firms in developing countries, it may be difficult to effectively exploit the foreign technology. Furthermore, OFDI could reduce domestic capital accumulation and hence productivity when the investors invest the scarce domestic resources abroad. Finally, it may well be argued that significant spillovers to developing nations take place only if a substantial portion of investment goes to more developed countries, which have a higher technological expertise.

A number of studies provide empirical evidence for the possible impact of OFDI on domestic productivity. Kimura and Kiyota (2006) analysed panel data on Japanese firms for the period 1994-2000, and suggested that OFDI increases productivity. Their key finding was that, on average, firms that invest abroad have 1.8% higher growth than those that do not engage in FDI. Barba and Castellani (2004) considered a sample of Italian firms between 1993 and 1998, and reported a positive causal effect of outward investment on the home country's TFP growth, output and employment. They also compared the multinational firms that invested abroad with the domestic firms, and found that the former outperform.

Herzer (2011) examined the long-run relationship between OFDI and TFP for a sample of 33 developing countries over the period 1980-2005 and found that on average a positive effect is prevalent. However there is considerable heterogeneity explained by the cross-country differences, mainly due to labour market regulations. Bitzer and Gorg (2009), on the other hand, found an overall negative effect of OFDI on TFP, using panel data for 17 OECD countries over 1973-2001. But there were large country differences, for example, South Korea observed the largest negative effect, though countries such as France, the United Kingdom and the United States witnessed higher factor productivity due to OFDI.

Therefore, the impact of OFDI on domestic productivity could be mixed, differing significantly across countries. It may be particularly true that the effect does not necessarily depend on the motive of investment but on a number of other factors such as technological capacity, government policies, and level of financial development. More importantly, human capital stock is a crucial factor. Benefits from international R-D spillovers could be greater when the labour force in developing countries is more educated and skilled, termed as the "absorptive capacity" by Keller (1996). Also, how well the diffused technology is absorbed and implemented could be reflected in the increase in number of patent applications filed by the residents.

We now attempt to provide evidence of the effect of OFDI from EMNEs into the developed countries on the productivity growth of the home country through R-D spillovers using an empirical model. We also evaluate and compare this effect with that of spillovers resulting from IFDI from resource rich countries into the emerging economies – the traditional channel.

4. Empirical framework

In this study we assess the impact of FDI on TFP growth through technology flows across borders by employing a generalised version of the methodology suggested by Van Pottelsberghe de la Potterie and Lichtenberg (2001).⁴ Our paper contributes constructively to the literature by focusing on the impact of FDI generated R-D spillovers from developed to developing countries during recent years, i.e. technology transfers through North-South bilateral FDI. We consider panel data for 18 emerging economies over the 1990-2010 period, and investigate the impact of R-D spillovers resulting from FDI flows to and from the 34 OECD member countries as the investing partners (Appendix A and B, respectively, list the 18 emerging and the 34 OECD nations). We also extend the specification by including catalyst terms, such as human capital, as control variables.

4.1. Model specification

Equation [1] is the basic econometric model; it states that the domestic TFP growth of a country is a function of its domestic R-D capital stock and of foreign R-D capital stock:

$$\ln TFP_{it} = \alpha_i + \alpha^f \ln RD_{it}^f + \alpha^d \ln RD_{it}^d + \varepsilon_{it} \quad [1]$$

where $i = 1 \dots 18$ is a country index; $t = 1990 \dots 2010$ is a time index; $\ln TFP$ is the natural logarithm of TFP;⁵ RD^f represents the foreign R-D capital stock; RD^d represents the domestic R-D capital stock; α_i is a country-specific intercept; α^f is the elasticity with respect to the foreign R-D capital stock; α^d is the elasticity with respect to the domestic R-D capital stock; and ε is the error term.

The two different foreign R-D capital stock terms are constructed as follows. The first one, the foreign R-D capital stock embodied in OFDI from country i to country j , RD_{it}^{fo} , is given as:

$$RD_{it}^{fo} = \sum_{j=1}^{34} \frac{OFDI_{ijt}}{GDP_{jt}} RD_{jt}^d \quad [2]$$

here, $i = 1 \dots 18$ emerging economies and $j = 1 \dots 34$ OECD countries. $OFDI_{ijt}$ is the OFDI flow from country i into country j in year t , GDP_{jt} is the GDP of country j in year t , RD_{jt}^d is the domestic R-D capital stock of country j in year t , i.e. RD_{jt}^d / GDP_{jt} is the degree of R-D intensity of country j in year t . Thus, RD_{it}^{fo} is the foreign R-D spillover for each country i , given as the weighted average of R-D intensity of its investing partner j with OFDI flow from country i into country j .

The second one, the foreign R-D capital stock embodied in IFDI to country i from country j , RD_{it}^{fi} , is computed as:

$$RD_{it}^{fi} = \sum_{j=1}^{34} \frac{IFDI_{ijt}}{GDP_{jt}} RD_{jt}^d \quad [3]$$

here, $IFDI_{ijt}$ is the IFDI flow into country i from country j in year t , hence RD_{it}^{fi} is the foreign R-D spillover for each country i , given as the weighted average of R-D intensity of its investing partner j with IFDI flow into country i from country j .

A further specification, Equation [4], includes “Human Capital” and “Number of Patent Applications filed by Residents” as two catalyst terms included with the purpose of getting better estimates of output elasticities of domestic and foreign R-D capital terms. These two variables are not only important channels of productivity growth, but also capture the level of indigenous capacity of the home country. We therefore include these control variables in order to avoid any bias in the estimation.

$$\ln TFP_{it} = \alpha_i + \alpha^f \ln RD_{it}^f + \alpha^d \ln RD_{it}^d + \alpha_1 \ln PA_{it} + \alpha_2 \ln H_{it} + \varepsilon_{it} \quad [4]$$

where PA_{it} = number of patent applications filed by residents in country i in year t , and H_{it} = human capital, proxied as “Average years of total schooling (age 25+)” in year t .

Thus, the two main hypotheses that would be examined are:

1. Whether the growth of an economy’s productivity is affected by foreign R-D efforts when it undertakes OFDI in technologically-advanced countries. This can then be compared with the impact of spillovers generated when R-D-intensive countries invest in it.
2. Whether the inclusion of the economy’s indigenous technological capability measured in terms of skilled workforce and level of innovation, the latter reflected by number of patent applications, results in improved estimation of the effect of foreign R-D spillovers.

4.2. Data Sources

The data source for OFDI and IFDI flows⁶ to and from, respectively, the 34 OECD countries under study between 1990 and 2010 is the *OECD International Direct Investment Database*. Data on GDP and number of patent applications filed by residents⁷ is drawn from the World Bank data series (www.databank.worldbank.org) and data on R-D intensity is drawn from the *OECD Factbook 2010*. We have collected data on TFP from the Conference Board: *Total Economy Database*. For the average years of total schooling (age 25+), we interpolate from the five yearly data given by Barro and Lee, extracted from the education statistics of the World Bank data series (www.databank.worldbank.org).

Before we discuss the results, an important issue must be addressed. As we are looking at the overall macroeconomic picture, it is necessary to show a match-up between the FDI flows into major sectors and the R-D intensity in the respective sectors in order to justify the R-D spillovers approach. This is important to rule out the criticism that FDI could possibly be going into those sectors where there is minimal R-D expenditure, and also to support the rationale behind creating the R-D spillover term and the motivation for developing countries investing to acquire strategic assets. Such accordance is essential, more in the case of OFDI from developing economies into the R-D intensive countries, aimed at attaining technological knowledge, than in the case of IFDI from technology intensive DMNEs. This is so because DMNEs are themselves advanced, and when they invest in emerging economies to enhance their own technological base, they also diffuse knowledge to the local EMNEs in the process.

Therefore, to substantiate this argument, we take the example of OFDI flows from eight major emerging economies under study, and look at the major destination/host countries where they invest. We further classify the investment on the basis of major sectors, and then focus on how much R-D expenditure the host countries undertake in those specific sectors (Table 3).

Looking, for example, at the case of China (Table 3), its major destinations are the United States, Japan, Germany and Australia and the major sectors are Trade and Services,

Table 3. **Matching sectoral distribution of OFDI with the sector specific R-D expenditure in the major destinations**

Investing/ home country	Major sectors	Major destinations (host country)	R-D expenditure in major sectors as a % of total sectoral R-D expenditure in the host country	R-D expenditure in the major sectors as a % of GDP in the host country
India	Pharmaceuticals, agricultural inputs, software, IT and broadcasting	US, Russia, Sri Lanka, Southeast Asia, UK	US: 19.97 UK: 29.24	US: 0.36 (1.89) UK: 0.31 (1.13)
China	Trade and services, manufacturing, resource extraction (oil, gas and minerals), IT	Hong Kong, US, Japan, Australia, Germany	US: 75.28 Japan: 90.31 Germany: 92.82 Australia: 61.55	US: 1.36 (1.89) Japan: 1.55 (1.73) Germany: 1.28 (1.53) Australia: 0.51 (0.93)
Brazil	Energy, mining, services	US, UK, Portugal, Netherlands	US: 29.04 UK: 19.24 Portugal: 41.83 Netherlands: 17.54	US: 0.52 (1.89) UK: 0.20 (1.13) Portugal: 0.17 (0.35) Netherlands: 0.15 (0.98)
Russia	Resource extraction (oil, gas and metal), manufacturing, telecommunication	European Union, US, CEE	US: 71.40 Germany: 92.65 Hungary: 77.54 Netherlands: 81.77 Portugal: 56.92	US: 1.28 (1.89) Germany: 1.27 (1.52) Hungary: 0.46 (0.60) Netherlands: 0.68 (0.98) Portugal: 0.18 (0.35)
Argentina	Oil and gas, iron and steel, food, pharmaceuticals, telecommunications	US, Europe, Japan	US: 11.77 Japan: 11.97	US: 0.21 (1.89) Japan: 0.21 (1.73)
Malaysia	Oil and gas, finance, real estate, construction, trade, hotels and restaurants	US, Singapore, ASEAN	US: 13.12	US: 0.24 (1.89)
Singapore	Finance, transport, manufacturing, real estate, construction	UK, Netherlands, Germany, Australia	UK: 94.02 Netherlands: 91.07 Germany: 98.34 Australia: 76.33	UK: 1.01 (1.13) Netherlands: 0.76 (0.98) Germany: 1.36 (1.53) Australia: 0.64 (0.93)
South Africa	Mining and natural resources, trade, finance, business activities, wood and wood products, machinery	US, UK, Netherlands, Germany, Australia	US: 9.85 UK: 8.99 Netherlands: 16.75 Germany: 11.77 Australia: 30.43	US: 0.18 (1.89) UK: 0.09 (1.13) Netherlands: 0.14 (0.98) Germany: 0.16 (1.53) Australia: 0.27 (0.93)

Notes: CEE = Central and Eastern Europe, ASEAN = Association of Southeast Asian Nations. Parentheses in column 5 = Total sectoral R-D expenditure/GDP. All figures in columns 4 and 5 are averaged across the 1990-2010 period.

Source: OECD STAN Database, World Bank Data Series, UNCTAD 2007.

Manufacturing, and Resource Extraction. Calculating the percentage of R-D expenditure in these sectors, the United States spends 75.28% of its total R-D expenditure in these three sectors only. Japan spends 90.31%, Germany spends 92.82%, and Australia spends 61.55%. The last column expresses R-D expenditure in these three major sectors as a percentage of GDP of host countries, i.e. the R-D intensity of those sectors. These figures clearly demonstrate that FDI from developing and emerging economies does take place in those sectors where the host countries are primarily undertaking R-D, and justify the asset seeking behaviour of emerging economies. In other words, host country sectors that have higher R-D intensity attract significant levels of FDI from home countries.

5. Results

We use a Fixed-Effects (FE) regression⁸ to analyse the impact of variables that vary over time. We also estimate the model based on a change specification (1st difference), in order to take into account trend and non-stationarity. Further, to check whether there is a long-run cointegration between the variables under study and that the estimated regression is non-spurious, we carry out Pedroni's (1999) cointegration tests for panel data. The seven tests, as suggested by Pedroni, are applied to the basic specification in level and in 1st difference.

Table 4a and 4b summarise the estimation and the cointegration test results, respectively, for our main specification (equation [1]). The 1st two columns of Table 4a present the “within” estimates and the last two columns are the “1st difference” estimates. Regressions (i) and (ii) show the estimated output elasticities of domestic R-D, and of the foreign R-D capital stock incorporated alternatively into one of the two types of technology transfer channels. In both cases, the estimated elasticity of domestic R-D capital stock is positive and significant. Regression (i) concentrates on the foreign R-D capital stock embodied in OFDI, and the estimated elasticity is positive and significant (3.24). This result supports our main hypothesis that OFDI allows EMNEs to source foreign technological bases, thereby increasing the growth of TFP in their home country.

Table 4a. **Total factor productivity estimation results – outward FDI and inward FDI 1990-2010, 378 observations**

	Within		First difference	
	(i)	(ii)	(iii)	(iv)
Domestic R-D				
$LnRD^d$	15.2089* (3.57)	14.3067* (4.53)	5.1028 (1.08)	7.4502* (2.10)
Foreign R-D				
$LnRD^{fo}$ (outward FDI)	3.2429* (6.83)	-	0.3868 (1.29)	-
$LnRD^{fi}$ (inward FDI)	-	6.1523* (12.36)	-	1.0649* (2.65)

Table 4b. **Pedroni's cointegration test results [equations (i) to (iv)]**

Statistic	(i)	(ii)	(iii)	(iv)
Panel v-Statistic	-2.3776*	-0.5152	-2.5095*	-2.1592*
Panel rho-Statistic	6.1019*	6.0915*	4.2799*	5.7028*
Panel PP-Statistic	3.4670*	1.8119**	-2.6126*	-1.8483**
Panel ADF-Statistic	5.4676*	1.1434	3.1352*	0.4989
Group rho-Statistic	7.2695*	7.3411*	5.2867*	6.9358*
Group PP-Statistic	2.3469*	-0.8474	-2.9220*	-5.2221*
Group ADF-Statistic	4.0912*	-2.5403*	4.2913*	-29.9276*

Notes: Figures in parentheses give t-statistics. * denotes significance at the 5% level, ** denotes significance at the 10% level. RD^d = Domestic R-D capital, RD^{fo} = Foreign R-D capital embodied in OFDI, RD^{fi} = Foreign R-D capital embodied in IFDI. The cointegration tests reject the null hypothesis of no cointegration.

Comparing the impact of technology spillovers embodied in OFDI with that of spillovers embodied in IFDI, regression (ii) shows that the magnitude of impact is much larger in the latter case (6.15). This suggests that IFDI from OECD countries into the

18 emerging economies induces more substantial technology transfers compared with OFDI from those 18 emerging economies into OECD countries. One possible explanation for such diversity could be that when MNEs from R-D intensive countries invest in developing countries, the former aim at exploiting their own technological innovations and in this process they diffuse the technological knowledge to the host country firms either through “copying” or “labour turnover”. Also, when developed countries undertake offshore production to take advantages of factors such as lower wages in developing countries, the host countries benefit more from externalities emanating from foreign companies. However, when firms from emerging countries invest in R-D intensive developed countries, their low technological capacity may not allow them to fully exploit the knowledge spillovers, hence resulting in a comparatively lower impact on the growth of TFP. But nonetheless, OFDI from EMNEs has a positive effect, thus supporting the premise that if foreign companies intend to copy domestic knowledge in the host countries, their home country is likely to benefit from potential spillovers (Van Pottelsberghe de la Potterie and Lichtenberg, 2001).

Looking at the 1st difference estimates, in equations (iii) and (iv), they broadly confirm the level estimates in terms of the sign of coefficients, though the magnitudes are much lower. Also, the elasticities are now significant only in the case of regression equation (iv) that takes into account the R-D spillovers embodied in IFDI. Almost all the cointegration tests suggest that each of the regression equation is cointegrated in the long run (Table 4b).

Table 5 investigates the time stability of the estimated elasticities. Regressions (i) and (ii) are the same as the within estimates of Table 4a, giving the elasticities for the whole period 1990-2010. Regressions (iii) and (iv) present the results for the 1990-2000 period, and regressions (v) and (vi) give the estimates for the 2001-10 period. Clearly, the impact of domestic R-D capital is driven by the 2001-10 trend, as the coefficient for the 1990-2000 series is negative and insignificant. In the case of the impact of foreign R-D capital embodied both in OFDI and IFDI, the emerging economies benefitted significantly only during 2001-10. Though the impact was positive in the 1990s, the effect was insignificant. These figures point out that the practice of technology transfers intensified only during the 2000s, also the phase when FDI flows to and from the EMNEs saw an increasing trend.

Table 5. Time stability of the output elasticities with respect to domestic and foreign R-D capital

	1990-2010		1990-2000		2001-2010	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Domestic R-D						
$LnRD^d$	15.2089*	14.3067*	-4.2690	-2.8086	16.4131*	18.4966*
	(3.57)	(4.53)	(-0.79)	(-0.71)	(2.26)	(3.50)
Foreign R-D						
$LnRD^{fo}$ (outward FDI)	3.2429*	-	0.3948	-	2.1016*	-
	(6.83)		(0.66)		(3.39)	
$LnRD^{fi}$ (inward FDI)	-	6.1523*	-	0.8556	-	4.1619*
		(12.36)		(1.05)		(6.88)

Notes: Figures in parentheses give t-statistics. * denotes significance at the 5% level. RD^d = Domestic R-D capital, RD^{fo} = Foreign R-D capital embodied in OFDI, RD^{fi} = Foreign R-D capital embodied in IFDI.

Table 6a and 6b focus on the estimation that includes the number of patent applications filed by residents and human capital as control variables. Looking at

Table 6a. **Estimation results – including control variables**

	Within		First difference	
	(i)	(ii)	(iii)	(iv)
Domestic R-D				
$LnRD^d$	7.8671 (1.62)	11.2084* (2.97)	7.1945 (1.37)	8.7365* (2.24)
Foreign R-D				
$LnRD^{fo}$ (outward FDI)	2.2107* (3.88)	-	0.3054 (0.97)	-
$LnRD^{fi}$ (inward FDI)	-	5.5842* (9.47)	-	1.0354* (2.50)
Control variables				
$LnPA$	-2.3279 (-1.26)	-1.6351 (-1.07)	1.2986 (1.10)	1.2099 (1.12)
LnH	56.4557* (3.63)	27.0407* (2.15)	-127.5859 (-1.43)	30.8099 (0.45)

Table 6b. **Pedroni's cointegration test results [equations (i) to (iv)]**

Statistic	(i)	(ii)	(iii)	(iv)
Panel v-Statistic	-2.0030**	-0.3709	-1.4516	-2.8329*
Panel rho-Statistic	5.9899*	7.0537*	3.4346*	6.9614*
Panel PP-Statistic	0.4269	3.2941*	-0.6211	-0.9215
Panel ADF-Statistic	6.1345*	4.8593*	2.1958*	1.9081**
Group rho-Statistic	7.1583*	8.4405*	4.1801*	8.3704*
Group PP-Statistic	-5.4812*	-5.0593*	-1.8470**	-5.7148*
Group ADF-Statistic	5.5249*	-0.8882	2.9465*	-0.9036

Notes: Figures in parentheses give t-statistics. * denotes significance at the 5% level; ** denotes significance at the 10% level. RD^d = Domestic R-D capital, RD^{fo} = Foreign R-D capital embodied in OFDI, RD^{fi} = Foreign R-D capital embodied in IFDI, PA = Number of patent applications filed by residents, H = Human capital defined as the average years of total schooling (age 25+). The cointegration tests reject the null hypothesis of no cointegration.

equations (i) and (ii) in Table 6a, the impact of domestic R-D capital is now significant only in the case of equation (ii). The elasticities of RD^{fo} (2.21) and RD^{fi} (5.58) are positive and significant, though the magnitudes are lower compared with the results obtained in Table 4a. Looking at the coefficients of control variables, the number of patent applications filed by residents does not have a significant impact on growth of TFP⁹. However, the human capital term has a substantial positive and significant effect. The latter result implies that skilled and educated workforce is an essential factor contributing towards improvement in efficiency, and not properly taking into account such factors leads to an upward bias in the output elasticity of foreign R-D capital stock. Again, the cointegration tests suggest the presence of a long-run relationship (Table 6b).

Further, we carry out an additional analysis to exploit the heterogeneity within the data set. We test whether the impact of FDI embodied spillovers is greater when emerging economies deal with larger investing partner countries than smaller. To do this, we introduce a dummy for the larger industrialised countries (G-7) in our sample of 34 OECD countries in order to investigate the difference in the estimated parameters. The results of this analysis are included as Appendix C.

We also compute two matrices of bilateral elasticities of growth of TFP with respect to foreign R-D capital terms, using the estimated parameters from regressions (i) and (ii) of

Table 4a. We follow a similar methodology as suggested by Van Pottelsberghe de la Potterie and Lichtenberg (2001) to compare our results. In case of OFDI embodied spillovers, the elasticity of country i 's TFP growth with respect to country j 's domestic R-D capital may be expressed as follows:

$$\alpha_{ij}^{fo} = \frac{\delta \text{LnTFP}_i}{\delta \text{LnRD}_j^d} \quad (i = 1 \dots 18 \text{ emerging economies}, j = 1 \dots 34 \text{ OECD nations})$$

$$= \frac{\delta \text{LnTFP}_i}{\delta \text{LnRD}_i^{fo}} * \frac{\delta \text{LnRD}_i^{fo}}{\delta \text{LnRD}_j^d} = \alpha^{fo} * \frac{\delta \text{LnRD}_i^{fo}}{\delta \text{LnRD}_j^d} \quad \left(\text{LnTFP}_{it} = \alpha_i + \alpha^d \text{LnRD}_{it}^d + \alpha^{fo} \text{LnRD}_{it}^{fo} + \varepsilon_{it} \right) \quad [5]$$

$$= \alpha^{fo} * \frac{\delta \text{RD}_i^{fo}}{\delta \text{RD}_j^d} * \frac{\text{RD}_j^d}{\text{RD}_i^{fo}}$$

$$\text{now: } \text{RD}_i^{fo} = \sum \frac{\text{OFDI}_{ij} * \text{RD}_j^d}{\text{GDP}_j} \dots \quad (\text{definition of } \text{RD}_i^{fo}) \quad [6]$$

$$\text{therefore: } \frac{\delta \text{RD}_i^{fo}}{\delta \text{RD}_j^d} = \frac{\text{OFDI}_{ij}}{\text{GDP}_j} \quad [7]$$

$$\text{hence: } \alpha_{ij}^{fo} = \alpha^{fo} * \frac{\text{OFDI}_{ij}}{\text{GDP}_j} * \frac{\text{RD}_j^d}{\text{RD}_i^{fo}} \quad [8]$$

implying that elasticity of country i 's TFP growth with respect to country j 's domestic R-D capital is an increasing function of OFDI from country i to country j , and of country j 's R-D intensity. Similarly, the elasticity of country i 's TFP growth with respect to country j 's domestic R-D capital, embodied in IFDI (country j investing in country i) may be obtained.

The computed bilateral elasticities for the two channels of R-D spillovers – OFDI and IFDI – are given in Table 7a and 7b, respectively. The figures indicate, for example, that a 1% increase in US R-D intensity raises the TFP of India by 1.31% through OFDI flows from India to the United States, and by 1.99% through IFDI flows into India from the United States. This implies that, in case of bilateral FDI flows between India and the United States, technology transfers are more intense when the United States undertakes investment in India than when India undertakes FDI in the United States. Of the 18 emerging economies under study, the TFP growth elasticity for 13 nations is higher when the United States invests in them than when they invest in the United States. In other words, the United States is an important R-D spillover generator when it undertakes FDI in developing and emerging countries.¹⁰

Comparing a subset of bilateral elasticities between BRICS and their six main OECD destination countries, namely the United States, the United Kingdom, Japan, Germany, Portugal and the Netherlands, it is clear that the major emerging economies benefit more when the R-D intensive developed countries invest in them than the other way round. For Brazil and China, their bilateral elasticities with all the six OECD investing partners are higher when spillovers are generated through IFDI. Russia, India and South Africa also attain greater productivity growth by acting as host nations, with much larger IFDI embodied elasticities with respect to three, four and five investing partners, respectively.

Overall the elasticities between all the 18 emerging economies and their 34 OECD investing partners, we observe that in the case of 22 OECD countries the TFP growth elasticity of foreign R-D capital embodied in IFDI flows from those 22 nations into the 18 emerging economies is higher than the elasticity embodied in OFDI flows in the reverse direction. Thus, the impact of technology spillovers from R-D intensive foreign countries is

Table 7a. Output elasticities of foreign R-D capital embodied in OFDI flows

Emerging → OECD ↓	India	China	Brazil	Russia	Argentina	Malaysia	Singapore	Thailand	Bulgaria	Indonesia	Latvia	Lithuania	Peru	Philippines	Romania	South Africa	Ukraine	Venezuela
Australia	-0.0011	0.0293	0.0000	0.0090	0.0000	0.5038	0.1317	0.0138	0.0000	0.1161	0.0000	0.0000	0.0000	0.7816	0.0000	0.1123	0.0000	0.0000
Austria	0.0066	0.0599	0.1637	0.3985	1.9791	0.0318	-0.0443	-0.8845	-1.2746	-0.0097	-2.0409	0.0638	0.0000	0.0042	-0.2031	0.2323	1.0121	-0.0048
Belgium	0.0388	-0.0446	1.1916	-0.0083	-8.1135	-0.2487	0.4845	0.1584	0.7081	-0.7862	0.5497	-0.0392	0.0164	-0.0930	-0.0807	0.0859	-0.0484	-0.0029
Canada	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Chile	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Czech Republic	0.0154	0.0167	0.0031	0.0360	-0.0021	0.0051	0.0012	0.0111	-0.0129	0.0031	0.0086	0.0550	0.0000	0.0000	-0.1252	-0.0002	0.3434	0.0000
Denmark	0.0180	0.0408	-0.0663	0.1658	-38.1271	0.0424	-0.0191	0.2360	0.1188	0.0277	4.5718	1.8405	0.0000	0.0120	0.0136	0.0036	-1.4514	-0.0019
Estonia	-0.0006	0.0010	-0.0004	0.1075	0.0033	-0.0001	0.0005	-0.0006	0.0399	-0.0035	-2.0205	1.9169	0.0000	-0.0001	-0.0019	-0.0001	0.2228	0.0001
Finland	-0.0127	0.0240	-0.0048	0.5325	-0.0821	0.0037	-0.0015	0.0112	0.0000	0.0102	-1.3337	-0.2951	-0.0254	-0.0025	0.0042	-0.0168	-0.7074	0.0000
France	0.0687	0.1185	0.0568	0.7654	10.9926	0.0586	0.1059	0.2216	-0.1379	1.7178	4.9012	0.6668	0.0210	0.1331	2.0648	0.0149	0.2306	-0.1253
Germany	0.1029	0.4168	0.0077	1.1571	-9.2466	0.2435	0.0403	0.2400	-0.4196	-0.5752	-0.4358	0.1452	-0.0239	0.0170	0.6989	0.7265	-0.3638	-0.7418
Greece	0.0000	-0.0001	0.0000	0.0033	-0.0035	0.0000	0.0000	-0.0002	-0.0236	0.0006	0.0305	-0.0769	-0.0002	0.0000	0.0057	0.0000	-0.0052	0.0000
Hungary	-0.0009	-0.0045	0.2589	0.1138	0.0789	0.0108	0.0062	-0.0146	-0.1893	-0.0174	-0.0618	0.0464	-0.0051	0.0054	0.0752	0.0082	-0.7620	-0.0002
Iceland	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6937	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ireland	-0.0076	-0.2405	0.0513	0.1159	2.8945	0.0082	0.0610	-0.2055	0.0120	-0.0256	0.2584	-0.0654	0.0541	-0.4176	-0.0327	-0.0274	0.0197	-0.0167
Israel	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Italy	0.0134	0.0389	0.0931	0.4357	9.2928	-0.0002	0.0022	0.2784	1.4987	0.0486	0.8394	0.0714	-0.0320	0.0056	0.4091	0.0057	0.0997	-0.0386
Japan	0.0081	0.0386	0.0009	0.0016	6.7840	0.0059	0.6125	-0.0235	0.0000	0.1735	0.0000	0.0000	0.0000	0.0163	0.0000	0.0000	0.0000	0.0000
Korea	0.0251	0.3298	0.0022	0.0078	0.0754	1.9055	0.1603	0.0740	-0.0017	0.0411	0.0000	0.0000	-0.0047	0.0208	0.0004	0.0013	0.0020	-0.0015
Luxembourg	-0.0088	0.4629	0.4345	-2.5374	24.9766	0.0094	0.0161	0.0950	3.3539	-0.0480	-3.3822	0.2561	-0.1268	2.6714	0.8055	0.6815	2.4995	0.0001
Mexico	0.0185	0.0049	0.0102	0.0001	6.7485	0.0007	0.0076	0.0160	-0.0004	0.0008	0.0000	0.0043	-0.0225	0.0008	0.0001	0.0001	0.0004	-0.0060
Netherlands	0.1825	0.0123	0.1799	0.1243	4.6648	0.1383	0.0970	3.1122	-0.5115	1.5740	0.7021	0.3655	-0.0744	0.1817	0.0187	0.0109	0.1341	-0.0309
New Zealand	0.0000	-0.0053	0.0000	0.0000	-0.0119	0.0194	0.0088	-0.0085	0.0000	0.2490	0.0000	0.0000	0.0000	0.0004	0.0000	-0.0003	0.0000	0.0000
Norway	0.0014	1.7128	0.0124	0.0008	-0.0890	0.0161	-0.2634	-1.2396	0.0000	0.0020	0.0252	0.0481	0.0935	0.0000	-0.0378	0.0114	0.0821	-0.0049
Poland	0.0031	0.0066	-0.0075	-0.1495	0.0048	0.0108	0.0024	0.0176	0.0212	-0.0258	1.0210	0.2537	0.0000	0.0060	-0.0562	0.0000	1.7002	-0.0001
Portugal	0.0038	-0.0018	0.0748	0.0058	0.0189	0.0000	0.0000	0.0000	0.0001	0.0000	0.0181	0.0889	0.0000	0.0000	-0.0068	0.0044	0.0058	-0.0098
Slovak Republic	0.0000	-0.0006	0.0088	-0.0072	1.6857	0.0023	0.0002	0.3329	-0.0032	-0.0002	0.0160	0.1483	0.0000	0.0000	-0.0077	0.0001	-0.0015	0.0000
Slovenia	0.0000	0.0000	-0.0001	0.0017	-0.0157	0.0000	0.0001	-0.0002	0.0323	0.0001	0.0418	0.0109	0.0000	0.0000	-0.0016	0.0000	-0.0189	0.0000
Spain	0.0007	0.0023	0.1816	0.4616	-36.0857	0.0176	0.0025	0.0041	0.0344	0.0072	0.0000	0.0000	-0.5373	0.0032	0.0060	0.0005	0.1858	-0.2346
Sweden	-0.0352	-0.0784	-0.0116	-0.3946	-0.1504	-0.0219	0.0250	-0.4524	0.0861	-0.1256	-1.5478	-2.5290	0.0306	-0.0106	-0.1379	0.2624	-0.1046	0.0126
Switzerland	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Turkey	0.0089	0.0003	0.0392	0.2412	0.0000	0.0060	0.0013	0.0433	-0.0629	0.0000	0.4411	0.2348	0.0000	-0.0006	0.0139	0.0002	0.2161	0.0002
United Kingdom	1.4836	0.0373	0.0029	0.0487	0.0905	0.0912	0.1176	0.1028	-0.0143	0.0200	0.0000	0.0000	0.0000	0.0029	0.0046	1.0925	0.0000	0.0000
United States	1.3072	0.2622	0.5568	1.6029	24.8770	0.3801	1.6836	1.1010	-0.0133	0.8655	-0.0558	0.0290	3.8768	-0.0979	-0.1890	0.0301	-0.0510	4.4471

Note: Estimated TFP growth elasticity in the country column with respect to the R-D capital in the row country, based on regression equation (i) of Table 4a.

Table 7b. **Output elasticities of foreign R-D capital embodied in IFDI flows**

Emerging → OECD ↓	India	China	Brazil	Russia	Argentina	Malaysia	Singapore	Thailand	Bulgaria	Indonesia	Latvia	Lithuania	Peru	Philippines	Romania	South Africa	Ukraine	Venezuela
Australia	-0.0042	-0.0032	0.0004	0.0003	0.0045	0.0001	0.0597	0.0074	0.0000	0.0004	0.0000	0.0000	0.0000	0.0196	0.0000	0.0004	0.0000	0.0000
Austria	0.0352	0.0866	0.0223	0.6556	0.0044	0.0138	0.0345	0.0125	2.8723	0.0128	0.5706	0.0697	0.0000	0.0007	2.1403	-0.0459	0.5067	0.0029
Belgium	0.0355	-0.0040	0.1669	0.1788	-0.0896	0.0180	0.5933	0.0355	0.2007	0.0099	0.0443	0.0241	0.1293	-0.0588	0.9163	0.0558	0.0391	0.0326
Canada	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Chile	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Czech Republic	0.0024	-0.0001	0.0000	0.0072	0.0000	0.0000	0.0001	0.0000	0.1634	0.0000	0.0056	0.0005	-0.0003	0.0000	0.1058	0.0000	0.0048	0.0000
Denmark	0.0464	0.0739	0.0408	0.0006	0.0568	0.0319	0.1181	0.0060	0.0456	0.0142	0.5254	0.3432	0.0075	-0.0052	0.0188	0.0592	-0.0221	0.0557
Estonia	0.0000	0.0000	0.0000	0.0026	0.0000	0.0000	0.0000	0.0002	0.0163	0.0000	2.5967	1.1439	0.0000	0.0000	0.0020	0.0001	0.0243	0.0000
Finland	0.2260	0.2497	0.0189	0.2782	0.0029	0.0237	0.0937	-0.0181	0.0061	0.0046	-0.9229	0.0860	0.0121	0.0024	0.0569	0.0156	0.0388	0.0115
France	0.3608	0.2695	0.7758	0.5749	0.5537	0.0697	0.2585	0.2098	0.3263	0.2306	0.0991	0.5222	0.5689	0.2206	0.5789	0.2286	0.2083	0.5892
Germany	0.9096	0.8207	0.3743	1.0354	0.2640	0.3104	0.1066	0.2855	0.8692	0.2985	0.9294	-0.1046	0.3147	0.1689	0.9396	0.5017	3.5681	0.6030
Greece	0.0000	0.0000	0.0001	0.0011	-0.0001	0.0000	0.0000	0.0000	0.0564	0.0000	-0.0010	-0.0001	0.0006	0.0000	0.0229	0.0001	0.0098	0.0000
Hungary	0.0013	-0.0001	0.0029	0.0214	0.0004	0.0022	0.0006	0.0000	0.2289	0.0000	0.0000	0.0085	-0.0002	0.0000	0.0345	0.0000	0.1734	0.0000
Iceland	0.0121	0.0010	0.0000	0.0099	0.0000	0.0000	0.0000	0.0001	0.0219	0.0005	0.4858	0.0553	0.0000	-0.0022	-0.0008	0.0001	0.0017	0.0000
Ireland	0.0130	0.0028	-0.0001	0.0019	0.0005	0.0001	0.0005	-0.0001	0.0000	0.0000	0.0000	0.0008	0.0000	0.0000	0.0000	0.0007	-0.0004	0.0000
Israel	0.0000	0.0029	0.0018	0.0000	0.0015	0.0000	-0.0140	0.0009	0.0000	0.0000	0.0000	0.0000	0.0113	0.0000	0.0063	0.0000	0.0000	0.0000
Italy	0.1126	0.0414	0.0959	0.0872	0.1010	0.0152	0.0043	0.0025	-0.0314	0.0074	0.0270	0.0680	0.0257	0.0188	0.2428	0.0337	0.1344	0.0682
Japan	0.7417	1.9582	0.6241	0.0463	0.1576	2.2146	0.7625	3.2390	0.0505	3.1225	0.0000	0.0000	0.4004	3.7303	0.0137	0.2253	0.0128	0.1095
Korea	0.1154	0.5644	0.0259	0.0345	0.0126	0.0684	0.0709	0.0904	0.0127	0.1429	0.0125	0.0003	0.5283	0.2118	0.0201	0.0136	0.0755	0.0103
Luxembourg	0.0080	0.0113	0.1203	0.2149	-0.0220	0.0034	0.2829	-0.0003	0.1016	0.0311	0.7264	0.0889	0.0106	0.0243	0.0076	0.2796	-0.0146	-0.0026
Mexico	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Netherlands	0.2368	0.1122	0.4195	0.7458	0.1289	0.2020	0.4169	0.3343	0.0547	0.0416	0.0778	0.0680	-0.1857	0.7517	0.4533	0.1377	0.1539	0.3661
New Zealand	0.0001	0.0000	0.0000	0.0000	-0.0003	0.0007	0.0060	0.0001	0.0000	0.0025	0.0000	0.0000	-0.0025	-0.0001	0.0000	0.0000	0.0000	-0.0016
Norway	0.0046	-0.0007	0.0161	0.1311	-0.0001	0.1078	0.3368	-0.0038	0.0165	-0.0002	0.1079	-0.0361	-0.0788	-0.0224	-0.0074	-0.0003	0.1096	-0.2046
Poland	0.0042	0.0003	0.0000	0.0128	0.0000	-0.0009	0.0009	0.0001	0.0095	0.0004	0.0076	0.8794	0.0000	0.0000	0.0134	-0.0001	0.0806	0.0000
Portugal	0.0013	0.0001	0.1152	0.0004	0.0027	0.0000	0.0000	0.0000	0.0011	0.0000	0.0015	0.0000	0.0000	0.0000	0.0226	0.0156	0.0110	-0.0061
Slovak Republic	0.0000	0.0000	0.0000	0.0006	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0003	0.0000	0.0000	0.0001	0.0000	0.0051	0.0000
Slovenia	0.0000	0.0002	0.0001	0.0239	0.0006	-0.0001	0.0000	0.0000	0.0180	0.0003	-0.0002	0.0000	-0.0001	0.0000	0.0002	0.0012	-0.0018	0.0000
Spain	0.0483	0.0258	0.7760	0.0409	1.9446	0.0113	0.0003	0.0120	0.1987	-0.0005	0.0000	0.0000	1.1386	-0.0191	0.1860	0.0229	0.0028	-0.0344
Sweden	0.2190	0.0705	0.0086	0.5755	-0.0038	0.2623	-0.0218	0.0150	0.6607	0.0059	1.0555	2.8989	0.2934	0.0062	0.1089	0.1366	0.4204	0.1050
Switzerland	0.0554	0.0205	0.0202	0.0217	0.0177	0.0276	0.1578	0.0407	0.0063	0.0179	0.0000	0.0000	0.1821	0.0085	0.0358	0.0031	0.0573	0.0644
Turkey	0.0017	0.0001	0.0001	0.0055	0.0001	0.0000	0.0001	0.0000	0.0057	0.0001	0.0020	0.0000	0.0000	0.0000	0.0052	0.0001	0.0059	0.0000
United Kingdom	0.9728	0.2739	0.4276	0.7111	0.5062	1.0309	0.1064	0.4673	0.0517	0.3307	0.2451	0.0011	0.6274	0.7393	0.0983	3.6281	0.2170	0.5632
United States	1.9900	1.5719	2.0963	0.7298	2.5052	1.7371	2.7743	1.4130	0.1863	1.8760	-0.4461	0.0315	2.1667	0.3546	0.1280	0.8365	0.3274	3.8178

Note: Estimated TFP growth elasticity in the country column with respect to the R-D capital in the row country, based on regression equation (ii) of Table 4a.

much greater when the developing and emerging economies act as host countries for the IFDI flows than as home countries undertaking OFDI. These country specific figures are consistent with our overall result given in Table 4a, showing a higher impact of RD^{fi} compared with RD^{fo} .

6. Conclusion and policy implications

An important dimension of the “going-out” strategy is provided by reverse technology spillovers through various mechanisms, one of which is the “strategic asset seeking” motive of OFDI. In this study, we examine the impact of such technology spillovers on the growth of TFP of the home country and, in contrast with some of the previous studies, the paper deals with the overall macro-effect of OFDI. We then compare the effect of spillovers embodied in OFDI with that of spillovers embodied in IFDI, and also analyse the role of human capital and the number of patent applications filed by residents as control variables to improve the estimation of output elasticity. To carry out the analysis, we apply the generalised version of Van Pottelsberghe de la Potterie and Lichtenberg Model (2001). We create the foreign R-D spillover variable based on outward and inward investment, focusing on the OFDI flows from 18 emerging economies into 34 OECD countries over the 1990-2010 period, along with the IFDI flows into the 18 emerging economies from the 34 OECD countries.

The results confirm that OFDI from the emerging economies into OECD countries has a positive and significant impact on the TFP growth of the former. The effect here is mitigated through a transfer of foreign R-D capital, but the magnitude of the impact is smaller than that of R-D spillovers resulting from IFDI flows in the opposite direction. This result is also confirmed by our country-specific bilateral elasticities, showing that developing countries benefit more when technology-rich nations invest in them than the other way round. However, our results are contrary to what Van Pottelsberghe de la Potterie and Lichtenberg (2001) suggested.¹¹ They found that IFDI is more intended to take advantage of the technological base of the home country than to diffuse the technology that originates in the home country. Thus, there is an insignificant impact on the growth rate of host countries’ productivity. Such a contrast might occur because of the difference in the coverage of countries. Van Pottelsberghe de la Potterie and Lichtenberg focus on FDI flows between 13 industrialised nations, whereas our study concentrates on IFDI flows from technologically-rich OECD countries into the emerging and developing economies. As a result, it may well be the case that technology transfers are more visible and have a significant impact for North-South FDI rather than North-North FDI.

Overall, the study provides evidence of a productivity-enhancing effect of FDI felt through foreign R-D spillovers. Further, the cointegration tests suggest the presence of a long-run relationship between TFP growth and the R-D spillovers through two channels – OFDI and IFDI. The inclusion of control variables highlights the importance of catalyst terms, such as human capital, in reducing the upward bias in the output elasticity of foreign R-D capital stock. We thus address two main subjects concerning the impact of FDI: the cross border flow of technology – through the traditional channel when DMNEs invest in emerging economies, and the more recent channel when EMNEs invest in advanced economies. Next, and more importantly, our analysis combines this flow of technology to explain the effect on TFP growth of emerging economies.

The theoretical framework of this paper therefore has much in common with the Schumpeterian premise that the economic growth of a nation depends upon the creation

of new technology or diffusion of technology (Tolentino, 1993). Since developing countries have limited technological capabilities, the transfer of technology across borders through direct investment provides the initial basis for technological development. Such transfers therefore assist the developing countries in narrowing down the technological gap, and hence attain profitable improvements in productivity and efficiency. However, the success of technology transfer depends on a number of factors. On the one hand, factors such as the ability of the country to adapt and develop the transferred technology could pose a challenge. This requires the development of human resources and a better absorption and implementation of advanced techniques through innovation. On the other hand, adequate infrastructure, such as scientific and technical training institutions, R-D facilities, and socio-economic environment play an important role in influencing the absorption of new and advanced technology. It is therefore imperative that government policies directed towards effective assimilation of foreign technology are set in motion.

It is often considered that OFDI enhances the interests of multinational corporations only; however the paper brings out opportunities for the economy as a whole by way of factors such as technology and productivity spillovers. With the understanding that OFDI has a positive impact for developing economies, the study encourages a high quality institutional environment to offer favourable conditions for running business, and hence make the business entities strong and competitive for foreign expansion. Given that OFDI constantly adds to the knowledge stock through reverse technology spillovers, and thereby affects productivity, the link between institutions and OFDI could be seen as a channel through which institutions promote productivity growth. In other words, the research suggests the need to devise strategies to develop a common lobby of interests between MNEs and policymakers in enhancing the positive effects of globalisation for growth and development of the country.

Notes

1. Dunning's IDP model suggests that a country's outward and inward FDI are partly a function of its level of development, and that countries go through different stages as their economy develops.
2. We consider technology transfers from resource rich to emerging economies. A number of reasons could be identified for why we take into account only FDI embodied spillovers, and not import embodied spillovers. First of all, it may be difficult to validate that the emerging economies are importing from advanced countries mainly to acquire strategic assets. Also, emerging economies such as India and China are rising exporters. Therefore by taking into account only FDI, we have concentrated on the increasing levels of outward investment from developing countries mainly aimed at fulfilling their asset-seeking motive. Further, two countries could import homogeneous goods from another country j , and such imports may benefit one country more than another. It may be difficult to ensure that goods sold by country j to country i were embodied by R-D intensity substantially different from the R-D intensity of goods sold to another country (Griliches and Lichtenberg, 1984). However, the FDI based weighting matrix that we attach embodies different R-D intensity for different i and j (as M&As, Greenfield investment, Joint Ventures are country specific). Also, even if FDI is classified as a technology flow matrix (rent spillover), the hierarchical clustering analysis shown in studies such as Van Pottelsberghe de la Potterie (1997) reflects that it is more likely to catch up knowledge spillovers than the Input-Output (IO) matrices because of a much closer clustering of the former to the technology proximity matrices. Lastly, studies have shown that knowledge spillover matrices yield higher returns than the IO matrix (Goto and Suzuki, 1989; Vuori, 1997; Verspagen, 1997).
3. TFP growth – measured as the change in GDP growth over the compensation-share weighted growth of combined factor inputs (labour and capital inputs, adjusted for change in their quality).
4. Modified versions of this methodology have also been tested and employed by Coe and Helpman (1995), Lichtenberg and Van Pottelsberghe de la Potterie (1998), and Zhu and Jeon (2007).

5. We use the TFP index values. These values are generated using the growth rates obtained from the *Total Economy Database* (the Conference Board). We do so to avoid the loss of observations [as $\ln(X)$ is undefined for $X < 0$] in using TFP growth rates.
6. Statistical data on FDI flows are more readily available than stocks. It is difficult to construct the stock values due to missing data and heterogeneous methodologies adopted in different countries.
7. Our source of data on Patent Applications is the *World Bank Database*, defined as – “Patent applications are worldwide patent applications filed through the Patent Cooperation Treaty (PCT) procedure or with a national patent office for exclusive rights for an invention”. PCT is an international treaty providing a unified procedure for filing patent applications, and incorporates both priority and second filings. Statistics based on PCT applications are less subject to geographic bias, eliminate any institutional bias, and the timeliness of the indicator is also good. Also as defined, in addition to PCT, the World Bank data also incorporate data from the National Patent Office. But as we are looking at emerging economies the bias of using domestic filings is subsided. This is because the home offices attract the majority of priority filings, as in the case of patents by inventors from developing countries such as Brazil, China, Russia (de Rassenfosse et al., 2013).
8. We carried out the Hausman test to select between the Fixed and Random Effects Model. The Hausman Test tests the null hypothesis that the coefficients estimated by the Random Effects Estimator are the same as the ones estimated by the Fixed Effects Estimator. If they are (insignificant P-value) then it is safe to use the Random Effects Model. However, if the P-value is significant, then the Fixed Effects Model is used. We obtain a significant P-value for the estimated equation, therefore justifying the use of the Fixed Effects Regression Model for our analysis.
9. A few drawbacks could be identified about the Patent Applications series used in this study. First of all, under PCT applications usually are of higher value, thus filtering out low value patents. As a result, it may put developing economies at a disadvantage. Also, companies (particularly small companies) are less likely to target foreign markets and mainly carry out inventions of local relevance. Overlooking these local patents therefore precludes a full view of the inventive activity of developing countries. This could also be a possible reason for an insignificant coefficient in our result. Further, PCT counts could be highly correlated with other counts such as USPTO, EPO and triadic. Therefore, as suggested by de Rassenfosse et al. (2013), the worldwide indicator that improves the measurement of inventive activity, especially in the case of emerging economies (because of no geographic bias and no filter on patent value) could be used to improve the estimation. However, as the main aim of this paper is not to focus on a detailed examination of patent counts, we tried to stick to using World Bank data to maintain a consistency for the countries under study.
10. The United States is a large industrialised country, and greater magnitudes of output elasticities of R-D spillovers embodied in OFDI and IFDI flows signify a much higher impact compared with other smaller OECD investing partners. This is consistent with the findings of previous studies (Van Pottelsberghe de la Potterie and Lichtenberg, 2001) that suggest a larger effect of FDI induced R-D spillovers for larger economies such as G7 countries.
11. Other than studying a different set of countries, our paper also differs from Van Pottelsberghe de la Potterie and Lichtenberg (2001) in terms of the definition of the key spillover variable. Unlike our study, they use a four year moving average of the flow of FDI. Also, the denominator of their spillover term is gross fixed capital formation of country j , whereas we use GDP of country j . These factors also explain differences in our results.

References

- Aw, B.Y., X. Chen and M.J. Roberts (2001), “Firm-Level Evidence on Productivity Differentials, Turnover, and Exports in Taiwanese Manufacturing”, *Journal of Development Economics* 66, pp. 51-86.
- Barba, N.G. and D. Castellani (2004), “Does Investing Abroad Affect Performance at Home? Comparing Italian Multinational and National Enterprises”, *Centre for Economic Policy Research Discussion Paper*, No. 4284.
- Barro, R.J. and J.W. Lee (2001), “International Data on Educational Attainment: Updates and Implications”, *Oxford Economic Papers* 53, pp. 541-563.
- Bera, Subhasis and Shikha Gupta (2009), “South-South FDI vs North-South FDI: A Comparative Analysis in the Context of India”, *ICRIER Working Paper* No. 238.
- Bernard, A.B. and B.J. Jensen (1999), “Exceptional Exporter Performance: Cause, Effect, or Both?”, *Journal of International Economics* 47, pp. 1-25.

- Bitzer, J. and H. Gorg (2009), "Foreign Direct Investment, Competition and Industry Performance", *The World Economy* 32, pp. 21-233.
- Bitzer, J. and M. Kerekes (2008), "Does Foreign Direct Investment Transfer Technology Across Borders? New Evidence", *Economics Letters* 100, pp. 355-358.
- Blomstrom, M. and A. Kokko (1998), "Multinational Corporations and Spillovers", *Journal of Economic Surveys* 12, pp. 1-31.
- Bosworth, B., S.M. Collins and A. Virmani (2007), "Sources of Growth in the Indian Economy", *NBER Working Paper Series*, No. 12901.
- Branstetter, L. (2006), "Is Foreign Direct Investment a Channel of Knowledge Spillovers? Evidence from Japan's FDI In The United States", *Journal of International Economics* 68, pp. 325-344.
- Clerides, S.K., S. Lach and J.R. Tybout (1998), "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco", *Quarterly Journal of Economics* 113, pp. 903-948.
- Coe, D.T. and E. Helpman (1995), "International R-D Spillovers", *European Economic Review* 39, pp. 859-887.
- Coe, D.T., E. Helpman and A.W. Hoffmeister (1997), "North-South R-D Spillovers", *Economic Journal* 107, pp. 134-149.
- De Rassenfosse, G., H. Dernis, D. Guellec, L. Picci and B. van Pottelsberghe de la Potterie (2013), "The Worldwide Count of Priority Patents: A New Indicator of Inventive Activity", *Research Policy* 42, pp. 720-737.
- Dernis, H., D. Guellec and B. Van Pottelsberghe de la Potterie (2001), "Using Patent Counts for Cross-Country Comparisons of Technology Output", *STI Review* 27, OECD Publishing, pp. 129-146.
- Engelbrecht, H.J. (1997), "International R-D Spillovers, Human Capital and Productivity in OECD Economies: An Empirical Investigation", *European Economic Review* 41, pp. 1479-1488.
- Goto, A. and K. Suzuki (1989), "R-D Capital, Rate of Return on R-D Investment and Spillovers of R-D in Japanese Manufacturing Industries", *Review of Economics and Statistics* 71, pp. 555-564.
- Govindarajan, V. and R. Ramamurti (2011), "Reverse Innovation, Emerging Markets, and Global Strategy", *Global Strategy Journal* 1, pp. 191-205.
- Griliches, Z. and F.R. Lichtenberg (1984), "Interindustry Technology Flows and Productivity Growth: A Reexamination", *Review of Economics and Statistics* 65, pp. 324-329.
- Guellec, D. and B. Van Pottelsberghe de la Potterie (2004), "From R-D to Productivity Growth: Do the Institutional Settings and the Source of Funds of R-D Matter?", *Oxford Bulletin of Economics and Statistics* 66, pp. 353-378.
- Herzer, D. (2012), "Outward FDI, Total Factor Productivity and Domestic Output: Evidence from Germany", *International Economic Journal* 26, pp. 155-174.
- Herzer, D. (2011), "The Long-Run Relationship between Outward FDI and Total Factor Productivity: Evidence for Developing Countries", *Proceedings of the German Development Economics Conference, Berlin*, No. 41, <http://hdl.handle.net/10419/48299>.
- Huang, S. and Q. Wang (2009), "Reverse Technology Spillover from Outward FDI: The Case of China", *Proceedings of the International Conference on Management of e-Commerce and e-Government*, pp. 550-553.
- Keller, W. (1996), "Absorptive Capacity: On the Creation and Acquisition of Technology in Development", *Journal of Development Economics* 49, pp. 199-227.
- Kimura, F. and K. Kiyota (2006), "Exports, FDI and Productivity: Dynamic Evidence from Japanese Firms", *Review of World Economics* 127, pp. 695-719.
- Kwark, N.S. and Y.S. Shyn (2006), "International R-D Spillovers Revisited: Human Capital as an Absorptive Capacity for Foreign Technology", *International Economic Journal* 20, pp. 179-196.
- Lall, S. (ed.) (1983), *The New Multinationals: The Spread of Third World Enterprises*, New York, John Wiley and Sons.
- Lichtenberg, F.R. and B. Van Pottelsberghe de la Potterie (1998), "International R-D Spillovers: A Comment", *European Economic Review* 42, pp. 1483-1491.
- Makino, S., C.M. Lau and R.S. Yeh (2002), "Asset-Exploitation versus Asset-Seeking: Implications for Location Choice of Foreign Direct Investment from Newly Industrialized Economies", *Journal of International Business Studies* 33, pp. 403-421.

- Narula, R. and J. Michel (2009), "Reverse Knowledge Transfer and its Implications for European Policy", *UNU-Merit Working Paper Series*, No. 2009-035.
- Pedroni, P. (1999), "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors", *Oxford Bulletin of Economics and Statistics* 61, pp. 653-670.
- Pesaran, M.H., Y. Shin and R.J. Smith (2001), "Bound Testing Approaches to the Analysis of Level Relationships", *Journal of Applied Econometrics* 16, pp. 289-326.
- Pradhan, J.P. and N. Singh (2009), "Outward FDI and Knowledge Flows: A Study of the Indian Automotive Sector", *International Journal of Institutions and Economies* 1, pp. 156-187.
- Pradhan, J.P. (2006), "Rise of Service Sector and Outward Foreign Direct Investment from Indian Economy: Trends, Patterns, and Determinants", *GITAM Journal of Management* 4, pp. 70-97.
- Tolentino, E.P. (1993), *Technological Innovation and Third World Multinationals*, London, Routledge.
- UNCTAD (2007), *Global Players from Emerging Markets: Strengthening Enterprise Competitiveness through Outward Investment*, New York and Geneva, United Nations.
- UNCTAD (2006), *World Investment Report 2006: FDI from Developing and Transition Economies: Implications for Development*, New York and Geneva, United Nations.
- UNCTAD (2005), Issues note "Internationalization of Developing country Enterprises through Outward Foreign Direct Investment" (TD/B/COM.3/EM.26/2) prepared for the Expert Meeting on "Enhancing the Productive Capacity of Developing Country Firms through Internationalization", Geneva.
- UNCTAD (2005), "Linkages, Value Chains and Outward Investment: Internationalization Patterns of Developing Countries' SMEs", TD/B/COM.3/69.
- UNCTAD (2005), "Policy Issues Related to Investment and Development: Emerging FDI from Developing Countries", UNCTAD/TD/B/COM.2/64.
- UNCTAD (2004), "India's Outward FDI: A Giant Awakening?", UNCTAD/DITE/IIAB/2004/1.
- UNCTAD (2003), "China: An Emerging FDI Outward Investor", UNCTAD/PRESS/EB/2003/08.
- UNESCO (1993), *Statistical Year Book*.
- Van Pottelsberghe de laPotterie, B. and F.R. Lichtenberg (2001), "Does Foreign Direct Investment Transfer Technology across Borders?", *The Review of Economics and Statistics* 83, pp. 490-497.
- Van Pottelsberghe de laPotterie, B. (1997), "Issues in Assessing the Effect of Interindustry R-D Spillovers", *Economic Systems Research* 9, pp. 331-356.
- Verma, R. (2008), "The Service Sector Revolution in India", *UNU-WIDER Research Paper*, No. 2008/72.
- Verspagen, B. (1997), "Measuring Intersectoral Spillovers: Estimates from the European and US Patent Office Databases", *Economic Systems Research* 9, pp. 47-66.
- Vuori, S. (1997), "Interindustry Technology Flows and Productivity in Finnish Manufacturing", *Economic Systems Research* 9, pp. 67-80.
- Wells, L.T. (1983), *Third World Multinationals: The Rise of Foreign Investments from Developing Countries*, Cambridge, MIT Press.
- Zhao, W. and L. Liu (2008), "Outward Direct Investment and R-D Spillovers: China's Case", Conference Proceeding of "Emerging Multinationals": Outward Foreign Direct Investment from Emerging and Developing Economies.
- Zhu, L. and B.N. Jeon (2007), "International R-D Spillovers: Trade, FDI, and Information Technology as Spillover Channels", *Review of International Economics* 15, pp. 955-976.

APPENDIX A

List of 18 emerging economies

Argentina
Brazil
Bulgaria
China
India
Indonesia
Latvia
Lithuania
Malaysia
Peru
Philippines
Romania
Russia
Singapore
South Africa
Thailand
Ukraine
Venezuela

APPENDIX B

List of 34 OECD countries

Australia
Austria
Belgium
Canada
Chile
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Israel
Italy
Japan
Korea
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Slovenia
Spain
Sweden
Switzerland
Turkey
United Kingdom
United States

APPENDIX C

Summary results for larger industrialised (G7) partner countries

Estimation results – including dummy for G7 countries

	Within	
	(i)	(ii)
Domestic R-D		
$LnRD^d$	12.4063* (2.75)	12.2284* (3.76)
Foreign R-D		
$LnRD^{fo}$ (outward FDI)	1.9862 (1.52)	-
$LnRD^{fi}$ (inward FDI)	-	7.5333* (7.60)
Larger countries		
$LnRD^{fo*} \text{ G7}$	2.3661** (1.85)	-
$LnRD^{fi*} \text{ G7}$	-	-1.0824 (-1.04)

Notes: Figures in parentheses give t-statistics. * denotes significance at the 5% level, ** denotes significance at the 10% level. RD^d = Domestic R-D capital, RD^{fo} = Foreign R-D capital embodied in OFDI, RD^{fi} = Foreign R-D capital embodied in IFDI.

The table shows the results that allow output elasticities with respect to foreign R-D capital stock to differ between G7 and other countries. The additional variables $LnRD^{fo*} \text{ G7}$ and $LnRD^{fi*} \text{ G7}$ are created by including a dummy that takes value one for G7 countries in the construction of foreign R-D spillovers term. This is done to test whether the estimated parameters are different for larger than for smaller investing partner countries.

The results clearly indicate that in the case of foreign R-D embodied in OFDI from the 18 emerging economies, the impact is much greater and significant when these economies undertake investment in the larger industrialised countries. However, the reverse is true for the spillovers embodied in IFDI flows into the emerging economies.

Environmental policies and productivity growth – a critical review of empirical findings

by
Tomasz Koźluk and Vera Zipperer*

The economic effects of environmental policies are of central interest to policymakers. The traditional approach sees environmental policies as a burden on economic activity, at least in the short to medium term, as they raise costs without increasing output and restrict the set of production technologies and outputs. At the same time, the Porter Hypothesis claims that well-designed environmental policies can provide a “free lunch” – encouraging innovation, bringing about gains in profitability and productivity that can outweigh the costs of the policy. This paper reviews the empirical evidence on the link between environmental policy stringency and productivity growth, and the various channels through which such effects can take place. The results are ambiguous, in particular as many of the studies are fragile and context-specific, impeding the generalisation of conclusions. Practical problems related to data, measurement and estimation strategies are discussed, leading to suggestions as to how they can be addressed in future research. These include: improving the measurement of environmental policy stringency; investigating effects of different types of instruments and details of instrument design; exploiting cross-country variation; and the complementary use of different levels of aggregation.

JEL classification: D24, Q50, Q55, Q58, O47, O31.

Keywords: Environmental policy, Porter Hypothesis, Productivity, Innovation.

* The authors are members of the Economics Department and Environment Directorate of the OECD (tomasz.kozluk@oecd.org and vera.zipperer@oecd.org). They would like to thank Jean-Luc Schneider, Giuseppe Nicoletti, Shardul Agrawala, Silvia Albrizio, Ivan Hašič and Jehan Sauvage, for their useful comments and suggestions. They would also like to thank Stefan Ambec and Victoria Shestalova for their constructive comments and insightful discussion. Special thanks go to Lyn Urmston and Diane Scott for editorial support. The opinions expressed and arguments employed in this paper are the responsibility of the authors and do not necessarily reflect the official views of the OECD or of the governments of its member and partner countries.

1. Introduction

The principal goal of environmental policies is to improve environmental outcomes, driven by the pursuit of objectives of broader wellbeing and ensuring sustainable growth. Environmental policies aim to achieve their objective by increasing the opportunity costs of pollution and environmental damage, curbing polluting behaviour, supporting investment and inducing innovation in less environmentally harmful technologies and so forth. However, they are likely to affect purely economic outcomes as well, particularly in the shorter term; these effects are of interest to policy makers when choosing to take action to improve environmental performance and selecting the relevant policy instruments.

A priori, the direction of the effect of environmental policies on macroeconomic variables such as GDP, productivity, innovation, employment, investment or trade, is far from clear. Traditionally, more stringent environmental policies have often been viewed as burdensome to economic performance – for instance by posing additional costs on producers without increasing output levels. Hence, they were considered detrimental to the profitability of the firm, productivity growth and GDP. For example, a number of studies attempted to attribute a significant part of the 1970s productivity slowdown in the United States to the increasing role of environmental policies (see Christainsen and Haveman, 1981, for a review). On the other hand, in the early 1990s, Michael Porter suggested that well-designed environmental policies might actually enhance productivity and increase innovation, yielding direct economic benefits next to the environmental benefits (Porter, 1991; Porter and van der Linde, 1995).

Empirical evidence on the effects of environmental policies on economic variables is rather weak. Many studies have been undertaken in the context of international trade, but empirical evidence on the effect on productivity is often context-specific and inconclusive. Early studies which were undertaken with the focus of explaining the contribution of environmental policy to the 1970s productivity slowdown in the United States, indeed found a negative effect on productivity growth, but more recent studies of the same phenomenon show either no effect at all or even a positive effect on productivity. Many of the attempts, in particular the earlier studies, do not address simultaneity issues, potentially attributing effects of other factors to environmental policies. Similarly, the majority of studies do not allow distinguishing among effects of different design aspects of environmental policies.

Several aspects of environmental policies matter for economic outcomes and are considered important elements of well-designed policies – in particular stringency, flexibility, predictability and competition friendliness. The objective of this paper is to review the existing empirical literature on the link between the stringency of environmental policies on the one side, and productivity and innovation, on the other. The paper also reviews the scarce literature on the effects of policy flexibility, and gives an overview of the empirical evidence on environmental policies' effects on the capital stock,

investment, and trade flows. The guiding questions of this analysis are: What are the empirical findings regarding the relationship between environmental policies on the one side and productivity and innovation on the other side? Do the findings differ at the firm/industry/macro level? What may explain the differences found? Which challenges need to be addressed in future research?

Section 2 of this paper discusses the background intuition for the economic effects of environmental policies, the nature of the expected effects and potential channels through which they may work. Section 3 reviews the results of different empirical studies. Section 4 discusses challenges for empirical analysis, including an overview of the different measures of environmental policies used in the literature. Section 5 concludes with a summary of the main outstanding issues regarding investigation of such effects and provides some insight for future research.

2. Background on the channels through which environmental policies may affect economic outcomes

In a sense, the natural environment is an input into practically all economic activities. Environmental policies that guarantee the sustainable use of environmental assets can be seen as ensuring a certain level of aggregate productivity growth over the long term. While this is clearly not the only motivation for environmental policies, a counterfactual, disruptive development could mean lower overall productivity growth at some point, though potentially to the benefit of higher temporary growth in the shorter term.

Leaving the sustainability argument aside, long-term effects of an increase in the stringency of environmental policies on productivity growth, if present, are likely to hinge on the nature of the policy tool in question. Fairly static or one-off policy changes are likely to permanently affect productivity levels, but only temporarily affect productivity growth rates. Long-term changes in productivity growth could occur if, for instance, environmental policies provide permanent incentives to innovate more. This could be the case for more dynamic types of policies, increasing in stringency, such as emission caps, increasing environmental tax rates or performance standards with preannounced paths. At the same time, to the degree environmental policies decrease competitive pressures, for instance by raising entry barriers or directly discriminating between incumbents and new firms, they may pose a drag on productivity growth or innovation (see Aghion et al., 2005; Aghion et al., 2004; Boursès et al., 2010; for discussions on entry barriers and competition). In the absence of more dynamic incentives, there are still various views on the direction and magnitude of short-term and medium-term effects, as discussed later in this section. While much of the short to medium term discussion applies primarily to traditionally-measured multifactor productivity (MFP) growth, many of the arguments can also be made in the context of productivity measures that are adjusted for environmental inputs.

2.1. The conventional argument – diverting resources from productive allocations

Traditional views of environmental policies tend to see them as a cost or burden to economic activity, at least in the short to medium term. Leaving aside the long-term sustainability argument, compliance with environmental policies generally forces firms to devote some part of inputs to pollution prevention and abatement, which are not traditionally considered as value added, or to curb production (Jaffe et al., 1995; Ambec et al., 2013). The effects can be both direct – with firms' costs rising directly due to pollution

abatement – and indirect, such as through increases in input prices in the industries affected by the regulation (Barbera and McConnell, 1990).

More specifically, policies such as environmental taxes, tradable emission permits, or water and wastewater treatment charges impose an additional production-related cost for something that would be free otherwise, hence can induce firms to devote resources to reduce costly output. A similar argument can be developed for performance standards. Technology-based standards and restrictions or bans on certain substances can limit the range of production technologies available and otherwise profitable production output. In principle, and abstracting from market imperfections for the moment, if productivity could be increased by curbing environmental externalities, firms would have done so already, even in the absence of environmental policy. Hence the effect of environmental policies on productivity should be negative or, at best, neutral.

The effects of environmental policies on productivity might, however, be more complex as indirect effects work through a variety of channels. In the short term, environmental policies may actually improve productivity in some specific activities – for instance industries using water as an input may benefit from the fact that water becomes cleaner, via reduced inputs devoted to water purification (Jaffe et al., 1995). Similarly, workers may become more productive if the adverse effects of air pollution on their health are curbed (Ostro, 1983, or more recently Graff Zivin and Neidel, 2012). Similarly, to the degree that some environmental policies provide budgetary revenues, the use of these may have effects on aggregate productivity – for instance if these are used to reduce distortionary taxation or reduce relative labour costs – as often argued in light of the double-dividend of pigouvian-style taxes.

Indirect effects may come in the form of effects on the survival rate of firms. If, by imposing additional costs, environmental protection eliminates less efficient firms from the market, aggregate productivity is likely to rise. At the same time, additional (entry) costs, or vintage differentiated regulations can discourage entry and exit, reducing competition in the market, shielding potentially inefficient incumbents and obsolete capital stock, thereby leading to lower productivity levels and growth (Mohr and Saha, 2008; OECD, 2006; Heyes, 2014). Some further empirical evidence suggests that the plant birth rate is negatively influenced by environmental regulations (Gray, 1997).

Finally, environmental policies may also foster the creation of industries or activities that would otherwise not exist or not benefit from economies of scale. The effects on productivity, while unlikely to be large on aggregate, are complex and uncertain *a priori*. For example, regulations that imply certain monitoring requirements may directly decrease the productivity of the sector concerned but may also boost both monitoring services and the production of relevant equipment – providing a first-mover advantage, if other countries eventually adopt similar laws. Municipal waste and wastewater treatment laws can induce new demand for such services and foster the growth of companies providing dedicated services. To the extent that such environmental policies involve public subsidies, aggregate effects need to take into account that these funds are taxed away from productive activities.

The extent to which different effects of environmental policies can be observed will depend on the level of aggregation. The plant and firm level allow observing the effects on changes in production processes and technologies, the diversification or changes in activity and products and the related potential costs and synergies. Industry level analysis

will pool some of these effects together with those of firm entry and exit, and changes in firm size. It will also depend on the actual level of industry aggregation. At the country level, changes in the industrial structure both as a result of policies and as result of changes in environmental quality are also likely to be present – but on the other hand, can be lost in the process of aggregation or cumbersome to identify. All the listed direct effects can be further complicated by spillover effects across firms, industries and countries as well as trade (Sauvage, 2014).

The overall rather negative perception of the effects of environmental policies on productivity growth stems partly from the nature of traditionally-measured MFP growth, which does not take into account environmental effects. In a simplified setting, it can be shown that environmental policies will negatively affect traditionally-measured MFP by definition (Box 1), even when they have no effect on broader productivity. Such ideas have led to various attempts at calculating adjusted MFP, or so-called “green” MFP, (see for instance Repetto et al., 1997; or Brandt et al., 2014) which would take into account the fact that pollution, or more broadly environmental services, can be seen as either an omitted output (i.e. side-product) or an omitted input. This concept also lies behind the distance function approach taken in several macro-level studies and discussed in the section below (Box 2).

2.2. The Porter Hypothesis – a free lunch?

If policies actually increase innovation they can enhance productivity growth. The general idea behind induced innovation goes back to Hicks (1932), and hinges upon the notion that increasing a factor’s cost should spur innovation to economise on the use of this factor.¹ In the 1990s, Michael Porter argued that indirect effects of well-designed environmental regulatory instruments might induce firms to innovate, which in turn might increase productivity and hence profitability – potentially outweighing the increase in abatement costs (Box 3). In this vein, environmental regulation which would set dynamic, long-lasting incentives to innovate more could increase both the level and the growth rate of productivity. These indirect effects might come through changes in the composition of inputs or through changes in the whole production process as the new regulation might shift the production function through stimulating the invention and adoption of new technologies. *A priori*, it is however unclear whether these indirect effects are negative or positive, or whether they are large enough to outweigh the drag of the direct effect. As Barbera and McConnell (1990) point out, more traditional inputs such as labour and capital might be needed as complements to the abatement capital and hence lead to a negative indirect effect. On the other hand, old capital might be replaced with new investment, thereby increasing efficiency and leading to a positive direct effect.

Porter’s arguments have met significant scepticism, though they had an important influence in environmental economics. Critics generally focus on the free-lunch argument, that is, that if there were productive opportunities available, they would have already been exploited by the firm. It is also unclear why policy-makers would be better aware than firms of the existence of these *win-win* areas to firms. Defendants tend to make the argument that because of market failures (e.g. information asymmetries), firms may be unwilling or unable to take a certain risk or action with net expected gains (Porter and van der Linde, 1995). Notably, Porter’s arguments tend to rely on “*well-designed*” regulation. This

Box 1. Effects of environmental policies on measured MFP growth

Environmentally-adjusted measures of productivity growth are not aimed at answering the question about productivity effects of environmental policies *per se*. These productivity measures are rather developed to improve the measurement of productivity in the first place and can then be used to conduct analyses of the impact of environmental policies. Even assuming away any effects of environmental policies on actual MFP growth, traditional measures of MFP growth are likely to show changes due to the omission of environmental services in the production function. The intuition behind this can be demonstrated in a simple Cobb-Douglas production function with constant returns to scale and three inputs, capital (K), labour (L) and environment (E):*

$$Y = AK^\alpha L^\beta E^{1-\alpha-\beta} \quad [1]$$

taking logarithms and differentiating with respect to time:

$$\hat{A} = \hat{Y} - \alpha \hat{K} - \beta \hat{L} - (1 - \alpha - \beta) \hat{E} \quad [2]$$

where $\hat{X}$ refers to the growth rate of X (i.e. $\frac{d \ln X}{dt}$).

The Solow residual is the actual or “environmentally adjusted” MFP growth, in the sense that it captures the difference between the increase in outputs and the increase in capital, labour and environmental inputs used to produce them. At the same time, if the role of environmental inputs is ignored in the basic production function:

$$Y = B * K^a L^{1-a} \quad [3]$$

the “measured” growth in MFP is equal to:

$$\hat{B} = \hat{Y} - a \hat{K} - (1 - a) \hat{L} \quad [4]$$

Assuming that the relative cost shares of labour and capital are not affected by the choice of measurement approach i.e. $\frac{a}{1-a} = \frac{\alpha}{\beta}$, equation [2] can be substituted into [4], which after rearranging gives:

$$\hat{B} = \hat{A} + (1 - \alpha - \beta) \left[\hat{E} - \left(\frac{1}{\alpha + \beta} \right) (\alpha \hat{L} + \beta \hat{K}) \right] \quad [5]$$

In which case, “measured” MFP growth ($\hat{B}$) is equal to “actual” MFP growth ($\hat{A}$) plus a bias. As the share of environmental services gets close to zero, this bias tends to disappear as $(\alpha + \beta)$ approaches 1.

Assuming a new or a more strict environmental policy affects the (implicit or explicit) price of environmental inputs without affecting the prices of labour or capital, it will reduce the use of environmental inputs ($\hat{E} \downarrow$) and increase the use of labour ($\hat{L} \uparrow$) and/or capital ($\hat{K} \uparrow$). This means the bias will go down in levels. In a more complicated setting this could affect cost shares, but for simplicity this effect is ignored here.

If actual MFP growth ($\hat{A}$) is unaffected (as assumed initially), the lower bias implies that the growth of measured MFP ($\hat{B}$) will inevitably decrease. The intuition behind this measurement phenomenon can be found in Repetto et al. (1997): if the environmental policy change results in the substitution of part of a previously unobserved input (E, which in this case is treated as part of the residual, B) with an observed input (K or L, for instance through the installation of pollution abatement equipment or hiring of monitoring personnel), leading to lower observed productivity. As the bias depends on the environmental policy regime in place, it might change with a change in environmental policy. One-shot policies and long-term incentive-based policies may have different effects on the over-time prevalence of such a change in the bias.

* Environmental inputs can include any type of environmental services, such as natural resources, sink functions for pollution or land use.

makes inference from empirical estimations increasingly difficult because it can often be argued that the insignificant or negative effects of a specific environmental regulation on productivity come from its poor design, not from a negative effect *per se*.²

Box 2. **Measuring productivity to account for environmental services**

Environmental services, such as the use of natural inputs and sink services provided by the environment are ignored in traditional measures of productivity. In response, an increasing number of studies attempt to incorporate some of these services. Two methodologies dominate the literature – a distance function approach and an adjusted growth accounting framework.

In the distance function framework (including data envelopment analysis, DEA; and stochastic frontier analysis, SFA) quantity data on inputs, traditional outputs and “bad” outputs are used to determine the technological production frontier, describing all possible efficient input/output combinations. For each possible input-output combination, the distance function then measures the possible efficiency gain of moving from an inefficient point on to the frontier. Studies that evaluate the effects of environmental regulation on productivity often model three scenarios by varying the assumption about the bad output: free disposability of bad outputs; a constant level of bad outputs while good outputs increase; and, a reduction of bad outputs at the same time as good output increases.

The second approach, based on a growth accounting framework, adjusts traditional productivity growth by the weighted difference of bad output and input growth. This requires assumptions on explicit shadow prices of pollutants (for an application see Brandt et al., 2014).

In practical applications, a key problem may be linked to the choice of environmental “bads” and their implicit weights, which depend on the underlying shadow prices. Studies tend to differ vastly in this respect, hindering comparability and generality of conclusions. Among bad output measures used are CO₂ emissions (Wu and Wang, 2008), a combination of air and water pollutants (Yoruk and Zaim, 2005), air, water and toxic pollutant releases (Boyd and McClelland, 1999). Inevitably, such adjustments need to limit the bad outputs to a handful of environmental issues, meaning they may also hide an increase in other, non-measured environmental bads. Shadow prices of pollutants are rarely available and need to be estimated separately or proxied with observed prices, for example resulting from policies such as environmental taxes. Both approaches are fragile to the underlying assumptions. The distance function framework allows calculating implicit shadow prices of pollutants, which are the basis for the productivity adjustment. As most of the studies do not attempt to calculate these prices, it is often not possible to assess whether they take reasonable values or compare them to social costs of pollutants, such as those estimated from life and health valuations, or those derived from surveys of revealed preferences – the willingness to pay for reducing certain environmental externalities or willingness to accept a certain compensation for these.

Source: Brandt et al. (2014).

2.3. Trade, foreign direct investment and competitiveness

A large literature directly links environmental policies to competitiveness and trade outcomes (see Ambec et al., 2013; Wagner, 2003; Jaffe et al. 1995, for reviews). This literature is closely linked to the theory of pollution havens – pollution-heavy firms and investment relocating to environmentally lax countries. On the other hand, a number of

Box 3. Various versions of the Porter Hypothesis and some of the theoretical underpinnings

In the early 1990s, the debate around the economic impacts of environmental regulation gained new momentum through an article written by M.E. Porter. He claimed that properly designed environmental regulation can trigger innovation which in turn can decrease, and even offset the costs of pollution abatement and enhance competitiveness (Porter, 1991; Porter and van der Linde, 1995). Initially, the Porter Hypothesis was suggested without any theoretical explanation of the factors at work or any comprehensive empirical evidence aside from a few case studies which were collected by the authors.

The theoretical explanations for the possible existence of a Porter effect were only developed at a later stage. Generally speaking, the Porter Hypothesis assumes that there are profit opportunities for firms which are not fully used until the firms are pushed to do so by the implementation of a new environmental policy. A comprehensive overview of the development of the different theoretical arguments to provide a foundation for the Porter Hypothesis can be found in Ambec et al. (2013) and Wagner (2003). Behavioural arguments are based on the idea that managers may be risk averse, myopic or rationally bounded and hence may not be able to realise all profitable investment opportunities. Environmental regulation might then require certain investments which turn out to be profitable. Another approach hinges on the presence of market failures, such as imperfect competition (due to first-mover advantage or barriers to entry) or high entry costs, asymmetric information (where “green” products are not correctly valued by consumers), R&D spillover effects (as innovation has a public good character and leads to underinvestment), co-ordination failure and organisational failure (where managers are able to lie about the true abatement costs in order to secure extra personal profits).

In order to find feasible empirical testing approaches, the Porter Hypothesis was divided into different sub-aspects which are nowadays known as the weak, strong and narrow version of the Porter Hypothesis (Jaffe and Palmer, 1997, were the first to differentiate the various aspects):

- The *weak version* of the Porter Hypothesis implies that environmental regulation will lead to an increase in environmental innovation. As Jaffe and Palmer (1997) describe, the firms subject to the new environmental regulation face an additional environmental constraint next to their financial ones. As firms are assumed to maximise profits, they will search for the most cost effective way to comply with the new regulation. The Porter Hypothesis suggests that firms will do so by innovating to reduce compliance costs, but does not necessarily imply more innovation in total.
- The *strong version* of the Porter Hypothesis claims that cost savings from the improved production processes are sufficiently large to increase competitiveness. It rejects the assumption of perfect markets with profit maximising firms and assumes instead that firms are not operating fully efficiently by leaving some profit opportunities unused. Environmental policies might hence induce the firm to rethink their production process. This might lead to extra profits which can in some cases be even larger than the cost of compliance. Various versions of this hypothesis have been tested, including the effects on more innovation and on actual company performance.
- Jaffe and Palmer (1997) describe also the *narrow version* of the Porter Hypothesis – certain types of environmental regulation, those which are designed to target the outcome rather than the design of the production processes, are more likely to increase innovation and improve company performance. In practice, this has been approached by looking at how the use of more flexible or market-based instruments affects environmental and total innovation, respectively.*

* Following De Serres et al. (2011) this review uses the term “market-based instruments” for environmental policy instruments aimed at addressing market failures mainly through price signals. This can include taxes, charges and fees, tradable permits, and subsidies for reducing pollution. Non-market approaches include direct environmental regulations, standards (“command-and-control” instruments), and voluntary approaches.

scholars emphasise the first-mover advantages and accompanying gains in the technological frontier. By exploiting and affecting competitive advantages, international trade can also mitigate potential adverse effects of environmental policies on productivity outcomes, meaning desired environmental outcomes can be achieved at lower economic cost. These perspectives are in principle compatible with Porter's ideas – specialisation and relocation may be an efficient response and does not preclude improvements in productivity, both at home and abroad. Productivity is a key driver of competitiveness, but given that competitiveness is driven by several other key aspects, in particular costs and location, and is a relative rather than an absolute concept (Krugman, 1994), a detailed review goes beyond the scope of this paper.³

Research in this area indicates that environmental policies are unlikely to be major determinants of a country's competitiveness, trade or FDI patterns (see Jaffe et al., 1995; Brännlund and Lundgren, 2009, Ambec et al., 2013, for reviews). This is probably because environmental costs are generally a small fraction of total firms' costs, which would also suggest relatively small effects on productivity growth. Competitiveness effects may also depend on country characteristics such as the level of development and income, trade openness, industrial structure and environmental, labour and other endowments, making the comparison of results of studies across different countries and periods difficult. Notably, the competitiveness literature has usually concentrated on the stringency of policies, rarely attempting to discriminate between different types of environmental policy tools.

3. Empirical evidence

In this review empirical research on the economic effects of environmental policies is classified according to the macroeconomic measure analysed and the level of aggregation of the study. Details of the studies cited regarding scope, methodology and results can be found in Appendix A, Table A.1.

3.1. Productivity

The majority of studies on the productivity effect of environmental policies are conducted at the firm- and industry-level, with only a few papers adopting a macroeconomic view. Common approaches include cost-function estimates, growth accounting and efficiency measures adjusted for environmental outputs (Box 2).

3.1.1. Firm- and plant-level studies using traditional productivity measures

Plant-level studies tend to compare traditionally measured productivity growth among regulated and non-regulated plants, with estimated effects being overall rather negative but not very robust. Effects of environmental regulation on MFP growth are found to be negative (Gollop and Roberts, 1983; Smith and Sims, 1985) or positive but insignificant (Berman and Bui, 2001a, in this case on productivity levels). None of these studies are able to control for potentially different characteristics of the two groups of plants, or to make a convincing case for ignoring such differences, which is a flaw in these types of fixed-effect or difference-in-difference approaches. Studies that do attempt to control for plant-level characteristics, including self-selection of firms into countries with more lax environmental regulation, tend to find many of them as significant determinants of productivity (for example, Becker, 2011; in the case of labour productivity).

Even within one industry (pulp and paper in this case) effects of environmental regulation on plant productivity can depend strongly on plant characteristics. Gray and Shadbegian (2003) show that integrated mills, subject to stricter environmental regulation due to the integrated pulping process, show a significant reduction of productivity due to increases in abatement costs, whereas productivity of non-integrated mills does not show any substantial reduction. Similarly, Becker (2011) finds no effect of environmental regulation on labour productivity levels in a broad sample of plants. Reducing the sample to plants which experienced a statistically-significant change in compliance costs over the years (only one-tenth of the sample), the study finds a negative effect of compliance costs on labour productivity.

At the same time, the effects of pollutant-specific regulations on productivity levels can be very different, depending on the pollutant. Greenstone et al. (2012) find a persistent negative effect of total environmental regulations, while ozone regulations and particulates emission regulations are estimated to have a negative effect on productivity levels, sulphur dioxide emission have no significant effect and carbon monoxide regulations even foster productivity. The paper does not discuss the reasons for such differences in outcomes, which may come from a number of areas, for instance be due to different prevention and abatement technologies readily available in case of each of the various pollutants.

Aside from methodological issues the conclusions from the bulk of studies suffer from the lack of generality. They evaluate very specific regulations, focus on firms in specific industries (electric power; brewing; pulp and paper; and manufacturing respectively) in different countries (United States and Canada) or areas (coastal California), different time periods and horizons. Only Greenstone et al. (2012) account for plant closure while Becker (2011) and, to a lesser extent Gray and Shadbegian (2003), explicitly control for some firm characteristics that may drive productivity. None of the studies attempts to control for spill overs across firms.

3.1.2. Firm- and plant-level studies using alternative measures of productivity

A number of more recent studies take into account bad outputs in the production process – and hence in the calculation of productivity – with the help of data envelopment analysis or stochastic frontier analysis (DEA and SFA respectively, Box 2). Interestingly, even in this case, the results are far from uniform. Negative effects on adjusted productivity are found by Broberg et al. (2012) for the heavily regulated Swedish pulp and paper industry, with no effect on the manufacturing sector. The specification for the pulp and paper industry however, seems questionable due to an overall poor fit of the model – as opposed to the pooled data sample which includes several industries; none of the traditional factors of the production function are found significant. Fleishman et al. (2009) find a multitude of different effects of the presence (but no effects of stringency) of air pollutant regulations for US power plants – positive, insignificant and negative – depending on types of plants and types of regulation. Positive effects of more stringent policies on adjusted technical efficiency are found in the Dutch horticulture sector (Van der Vlist et al., 2007) and companies under the EU ETS (Jaraite and Di Maria, 2012). Both studies examine fairly narrow definitions of stringency and do not investigate medium- to long term effects. Managi et al. (2005), who allow for time-varying effects, find negative short-term effects of regulation on productivity levels in US offshore oil and gas fields, which disappear over time, eventually becoming positive and outweighing the initial drag on productivity.

3.1.3. Industry-level studies

The empirical evidence produced by industry-level studies is even more ambiguous. Early studies tend to find a negative effect, while more recent ones suggest a positive link between environmental regulation and productivity. These older studies tend to suffer from serious identification problems. One of the more widely cited studies that attempts to explain the US productivity slowdown in the 1970s with environmental regulation finds strong significant results, which disappear upon the inclusion of controls or elimination of outliers (Gray, 1987). Other studies that find negative effects on productivity growth either are unable to control for industry characteristics (Barbera and McConnell, 1990) or do so, but remain questionable due to a small sample (Dufour et al., 1998).

As in the case of plant-level studies, the industry studies analyse rather specific cases – countries, a handful of very specific industries (five in Barbera and McConnell, 1990; Hamamoto, 2006), and specific environmental laws. While Gray (1987) and Yang et al. (2012) analyse a set of several hundred industries, in other cases the cross-section dimension is more limited. The somewhat longer time series available allows some studies to attempt to look at the time dimension of effects (Hamamoto, 2006; Lanoie et al., 2008). The results potentially reflect the bulky nature of up-front expenditures in reaction to changes of environmental laws – finding the negative contemporaneous effect outweighed by the subsequent positive effects of regulation on MFP growth (Lanoie et al., 2008). On the contrary, Hamamoto (2006) finds significant positive effects of command-and-control regulations in Japan – on R&D spending and in a second step on productivity growth – which tend to decline over time.

Differentiating between pollution abatement fees and pollution abatement capital expenditures can allow for dealing with some of the simultaneity problems of the Pollution Abatement Costs and Expenditures (PACE, see Box 4) environmental proxies. Capital expenditures are found to have no effect while abatement fees are found to stimulate R&D and higher productivity (Yang et al., 2012), though the link between this (environmentally) induced R&D and MFP levels is found weaker than in terms of general R&D spending.

Adjusting productivity for “bad” outputs in a DEA framework does change the overall mixed conclusions – Domazlicky and Weber (2004) find no evidence of a link between environmental regulation and adjusted productivity growth in the US chemical sector. The only cross-country study, Alpay et al. (2002), similarly finds no effect on adjusted productivity growth on the US food manufacturing sector, but a positive one in Mexico.

3.1.4. Macroeconomic studies

Empirical evidence at the aggregate economy level is very limited, largely because of data and identification problems. The approach taken by all of the studies surveyed is to include bad outputs in a distance function framework and hence allow crediting the reduction of pollution. Evidence is fairly inconclusive – scenarios with constant bad outputs (as a result of environmental policies) find negative effects on adjusted MFP growth (Jeon and Sickles, 2004) or slightly positive effects in a different sample of countries (Wu and Wang, 2008).

Interestingly, even in cases where the ratification of the UNFCCC is used as a proxy for environmental policy stringency, the effects on adjusted MFP growth can be positive (Yörük and Zaim, 2005) or negative (Wu and Wang, 2008). Explanations may relate to different samples – Yörük and Zaim (2005) focusing on OECD economies, Wu and Wang (2008) on

Box 4. **Abatement expenditures as proxy for stringency of environmental regulation**

Firm and industry-level data on environmental protection expenditure, defined as expenditures on “purposeful activities aimed directly at the prevention, reduction and elimination of pollution or nuisances arising as a residual of production processes” (OECD, 2007), have been collected since the 1970s by national authorities and the OECD. The most known example is the PACE – a plant-level survey that asks questions on pollution abatement capital expenditures and operating costs associated with compliance to local, state, and federal regulations and voluntary or market-driven pollution abatement activities. The PACE survey was conducted by the US Census Bureau between 1973 and 1994 (annually, with the exception of 1987), and discontinued thereafter. The Environmental Protection Agency picked up the bill for a 1999 survey, which, however, had significant conceptual differences relative to the previous waves, complicating time-series analysis (Becker and Shadbegian, 2004). A new survey followed in 2005, and is for the moment the last one conducted. A similar, though not directly comparable concept governs the Joint OECD/Eurostat Questionnaire on Environmental Protection Expenditure and Revenues (EPER), which commenced in the late 1990s for EU countries, and several other exercises across countries such as in Canada or Korea.

Empirical studies that use environmental pollution abatement costs as a measure of environmental policy stringency rely primarily on the assumption that higher environment-related expenditure is induced by more strict environmental policy. In practice this concept faces a number of drawbacks, some of which are emphasised in an international context. First of all, the data are often not easily comparable across countries and over time. Second, it is difficult to classify expenditures – for example distinguishing what share of costs of a change in technology is driven by environmental policies and what share by profits. This relates to the issue of counterfactuals, that is identifying an expenditure scenario without environmental policies. The self-reporting nature of the exercise may exacerbate such issues, as different firm (and country) characteristics may influence responses. In fact, Broberg et al. (2012) claim that the pulp and paper industry generally categorises investments into closed-loop systems, which recycle and reuse wastewater, as environmental investment even though it is purely driven by profitability considerations. Over-reporting may also be a result of attempts by industries to gain a “green” image or to signal voluntary efforts in order to avoid hard policy intervention. On the other hand, Berman and Bui (2001a) claim that costs of abatement are often incompletely measured, such as for example a switch to another fuel where the extra cost of this switch is not reported. Arguments in both directions are also provided by Brunnermeier and Cohen (2003). Moreover, the share of such types of expenditure may also reflect regulations that are external to environmental policies – for example stricter safety regulations may improve safety (and hence lower the need for environmental pollution abatement). Finally, there is a sample selection issue: if environmental policies lead to changes in the industry structure, with firm entry and exit, these phenomena may not be easily captured with firm-level questionnaires.

Source: Becker and Shadbegian (2004); OECD (2007).

APEC economies – or differences in the set of bad outputs used to adjust MFP growth – Yörük and Zaim (2005) using an average of air and water pollutants, while Wu and Wang (2008) focus only on CO₂. More generally, in both cases the protocol ratification dummy

seems rather prone to capturing a large number of other effects – an issue that would require further investigation.

3.2. Innovation as a driver of productivity growth

Innovation is a central aspect of productivity growth and has been studied in the context of effects of environmental policy stringency. The “weak” version of the Porter Hypothesis (Box 3) – more stringent environmental regulation will increase environmental innovation⁴ – is basically a variation of Hick’s (1932) argument and is fairly well supported by empirical evidence. On the contrary, there seems to be little evidence of firms actually innovating more overall:

- At the plant-level, stringency of policy tends to have a fairly strong effect on the decision to engage in environmental R&D (Johnstone and Labonne, 2006; Arimura et al., 2007; Lanoie et al., 2011; Yang et al., 2012).
- At the industry level, Jaffe and Palmer (1997) and Hamamoto (2006) find a positive effect of more stringent regulations on total R&D expenditures. However, Jaffe and Palmer (1997) test and fail to find any effect on actual patents. Essentially re-running the former study for environmental patents Brunnermeier and Cohen (2003) find a small but significantly positive effect. In a similar vein, Kneller and Manderson (2012), find that in UK manufacturing industries there is a positive relation between the stringency of environmental regulation and environmental R&D expenditure, but not total R&D expenditure.
- On a macro, cross-country level, some weak evidence of tighter environmental regulation on environmental innovation has been documented (Lanjouw and Mody, 1996; Popp, 2006; De Vries and Withagen, 2005).⁵

None of these studies looks into spillover effects across firms and industries – for instance the fact that more stringent environmental regulation may spur innovation in industries or firms not directly affected by the regulations, but result in improved “general-purpose” technologies (e.g. in information and communication technologies, electronics, chemicals or pollution abatement equipment and services). Furthermore, new evidence suggests that previous studies may have underestimated the effects on environmentally-friendly innovation by focusing only on domestic effects. Environmental policies can also promote foreign innovation (Dechezleprêtre and Glachant, 2014) and technology transfer (e.g. in the automobile sector, Dechezleprêtre et al., 2013).

Overall, the general tendency to find positive effects on innovation contrasts somewhat with the dimmer picture when looking at actual productivity outcomes, as in the preceding section. This may be linked to the inherent challenges of actually measuring innovation, often proxied with either an input based (R&D spending) or output based (patent counts) measure. Either type of measure is however imperfectly related to overall innovative activity and even less perfectly with resulting improvements in productivity. A further possible explanation could be linked to a tendency of firms to increase environmental R&D budgets at the expense of general R&D. Others could include the decreasing returns from the increased environmental innovation (as in Hamamoto, 2006; Yang et al., 2012). Finally, some of the effects could result from flaws of using self-reporting measures, such as PACE as proxy for environmental policy stringency. Purely technically, the time dimension may play a role – an increase in R&D capital or personnel, even if potentially increasing productivity in the future – can be expected to decrease measured

productivity contemporaneously (Box 1). While there is little evidence of actual increases in R&D budgets, structural changes in the R&D direction may still imply lags in terms of productivity.

3.3. Effects on capital stock and investment

As environmental regulation often induces investment into pollution abatement capital (shown empirically by e.g. Gollop and Roberts, 1983; Berman and Bui, 2001a), theory suggests that this “abatement investment” might crowd out “non-environmental investment” into productive capital, potentially putting future productivity at risk. The empirical evidence on a possible crowding-out effect of investment is mixed. As Jaffe et al. (1995) conclude in their literature survey, there is no clear empirical evidence for a decrease in total investment due to necessary pollution abatement investment. Some evidence of crowding out has been observed (Rose, 1983; Gray and Shadbegian, 1998, 2003), while other studies found no effects (Kneller and Manderson, 2012). Still, the bulk of studies include at most a one year lag, failing to account for the fact that productivity-relevant effects on capital investment are likely to be of a longer-term nature.

Stricter environmental policies can also bring forward the obsolescence of existing capital. On the other hand, they may decrease entry (and exit), lengthen the life of polluting capital and actually delay investment (Heyes, 2014). There are only very few studies that analyse the effect of environmental regulation on the age of capital stock. The “modernising effect” of the capital stock (Xepapades and de Zeeuw, 1999) might be triggered by a downsizing of the capital stock or by investment into new machines, as confirmed by Hamamoto (2006) who finds a decrease in the average age of the capital stock in Japanese manufacturing in response to more stringent policies. On the other hand Nelson et al. (1993) find an increase in the age of capital stock, meaning a reduction in reinvestment in capital, in a sample of electric utilities in the United States. The different results may be linked to the market structure in the sectors examined and specific aspects of environmental policies, in particular if vintage differentiated laws are adopted – as for instance in the case of the New Source Review of the US Clean Air Act.

3.4. Evaluating the effects of policy instrument flexibility

More flexible environmental policy instruments are, under certain conditions, seen as achieving environmental goals in combination with superior economic outcomes due to their static and dynamic efficiency properties (see De Serres et al., 2011 for a review). Hence, the so-called “narrow” version interpretation of the Porter Hypothesis, stipulates that more flexible, market-based instruments will stimulate more innovation than less flexible command-and-control policies (Box 3). Few studies allow for testing such a hypothesis – most focus on either a single policy change or on a measure of policies that does not allow distinction between different instruments.

As a result, only a handful of studies look at whether more flexible instruments cause more environmental innovation. Both Johnstone and Labonne (2006) and Lanoie et al. (2011) find marginal and rather fragile evidence for performance standards yielding more innovation, in the latter case with respect to the less flexible technological standards. At the same time, they find no effect of taxes with respect to other instruments. Johnstone et al. (2010b) find some evidence of perceived flexibility of environmental policies increasing patenting behaviour when applying a factor analysis to extract the variation in cross-country perceived flexibility, which tends to be highly correlated with stringency.

Arimura et al. (2007) do not find any different effects of voluntary approaches relative to command-and-control measures. Possible explanations for the above findings are the difficulty of assessing the actual flexibility of an instrument and omitted variables. The fact that different types of policy instruments are effective for stimulating innovation activity in different sectors might also lead to contradicting results of studies. In the renewable energy sector for example, broad policies, such as tradable energy certificates are more effective in more mature technologies, while more targeted subsidies, such as feed-in tariffs are more effective in case of more costly technologies (Johnstone et al., 2010a).

The only study that looks at overall innovation effects, Hamamoto (2006) claims that in Japan in the 1960s and 1970s command-and-control policies did trigger higher overall R&D activity but that a SO_x charge, which was introduced later, did not.

4. Challenges for empirical implementation

Neither the choice of the set-up nor the identification strategy of empirical analyses of the economic effects of environmental policies is straightforward. The choice of the level of aggregation determines which effects can and cannot be captured in the analysis. The choice of variables and datasets are equally challenging, in particular with respect to the timing of effects. Analyses of the effect on innovation face the problem of measuring the latter through proxies. A clear, basic challenge confronting all studies is the measurement of the policy variable – that is finding a suitable, measurable variable for the stringency of environmental regulation.

4.1. Measuring environmental policies – various approaches to the explanatory variable

Estimating the effect of environmental policies often requires a proxy variable for the stringency of the regulation. In a comprehensive review of available measures, Brunel and Levinson (2013) list multidimensionality and simultaneity among the main issues to be solved in such measurement. In particular, multidimensionality represents the challenge to summarise – in a meaningful and comparable way – the potentially available information on stringency across different pollutants (with different short and long-term effects, local and national or global in nature); different environmental domains or media; and a multitude of instruments (often industry, pollutant, location and vintage specific) in the context of countries with very different industrial structures and geographical characteristics. Simultaneity can be seen as a general problem – affecting both the measurement of policies itself and the actual empirical applications of the policy stringency variable. The latter is discussed in the section on identification below, but the former problem affects a number of measures used in the literature – particularly those based on outcomes of policies, such as spending or emissions, rather than actual policy characteristics. In this case simultaneity arises from the fact that measured environmental policy stringency cannot distinguish the effect of actual environmental policy stringency from that of other contributors to the measured outcome, such as non-environmental regulations; market imperfections; levels of income, skills and technology; capital intensity and pollution levels. While various attempts at measuring environmental policy stringency have been used in the past, none addresses the above considerations in a satisfactory manner.

There are several dominant approaches to measuring environmental policies in the empirical literature (Appendix A, Table A.1):

- One of the most popular is the use of survey-based firm or plant-level expenditures, which are interpreted as induced by environmental rules and hence intended to proxy for their stringency. The US Pollution Abatement Costs and Expenditures (PACE) and more generally environmental protection expenditures (Box 4) are widely used examples (Gray, 1987, Morgenstern et al., 2002; Gray and Shadbegian, 2003). This measure suffers from poor comparability across countries and time, problems of counterfactuals, or defining what part of expenditures are driven by environmental policies, the self-reporting nature of the surveys and sampling. The use of such proxies may pose practical problems in the case of attempts to assess the Porter Hypothesis – looking at the effects on environmental innovation may preclude using expenditures on innovative, pollution-reducing technologies as the explanatory variable – i.e. as proxy for environmental policies. Moreover, they are to some degree endogenous – higher compliance costs may come from older plants and technologies, rather than stricter policies (Ambec et al., 2013).
- As an alternative, “shadow prices” of pollution are proposed (reviewed in more detail by Brunel and Levinson, 2013). These stem from the idea that environmental policies put a price on pollution, hence the more stringent the policies, the higher the price faced by polluters, affecting their optimisation decisions. The main appeal of this approach is the fact that it focuses on the estimated “cost” of pollution faced by the firm, regardless of the instrument mix in place, circumventing the problem of multidimensionality. However, among the main problems are simultaneity (estimated shadow prices are also affected by other factors) and strong reliance on estimation assumptions (for example, regarding the functional form of the cost or production functions).
- A number of studies have used environmental or related performance data as a measure of policy stringency. Examples include energy intensity (Cole and Elliot, 2003; Van Beers and Van den Bergh, 1997; Harris et al., 2003), state compliance with environmental standards (McConnell and Schwab, 1990), and pollutant emission intensity (Smarzynska and Wei, 2004). Brunel and Levinson (2013) themselves follow this approach, proposing a measure of actual pollution intensity relative to what could be expected given the country’s industrial structure. While interesting, particularly in cross-country empirical applications such measures tend to suffer from simultaneity and possible reverse causality as these measures might be influenced by e.g. differences in factor prices, technology and industrial structure more than by environmental regulation.
- Perceptions of the stringency of laws and their enforcement, as compiled by the World Economic Forum (WEF),⁶ were used by Kalamova and Johnstone (2011) and Johnstone et al. (2010b). Similarly, Johnstone and Labonne (2006) and Lanoie et al. (2011), use survey based company level perceptions of the stringency of environmental regulations.⁷ While potentially reflecting what matters – that is the stringency perceived by those who are affected – such measures have several downsides: they are context-specific (e.g. depending on the business cycle), linked to respondents subjectivity, problematic in international and inter-temporal comparability, and prone to the state of the economy (cyclicality) and sampling bias (for example, reflecting only the “incumbents” views).
- Some studies, in particular those based on difference-in-difference estimations, use a more event-based approach, based on the event of introduction or significant change in

a particular policy (Berman and Bui, 2001b; Van der Vlist et al., 2007; Curtis, 2012). While these approaches may better capture causal relationships in the case of very specific effects of individual policy changes, this will often be at the sacrifice of the generality of conclusions. Moreover, by measuring *de jure* aspects of environmental policies, they may omit implementation details and the enforcement of the environmental regulation.

- Others have experimented with broader policy proxies. Examples include the ratifications of international environmental treaties (counts, as in Smarzynska and Wei, 2004; or incidence, as in Yörük and Zaim, 2005; or Wu and Wang, 2008) or policy instruments in a given environmental domain (Jaraité and Di Maria, 2012). These measures apply only to a specific industry or pollutant dimension, several other approaches have attempted to summarise information in a number of “representative” industry and pollutant dimensions to create a more general index of environmental policy stringency (Berman and Bui, 2001a; Fleishman et al., 2009). These may again suffer from a selective approach, possibly omitting crucial areas, and in an international context often from aggregation issues when quantifying and compiling the stringency across very different instruments, pollutants and industrial structures. They also focus on the *de jure* rather than the *de facto* stringency of environmental policies.
- The spectrum of other proxies used in empirical studies is wide, including congressional pro-environment voting records of state representatives (Gray, 1997) or environment-related inspection frequency (Alpay et al., 2002; Testa et al., 2011; Brunnermeier and Cohen, 2003).

Overall, the choice of the most appropriate variables is likely to depend on data availability and the nature of the exercise – for example the type of pollution in question or the type of cross section units. When investigating the effects of several environmental domains or instruments, empirical analyses could either use a set of separate measures or a composite indicator, which summarises information across a number of the main dimensions of environmental policies. To the extent environmental policy characteristics can be approximated by a selected number of policy instruments, composite indicators can prove very powerful in empirical analysis. At the same time, a set of separate, non-aggregated measures may more directly address issues of trade-offs and complementarities across policies, instruments and design features, but may prove problematic due to multicollinearity, a loss of degrees of freedom, and, inevitably, the inability to cover all potentially relevant dimensions. In the context of international analysis, where comparable data is needed, the use of a composite index seems rather underexploited. This may be a consequence of poor data availability, and potential risks related to the choice of areas and quantification and aggregation of gathered information. Nonetheless, this omission is striking given that similar attempts have taken place in areas of significant complexity and multidimensionality, such as Product Market Regulation or Employment Protection Legislation (Nicoletti et al., 2000).

4.2. Time horizon matters but longer-term effects are difficult to capture

The time dimension of the economic effects of environmental policy may depend on the specific policy design and its announcement. On the one hand impacts may precede the date of introduction of the policy – firms can decide to make investments or change technologies upon announcement or even in anticipation of a new environmental rule. This argument however, relies strongly on the assumption that the policy change is credible and announced well in advance, for firms to take early action. Such a situation can

arise due to international obligations – which require policy action by a certain date (frequent case in the EU) or due to a general perception of tightening of environmental rules, or a stable commitment to environmental policies – for instance Johnstone et al. (2010b) find a significant role of environmental policy predictability in triggering investment into renewable energy innovation. This might reflect the more general argument that in the presence of higher uncertainty of future laws firms may delay investments and adoption of new technologies (Shadbegian and Gray, 2005). On the other hand, impacts may take time – if policies trigger higher R&D investment, it may take years to actually bring about measurable improvements in technology and processes. In this vein, Jaffe and Palmer (1997) find a positive effect of pollution control expenditures on R&D expenditure in the manufacturing sector in the United States in 1975–1991, yet no significant effect on the number of successful patents.

By sheer nature, short-term and long-term effects may differ. If a new policy induces lump-sum investment, be it into capital (necessitated directly for pollution reduction or due to an accelerated scrapping rate as old “dirty” capital becomes obsolete), hiring of new personnel (to cope with environmental requirements or as a change of technology) or increased R&D expenditure, short-term inputs will increase without an equivalent increase in output. This would show up in curbed productivity growth. Thereafter, rebound effects or the benefits of new technologies may imply temporarily higher productivity growth – even assuming away the strong version of Porter’s Hypothesis. While these effects may eventually fade out, with productivity growth returning back to trend, most empirical studies focusing on short-term or even contemporaneous effects will have a problem capturing the dynamics of such developments. Further complications may arise from the fact that less environmentally-efficient companies may go bankrupt as a result of regulation, while new companies may enter the market. Such effects are rarely taken into account in empirical research. Finally, trade and regulation avoidance may complicate the picture further, as for instance firms may tend to reallocate production and investment to plants subject to more lax environmental rules among US states (Gray and Shadbegian, 1998), or commence activity in States with less stringent legislation (Gray, 1997).

The focus on contemporaneous and short-term effects in empirical literature is driven by data availability and attempts to capture some longer-term effects have been fairly modest. Notable efforts, for example, Managi et al. (2005) include up to eight years of the lagged policy variable in their analysis of productivity growth in the offshore oil and gas industry in the United States. They find a negative contemporaneous effect. After the second year, the regulation’s effect on productivity becomes positive. The accumulated effect outweighs the initial negative impact after the fourth year. The subsequent lags are also significantly positive. Lanoie et al. (2008) provide similar evidence for the Canadian manufacturing sector: the negative contemporaneous effect on productivity growth is outweighed by the accumulated effect after three years. Similarly, for the Taiwanese manufacturing sector Yang et al. (2012) find a small but significantly positive effect of the second year lag, but do not include further lags due to the short sample. Hamamoto (2006) tests different lag-structures in several models and shows that the specification of the six year lag fits better in terms of explained variability than for example a three-year lag structure. In a study of Swedish manufacturing industries, Broberg et al. (2012) do not find a significant longer-term effect. Only when looking explicitly at the pulp and paper industry, does a significant negative effect appear with a two year lag. Greenstone et al.

(2012) investigate the manufacturing sector in the United States and find a negative dynamic effect as well.

Much of the research discussed above takes a fairly static approach to environmental policies – one looking for a single, unique effect of a unique policy change. However, as discussed in De Serres et al. (2010), the nature of incentives provided by policies may be more dynamic. For instance, market-based instruments tend to provide more persistent incentives to innovate and to increase productivity than other policy instruments. These effects can be even stronger if policies have longer time horizons – for example when tax rates or caps on carbon emissions are (credibly) announced for years to come (such as in the case of the EU ETS), or when standards depending on best available technologies are to be revised periodically. Such characteristics of environmental policies are rarely straightforward to capture through standard measures of stringency and hence are usually ignored in empirical analysis.

4.3. Studies at different levels of aggregation may yield complementary insight

In the case of firm- or plant-level studies most standard data sets are fairly limited – they offer a short time series and often focus on firms in particular sectors and countries, hence limiting the generality of the investigation. Such datasets generally cover only incumbent firms over the sample period, leading to a disregard for entry and exit, which may reflect precisely the effects of environmental policies. In practice, very few studies actually attempt to adjust their estimates for entry and exit of firms using alternative data sources. Greenstone et al. (2012) show that the negative estimated effects of environmental policies on productivity are actually larger when adjusted for the fact that most inefficient firms drop out of the sample as a result of the environmental laws. For employment outcomes, Berman and Bui (2001b) find, however, no effect of taking account of dissuaded entry and induced exit. Finally, studies do not investigate network and spillover effects, such as those of tighter policies in a given sector which might spur innovation and productivity growth in other sectors.

Industry-level studies potentially provide longer time-series and are more suitable for dealing with entry and exit, but still usually ignore network and spillover effects. The level of aggregation of industries may matter – some effects may be missed at higher levels of aggregation, while lower levels will suffer from lack of data. Similarly, industry-level analysis will also miss the costs of labour substitution or movement from one plant to another, when total industry employment remains stable (Morgenstern et al., 2002). Furthermore, many studies tend to focus on a subset of industries only (e.g. Barbera and McConnell, 1990; Hamamoto, 2006; Yang et al., 2012).

The macroeconomic level approach potentially deals with some of the above issues by capturing the overall effects on productivity. They also offer more policy instrument variation than present in a single sector and often longer time series. These advantages come at a cost – measurement and comparison of policies may be more cumbersome – requiring to summarise multidimensional information, as discussed before. Moreover, the identification of effects of policy changes becomes difficult, and requires controlling for a wide range of other variables.

4.4. Identification of the true effect – a challenging task

Environmental policies are rarely introduced in isolation, making it difficult to assess their individual effects. Environmental policies can interact with each other and are often

accompanied by mitigating measures, designed to soften perceived adverse impacts; by promoting environmental action in other countries; or even by border tax adjustments (OECD, 2010). The simultaneous implementation of related measures makes it difficult to properly identify the initial effect and to attribute an observed economic outcome to the specific policy under consideration. Moreover, many measures are temporary and eventually reversed, which due to time lags in their implementation and effects poses a further challenge for the identification of actual effects.

The comparability across studies suffers from the fact that different types of policy instruments as such can have different economic effects. The stringency of an appropriately designed performance-based standard for example might rely on the best available technologies, and will hence increase as more environmentally efficient technologies develop. On the other hand, the effective stringency of a fixed pollution cap may decrease over time for the very same reason. This makes it difficult to compare studies analysing different policy approaches.

One additional complication for empirical analysis is the potential reverse causality, that is, the extent to which the stringency of environmental policies is driven by productivity growth, causing practical problems for estimation such as biased estimates (Gray, 1987). As economic growth can spur demand for environmental quality (environmental Kuznets curve), stricter environmental laws might be a response of policy makers to this increasing demand. Similarly, environmental policies create new rent opportunities and as such can incentivise lobbying for more stringent laws. In an attempt to deal with issues of this type Managi et al. (2005) use a Granger causality test to investigate if environmental stringency is affected by productivity growth in the offshore oil and gas industry, and cannot reject non-causality. A recent study by Abdullah and Morley (2014) investigates this reverse causality problem on the macro-economic level by looking at environmental tax revenues and GDP, finding some weak evidence of a long-run causal effect from GDP to environmental tax revenues.

5. Conclusion

Empirical research on the productivity effects of environmental policies is largely inconclusive. Results are usually very context-specific and hence can only provide limited general policy conclusions – raising the question to what extent the results from a specific policy change, aimed at a particular pollutant and industry in a given country, can be generalised. The fairly broad support for the weak version of the Porter Hypothesis is not very surprising. However, the findings of an ambiguous effect of environmental policies on productivity, in line with the strong Porter Hypothesis, are rather unexpected, considering the earlier research aimed at explaining the US productivity slowdown with the tightening of environmental regulation.

Finding significant effects of environmental policy changes may be hard because environmental compliance costs are usually only a small share of total costs (Gray and Shadbegian, 2003). The size of the effects that are found in different studies reviewed is hardly comparable, mainly due to the crudeness of the environmental policy proxy variable.

Firm-level difference-in-difference studies are methodologically the most convincing approach. These studies allow clear identification of the economic effects of environmental policies, which helps to understand the forces at work on a microeconomic

level. However, the very same argument implies that these studies suffer from lack of generality. Most studies also have problems dealing with a selection bias, dynamics (short samples), sample selection (entry/exit) or network effects/spillovers. Moreover, given that most studies only look at one type of policy instrument, they are also poorly suited to assess the potentially different effects of market-based and command-and-control policies. At the industry-level, generality is also dubious, and the bulk of older studies tend to suffer from the simultaneity problem. Spillovers across industries are left unaccounted for. While in practice only DEA and SFA studies control for environmental outcomes, the use of environmentally-adjusted productivity *per se* is asking a different question – no longer focusing on pure economic performance but implicitly weighting it against environmental performance (Boxes 1 and 2).

In terms of methodology, several general ideas for future research can be drawn:

- First, adding an international dimension can increase the variation both across policies and across outcomes, providing a richer sample. This could possibly reduce the need for a longer time series. So far, studies predominantly tend to look at national effects of national policies. With respect to many of the studies, aside perhaps from some of those conducted across US states, this additional degree of variation, both in policies and in outcomes, can allow for better identification of effects, hence easing some of the simultaneity issues. Such a comprehensive approach can also prove more suitable in comparing effects of market-based and command-and-control instruments, or other pertinent aspects of policy design, such as imposed administrative burdens, barriers to entry, exit and competition (Koźluk, 2014).
- Second, combining various levels of aggregation, such as macro, industry and firm-level, can yield complementary insights. Most studies look at one single level, while as discussed above, effects can differ significantly on each of the levels, due to considerations such as entry/exit, and spillover and network effects, and international trade. Looking at the three levels can help reconcile results that seem to be conflicting.
- Third, the utilisation of composite measures of environmental policies in cross-country comparisons can provide interesting insight on the effects of environmental policies, and their specific design characteristics – such as stringency and flexibility. The use of such composite indexes is rather scarce so far. This can partly be explained by the prevailing interest in national level studies, where effects of policies might better be analysed using more direct measures, such as discrete changes in specific policies. In the context of international comparisons, where different countries tend to use a multitude of different tools (Botta and Koźluk, 2014), an effort to construct such a composite index can provide a basis to analyse the effects of environmental regulations. Sub-indicators, by type of policy instrument or design characteristic, can also allow differentiating among the effects, for instance of market-based instruments versus command-and-control regulations. Practical challenges in constructing such composite indexes are surely also relevant, for example the complexity of multi-dimensional environmental policies and the data gathering process. Tackling these challenges, a composite index can provide policymakers with more insight on the choice among different policy approaches available.

Notes

1. Hicks (1932), p. 124: “A change in the relative prices of the factors of production is itself a spur to invention, and to invention of a particular kind - directed to economising the use of a factor that has become relatively expensive”.
2. This often omitted point is also made by Romstad (1998).
3. An important point in this context is that many studies look at a fairly short-term perspective, where the link between productivity and competitiveness is not as evident.
4. The terms “environmental” innovation and “environmental” R&D, used commonly in the literature, are often defined fairly loosely. The approach taken in this review is to use them very broadly, in relation to innovative and R&D activity that focuses on reducing (some) environment-related effects, such as pollution or use of natural resources.
5. In De Vries and Withagen (2005) two out of the three measures of environmental regulation do not have any significant effect on environmental innovation. Their preferred model, which uses an instrumental variable approach, shows a significant and rather large positive effect on the number of environmental patents, they fail to provide any details on the validity of the instrumental variable approach.
6. The WEF “Executive Opinion Survey”, conducted annually, asks respondents (business executives) a number of questions related to environmental policy design. The questions, coverage and sampling have varied across the years, but the most common questions were to assess the “stringency” (and “enforcement”) of the overall environmental regulation in the country of operation, on a 1 to 7 scale. The survey was implemented by the WEF’s partner institutes in over 150 economies. In most years, there were responses from between 8 000 and 15 000 firms (see WEF, 2013; for a description of the sampling strategy.) http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2012-13.pdf
7. The 2003 survey data were collected for the manufacturing sector in seven OECD countries and include information on environmental R&D expenditure, environmental and commercial performance and perceived stringency of environmental regulation.

References

- Abdullah, S. and B. Morley (2014), “Environmental Taxes and Economic Growth: Evidence from Panel Causality Tests”, *Energy Economics*, Vol. 42, pp. 27-33.
- Aghion, P., R. Blundell, R. Griffith, P. Howitt and S. Prantl (2004), “Entry and Productivity Growth: Evidence From Microlevel Panel Data”, *Journal of the European Economic Association*, Vol. 2(2-3), pp. 265-276.
- Aghion, P., N. Bloom, R. Blundell, R. Griffith, P. Howitt (2005), “Competition and Innovation: An Inverted-U Relationship”, *Quarterly Journal of Economics*, Vol. 120(2), pp. 701-728.
- Ambec, S., M.A. Cohen, S. Elgie and P. Lanoie (2013), “The Porter Hypothesis at 20: Can Environmental Regulation Enhance Innovation and Competitiveness?”, *Review of Environmental Economics and Policy*, Vol. 7(1), pp. 2-22.
- Alpay, E., S. Buccola and J. Kerkvliet (2002), “Productivity Growth and Environmental Regulation in Mexican and US Food Manufacturing”, *American Journal of Agricultural Economics*, Vol. 84(4), pp. 887-901.
- Arimura, T., A. Hibiki and N. Johnstone (2007), “An Empirical Study of Environmental R&D: What Encourages Facilities to be Environmentally Innovative?”, in *Corporate Behaviour and Environmental Policy*, ed., N. Johnstone, Cheltenham, UK.
- Barbera, A.J. and V.D. McConnell (1990), “The Impact of Environmental Regulations on Industry Productivity: Direct and Indirect Effects”, *Journal of Environmental Economics and Management*, Vol. 18, pp. 50-65.
- Becker, R.A. (2011), “Local Environmental Regulation and Plant-level Productivity”, *Ecological Economics*, Vol. 70, pp. 2513-2522.
- Becker R.A. and R.J. Shadbegian (2004), “Comparing the 1994 and 1999 Pollution Abatement Costs and Expenditures Surveys”, *NCEE Working Paper*, No. 04-08.
- Berman, E. and L.T.M. Bui (2001a), “Environmental Regulation and Productivity: Evidence from Oil Refineries”, *The Review of Economics and Statistics*, Vol. 83(3), pp. 498-510.

- Berman, E. and L.T.M. Bui (2001b), "Environmental Regulation and Labor Demand: Evidence from the South Coast Air Basin", *Journal of Public Economics*, Vol. 79, pp. 265-295.
- Botta, E. and T. Koźluk, (2014), "Measuring Environmental Policy Stringency in OECD Countries: A Composite Index Approach", *OECD Economics Department Working Papers*, No. 1177, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5jxrjnc45gvg-en>.
- Bourlès, R., G. Cette, J. Lopez, J. Mairesse and G. Nicoletti (2013), "Do Product Market Regulations In Upstream Sectors Curb Productivity Growth? Panel Data Evidence For OECD Countries", *Review of Economics and Statistics*, Vol. 95(5), pp. 1750–1768.
- Boyd, G.A. and J.D. McClelland (1999), "The Impact of Environmental Constraints on Productivity Improvement in Integrated Paper Plants", *Journal of Environmental Economics and Management*, Vol. 38, pp. 121-142.
- Brandt, N., P. Schreyer and V. Zipperer (2014), "Productivity Measurement with Natural Capital and Bad Outputs", *OECD Economics Department Working Papers*, No. 1154, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5jz0wh5t0ztd-en>.
- Brännlund, R. and T. Lundgren (2009), "Environmental Policy without Costs? A Review of the Porter Hypothesis", *International Review of Environmental and Resource Economics*, Vol. 3, pp. 75-117.
- Broberg, T., P.-O. Marklund, E. Samakovlis and H. Hammar (2012), "Testing the Porter Hypothesis: The Effects of Environmental Investments on Efficiency in Swedish Industry", *Journal of Productivity Analysis*, Vol. 40(1), pp. 43-56.
- Brunel, C. and A. Levinson (2013), "Measuring Environmental Regulatory Stringency", *Working Paper*, Georgetown University.
- Brunnermeier, S.B. and M.A. Cohen (2003), "Determinants of Environmental Innovation in US Manufacturing Industries", *Journal of Environmental Economics and Management*, Vol. 45, pp. 278-293.
- Brunnermeier, S.G. and A. Levinson (2004), "Examining the Evidence on Environmental Regulation and Industry Location", *Journal of Environment and Development*, Vol. 13(6), pp. 6-41.
- Christainsen, G.B. and R.H. Haveman (1981), "Public Regulations and the Slowdown in Productivity Growth", *The American Economic Review*, Vol. 71(2), pp. 320-325.
- Cole, M.A. and R.J. Elliot (2003), "Do Environmental Regulations Influence Trade Patterns? Testing Old and New Trade Theories", *The World Economy*, Vol. 26(8), pp. 1163-1186.
- Copeland, B.R. and M.S. Taylor (2004), "Trade, Growth and the Environment", *Journal of Economic Literature*, Vol. 42, pp. 7-71.
- Curtis, M. (2012), "Who Loses under Power Plant Cap-and-Trade Programs? Estimating the Impact of the NO_x Budget Trading Program on Manufacturing Employment", *Working Paper Georgia State University*, Georgia.
- Dechezleprêtre, A. and M. Glachant (2014), "Does Foreign Environmental Policy influence Domestic Innovation? Evidence from the Wind Industry", *Environmental and Resource Economics*, No. 58, pp. 391-413, <http://dx.doi.org/10.1007/s10640-013-9705-4>.
- Dechezleprêtre, A., E. Neumayer, R. Perkins (2013), "Environmental Regulation and the Cross-border Diffusion of New Technology: Evidence from Automobile Patents", *FEEM Working Paper*, No. 33.
- De Serres, A., F. Murtin and G. Nicoletti (2010), "A Framework for Assessing Green Growth Policies", *OECD Economics Department Working Papers*, No. 774, www.oecd-ilibrary.org/economics/summary-indicators-of-product-market-regulation-with-an-extension-to-employment-protection-legislation_215182844604, OECD Publishing.
- De Vries, F.P. and C. Withagen (2005), "Innovation and Environmental Stringency: The Case of Sulphur Dioxide Abatement", *Discussion Paper Tilburg University*, Tilburg.
- Domazlicky, B.R. and W.L. Weber (2004), "Does Environmental Protection Lead to Slower Productivity Growth in the Chemical Industry?", *Environmental and Resource Economics*, Vol. 28, pp. 301-324.
- Dufour, C., P. Lanoie and M. Patry (1998), "Regulation and Productivity", *Journal of Productivity Analysis*, Vol. 9(3), pp. 233-247.
- Fleishman, R., R. Alexander, S. Bretschneider and D. Popp (2009), "Does Regulation Stimulate Productivity? The Effect of Air Quality Policies on the Efficiency of US Power Plants", *Energy Policy*, Vol. 37, pp. 4574-4582.

- Gollop, F.M. and M.J. Roberts (1983), "Environmental Regulations and Productivity Growth: The Case of Fossil-fuelled Electric Power Generation", *Journal of Political Economy*, Vol. 91(4), pp. 654-674.
- Graff Zivin, J. and M. Neidel (2012), "The Impact of Pollution on Worker Productivity", *American Economic Review*, Vol. 102(7), pp. 3652-73.
- Gray, W.B. (1987), "The Cost of Regulation: OSHA, EPA and the Productivity Slowdown", *The American Economic Review*, Vol. 77(5), pp. 998-1006.
- Gray, W.B. (1997), "Manufacturing Plant Location: Does State Pollution Regulation Matter?", *NBER Working Paper Series 5880*, National Bureau of Economic Research, Cambridge Massachusetts.
- Gray, W.B. and R.J. Shadbegian (1998), "Environmental Regulation, Investment Timing, and Technology Choice", *The Journal of Industrial Economics*, Vol. 46, pp. 235-256.
- Gray, W.B. and R.J. Shadbegian (2003), "Plant Vintage, Technology, and Environmental Regulation", *Journal of Environmental Economics and Management*, Vol. 46, pp. 384-402.
- Gray, W.B. and R.J. Shadbegian (2005), "Pollution Abatement Expenditures and Plant-level Productivity: A Production Function Approach", *Ecological Economics*, Vol. 54, pp. 196-208.
- Greenstone, M., J. List and C. Syverson (2012), "The Effects of Environmental Regulation on the Competitiveness of US Manufacturing", *NBER Working Paper Series 18392*, National Bureau of Economic Research, Cambridge Massachusetts.
- Hamamoto, M. (2008), "Environmental Regulation and the Productivity of Japanese Manufacturing Industries", *Resource and Energy Economics*, Vol. 28, pp. 299-312.
- Harris, M., L. Konya and L. Matyas (2003), "Modelling the Impact of Environmental Regulations on Bilateral Trade Flows: OECD, 1990-1996", *World Economy*, Vol. 25(3), pp. 387-405.
- Heyes (2014), "Path Dependence, Firm Entry-Exit and the Environment", OECD Publishing, forthcoming.
- Hicks, J.R. (1932), *The Theory of Wages*, McMilland & Co, London.
- Jaffe, A.B., S.R. Peterson, P.R. Portney and R.N. Stavins (1995), "Environmental Regulation and the Competitiveness of US Manufacturing: What Does the Evidence Tell Us?", *Journal of Economic Literature*, Vol. 33(1), pp. 132-163.
- Jaffe, A.B., R.G. Newell and R.N. Stavins (2003), "Technological Change and the Environment", *Handbook of Environmental Economics*, Vol. 1, pp. 461-516.
- Jaffe, A.B. and K. Palmer (1997), "Environmental Regulation and Innovation: A Panel Data Study", *The Review of Economics and Statistics*, Vol. 79(4), pp. 610-619.
- Jaraité, J. and C. Di Maria (2012), "Efficiency, Productivity and Environmental Policy: A Case Study of Power Generation in the EU", *Energy Economics*, Vol. 34, pp. 1557-1568.
- Jeon, B.M. and R.C. Sickles (2004), "The Role of Environmental Factors in Growth Accounting", *Journal of Applied Econometrics*, Vol. 19, pp. 567-591.
- Johnstone, N. and J. Labonne (2006), "Environmental Policy, Management and R&D", *OECD Economic Studies*, Vol. 42, pp. 169-203, www.oecd.org/eco/greeneco/38698195.pdf, OECD Publishing.
- Johnstone, N., I. Haščič and D. Popp (2010a), "Renewable Energy Policies and Technological Innovation: Evidence Based on Patent Counts", *Environmental and Resource Economics*, Vol. 45, pp. 133-155.
- Johnstone, N., I. Haščič and M. Kalamova (2010b), "Environmental Policy Design Characteristics and Technological Innovation: Evidence from Patent Data", *OECD Environment Working Papers*, No. 16, www.oecd-ilibrary.org/content/workingpaper/5kmjstwtqwhd-en, OECD Publishing.
- Kalamova, M. and N. Johnstone (2011), "Environmental Policy Stringency and Foreign Direct Investment", *OECD Environment Working Papers*, No. 33, www.oecd-ilibrary.org/content/workingpaper/5kg8ghvf85d5-en, OECD Publishing.
- Koźluk, T. (2014), "The Indicator of Economic Burdens of Environmental Policy Design – Results from the OECD Questionnaire", *OECD Economics Department Working Papers*, No. 1178, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5jxrnbnbm8v-en>.
- Kneller, R. and E. Manderson (2012), "Environmental Regulations and Innovation Activity in UK Manufacturing Industries", *Resource and Energy Economics*, Vol. 34, pp. 211-235.
- Krugman, P.R. (1994), "Competitiveness – A Dangerous Obsession", *Foreign Affairs*, Vol. 73 (2), pp. 28-44.

- Lanjouw, J.O. and A. Mody (1996), "Innovation and the International Diffusion of Environmentally Responsive Technology", *Research Policy*, Vol. 25(5), pp. 549-571.
- Lanoie, P., M. Patry and R. Lajeunesse (2008), "Environmental Regulation and Productivity: Testing the Porter Hypothesis", *Journal of Productivity Analysis*, Vol. 30, pp. 121-128.
- Lanoie, P., J. Laurent-Lucchetti, N. Johnstone and S. Ambec (2011), "Environmental Policy, Innovation and Performance: New Insights on the Porter Hypothesis", *Journal of Economics and Management Strategy*, Vol. 20(3), pp. 803-842.
- Levinson, J. and A. Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies*, Vol. 70, pp. 317-341.
- Managi, S., J.J. Opaluch, D. Jin and T.A. Grigalunas (2005), "Environmental Regulations and Technological Change in the Offshore Oil and Gas Industry", *Land Economics*, Vol. 81(2), pp. 303-319.
- McConnell, V. and R. Schwab (1990), "The Impact of Environmental Regulation on Industry Location Decisions: The Motor Vehicle Industry", *Land Economics*, Vol. 66(1), pp. 67-81.
- Mohr, R.D. and S. Saha (2008), "Distribution of Environmental Costs and Benefits, Additional Distortions and the Porter Hypothesis", *Land Economics*, Vol. 84(4), pp. 689-700.
- Morgenstern, R.D., W.A. Pizer and J.-S. Shih (2002), "Jobs Versus the Environment: An Industry-level Perspective", *Journal of Environmental Economics and Management*, Vol. 43, pp. 412-436.
- Nelson, R.A., T. Tietenberg and M.R. Donihue (1993), "Differential Environmental Regulation: Effects on Electric Utility Capital Turnover and Emissions", *Review of Economics and Statistics*, Vol. 75(2), pp. 368-373.
- Nicoletti, G., S. Scarpetta and O. Boylaud (2000), "Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation", *OECD Economics Department Working Papers*, No. 226, www.oecd.org/eco/OECD-Economics-Department-Working-Papers-by-year.pdf.
- OECD (2006), "Environmental Regulation and Competition – Policy Roundtable", DAF/COMP(2006)30.
- OECD (2007), "Pollution Abatement and Control Expenditure in OECD Countries", *Environmental Policy Committee, Working Group on Environmental Information and Outlooks ENV/EPOC/SE(2007)1*.
- OECD (2010), "Linkages between Environmental Policy and Competitiveness", *OECD Environment Working Papers*, No. 13, www.oecd-ilibrary.org/environment/linkages-between-environmental-policy-and-competitiveness_218446820583, OECD Publishing.
- Ostro, B.D. (1983), "The Effects of Air Pollution on Work Loss and Morbidity", *Journal of Environmental Economics and Management*, Vol. 10(4), pp. 371-82.
- Popp, D. (2006), "International Innovation and Diffusion of Air Pollution Control Technologies: The Effects of NO_x and SO₂ Regulation in the United States, Japan, and Germany", *Journal of Environmental Economics and Management*, Vol. 51, pp. 46-71.
- Porter, M. (1991), "America's Green Strategy", *Scientific American*, Vol. 264(4), p. 168.
- Porter, M.E. and C. van der Linde (1995), "Toward a New Conception of the Environment Competitiveness Relationship", *Journal of Economic Perspectives*, Vol. 9(4), pp. 97-118.
- Repetto, R., D. Rothman, P. Faeth and D. Austin (1997), "Has Environmental Protection Really Reduced Productivity", *Challenge*, Vol. 40(1), pp. 46-57.
- Romstad, E. (1998), "Environmental Regulation and Competitiveness", in: Barker, T. & Koehler, J. (eds.), *International Competitiveness and Environmental Policies*, pp. 185-196, Edward Elgar Publishers, Aldershot, England.
- Rose, A. (1983), "Modelling the Macroeconomic Impact of Air Pollution Abatement", *Journal of Regional Science*, Vol. 23(4), pp. 441-459.
- Sauvage, J. (2014), "The Stringency of Environmental Regulations and Trade in Environmental Goods and Services", *OECD Trade and Environment Working Papers*, No. 2014/03, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5jxrn7xsnmq-en>.
- Shadbegian, R.J. and W. B. Gray (2005), "Pollution Abatement Expenditures and Plant-level Productivity: A Production Function Approach", *Ecological Economics*, Vol. 54, pp. 196-208.
- Smarzynska, B.J. and S.J. Wei (2004), "Pollution Havens and Foreign Direct Investment: Dirty Secret or Popular Myth?", *The B.E. Journal of Economic Analysis and Policy*, Vol. 2, p. 8.

- Smith, J.B. and W.A. Sims (1985), "The Impact of Pollution Charges on Productivity Growth in Canadian Brewing", *The RAND Journal of Economics*, Vol. 16(3), pp. 410-423.
- Testa, F., F. Iraldo and M. Frey (2011), "The Effect of Environmental Regulation on Firms' Competitive Performance: The Case of the Building & Construction Sector in some EU Regions", *Journal of Environmental Management*, Vol. 92, pp. 2136-2144.
- Van Beers, C. and J. Van den Bergh (1997), "An Empirical Multi-Country Analysis of the Impact of Environmental Regulations on Foreign Trade Flows", *Kyklos*, Vol. 50(1), pp. 29-46.
- Van der Vlist, A.J., C. Withagen and H. Folmer (2007), "Technical Efficiency under Alternative Environmental Regulatory Regimes: The Case of Dutch Horticulture", *Ecological Economics*, Vol. 63, pp. 165-173.
- Wagner, M. (2003), "The Porter Hypothesis Revisited: A Literature Review of Theoretical Models and Empirical Tests", Research Memorandum Centre for Sustainability Management University of Lüneburg, Lüneburg Germany.
- WEF (2013), *The Global Competitiveness Report 2012-2013*, World Economic Forum, Geneva.
- Wu, Y. and B. Wang (2008), "Environmental Regulation and Total Factor Productivity Growth: An Empirical Study of the APEC Economies", *Economic Research Journal*, Vol. 5, pp. 19-33.
- Xepapades, A. and A. de Zeeuw (1999), "Environmental Policy and Competitiveness: The Porter Hypothesis and the Composition of Capital", *Journal of Environmental Economics and Management*, Vol. 37, pp. 165-182.
- Yang, C.-H., Y.-H. Tseng and C.-P. Chen (2012), "Environmental Regulations, Induced R&D, and Productivity: Evidence from Taiwan's Manufacturing Industries", *Resource and Energy Economics*, Vol. 34, pp. 514-532.
- Yörük, B.K. and O. Zaim (2005), "Productivity Growth in OECD Countries: A Comparison with Malmquist Indices", *Journal of Comparative Economics*, Vol. 33, pp. 401-420.

Appendix A

Table A.1. **Overview of empirical studies**

Author and year	Dep. variable	Independent variable	Sample	Methodology	Result
Productivity measures					
Gollop and Roberts (1983)	Δ TFP	Ratio of legal emission target to effective enforcement	American electric power industry, 1973-1979 (F)	<ul style="list-style-type: none"> compare estimated productivity and growth contributions across regulated and non-regulated plants 	<ul style="list-style-type: none"> productivity growth is reduced for restricted plants, by 0.5 percentage points per year on average evidence for increased costs due to sulphur dioxide emission restriction
Smith and Sims (1985)	Δ TFP	Payments for BOD and SS emissions	Canadian brewing industry, 1971-1980 (F)	<ul style="list-style-type: none"> compare productivity growth of regulated and non-regulated firms 	<ul style="list-style-type: none"> negative impact of pollution charges on productivity growth
Berman and Bui (2001a)	TFP	Count variable of number of regulations in place	American oil refineries (Los Angeles basin), 1977-1992 (F)	<ul style="list-style-type: none"> fixed effect estimation compare regulated and non-regulated refineries 	<ul style="list-style-type: none"> positive, insignificant effect for regulated plants no evidence for increase abatement operating costs due to regulation found
Gray and Shadbegian (2003)	TFP	PAOC	American pulp and paper mills, 1979-1990 (F)	<ul style="list-style-type: none"> fixed effect and GMM estimation of regression of TFP and PAOC and directly of production function which includes PAOC 	<ul style="list-style-type: none"> negative effect on productivity levels is driven by integrated mills; effect for non-integrated mills is negligible
Greenstone et al. (2012)	TFP	Dummy variable of attainment/non-attainment of air pollution regulations	American manufacturing sector, 1972-1993 (F)	<ul style="list-style-type: none"> include up to two lags fixed effect estimation regressing TFP on environmental policy variable 	<ul style="list-style-type: none"> overall, negative effect on MFP ozone regulations have strongest contemporaneous negative effect, PM and SO regulations strongest accumulated effect negative accumulated overall effect is larger than contemporaneous effect
Becker (2011)	Labour productivity	Ratio of PACE to economic activity	American manufacturing industries, 1980-1994 (F)	<ul style="list-style-type: none"> fixed effect estimation of Cobb-Douglas function including compliance costs 	<ul style="list-style-type: none"> no significant effect on productivity significant negative effect is found in sample consisting only of plants which experienced statistically meaningful changes in environmental compliance costs
Broberg et al. (2013)	Efficiency score derived from SFA	Environmental protection investments (distinguished into pollution prevention and pollution control)	Four Swedish manufacturing industries, 1999-2004 (F)	<ul style="list-style-type: none"> efficiency measure using translog stochastic production frontier model efficiency modelled as function of regulatory proxy including up to two lags 	<ul style="list-style-type: none"> no support for Porter Hypothesis in overall manufacturing industry negative effect in pulp and paper industry, mainly driven by negative effect of lagged variables
Van der Vlist et al. (2007)	Efficiency score derived from SFA	Dummies for relevant environmental policies	Dutch horticulture sector, 1991-1999 (F)	<ul style="list-style-type: none"> estimation of stochastic production frontier including dummy variables of environmental policies 	<ul style="list-style-type: none"> positive effect on technical efficiency firms under strict environmental policy regime are found to be more technically efficient than those under lax regime

Table A.1. **Overview of empirical studies** (cont.)

Author and year	Dep. variable	Independent variable	Sample	Methodology	Result
Jaraité and Di Maria (2012)	1/ Δ efficiency score derived from DEA	Average annual spot price of CO ₂ allowance Ratio of initial permit allocation to verified emissions	24 European fossil-fuel based public power generating sectors, 1996-2007 (F)	<ul style="list-style-type: none"> ● efficiency measure using DEA incl. CO₂ and SO₂ as bad outputs ● fixed effect estimation 	<ul style="list-style-type: none"> ● adjusted efficiency score increases over time as bad output was reduced but laxity of policy (low CO₂ prices) partly offset the effect ● positive effect on technical change, negative effect on technical efficiency
Fleishman et al. (2009)	Efficiency score derived from DEA	Dummies for SOX and NOX regulation	American gas and coal power plants, 1994-2004 (F)	<ul style="list-style-type: none"> ● efficiency measure using DEA incl. bad outputs ● Tobit estimation 	<ul style="list-style-type: none"> ● positive effect of SO₂ regulations for coal and gas plants ● negative effect of NOX regulation for gas plants ● effect driven by presence not stringency of regulation ● using adjusted or traditional efficiency scores does not alter the results
Managi et al. (2005)	Δ Efficiency score derived from DEA	Environmental compliance cost	American offshore oil and gas industry, 1968-1988 (F)	<ul style="list-style-type: none"> ● efficiency measure using DEA incl. bad outputs ● Almon distributed lag model and Granger causality tests ● includes up to eight lags 	<ul style="list-style-type: none"> ● negative effect in the short term ● positive effect in medium to long term ● excluding bad outputs from efficiency score calculation shows no significant effect of environmental regulation ● evidence that higher technological change of market outputs leads to more stringent environmental regulations
Boyd and McClelland (1999)	Efficiency score derived from DEA	Air and water pollutants, toxics	American integrated paper plants, 1988-1992 (F)	<ul style="list-style-type: none"> ● compares efficiency score under the assumption of weak and free disposability of bad outputs 	<ul style="list-style-type: none"> ● positive effect of environmental regulation on efficiency
Gray (1987)	Δ TFP	PACE Worker and health regulation (number of inspections)	450 American manufacturing industries, 1958-1978 (I)	<ul style="list-style-type: none"> ● uses growth accounting to calculate productivity ● simple regression analysis 	<ul style="list-style-type: none"> ● negative effect when regulation measures are included separately without any covariates ● no significant effect when both regulations plus other explanatory variables are included ● sensitivity tests render PACE coefficient insignificant
Dufour et al. (1998)	Δ TFP	Investment in pollution-control equipment to total input costs worker and health regulation	Canadian manufacturing industries, 1985-1988 (I)	<ul style="list-style-type: none"> ● uses growth accounting to calculate productivity ● GLS estimation, controlling for economies of scale and business cycle fluctuations 	<ul style="list-style-type: none"> ● negative effect of environmental regulation ● effect of worker health regulation depends on type of regulation (negative effect of protective reassignments, positive effect of mandatory prevention programmes and fines)
Barbera and McConnell (1990)	Δ TFP	Abatement capital	Five American manufacturing industries, 1960-1980 (I)	<ul style="list-style-type: none"> ● estimating cost elasticity of abatement capital ● differentiate direct and indirect effect: ● direct effect comes through changes in costs and their effect on productivity ● disentangling effect of abatement capital costs on ΔTFP yields indirect effect 	<ul style="list-style-type: none"> ● small indirect effect (positive/negative) ● negative direct effect (as long as abatement capital grows) ● negative net effect (direct plus indirect)

Table A.1. **Overview of empirical studies** (cont.)

Author and year	Dep. variable	Independent variable	Sample	Methodology	Result
Hamamoto (2006)	TFP	Induced R&D (derived from pollution abatement capital cost effect on R&D spending)	Five Japanese manufacturing industries, 1966-1982 (I)	<ul style="list-style-type: none"> ● elasticity of R&D expenditure w.r.t. lagged regulatory stringency is used to calculate induced R&D ● extended (standard inputs plus R&D capital) Cobb-Douglas production function is used to examine effect on productivity growth ● includes up to six lags 	<ul style="list-style-type: none"> ● negative effect of PACE on age of capital stock which in turn does not affect TFP growth rate ● positive effect of PACE on R&D expenditure and induced R&D positively affects TFP growth ● longest lag structure model performs best
Yang et al. (2012)	TFP	R&D induced by environmental regulation (calculated through abatement capital costs and pollution abatement fees)	Taiwanese manufacturing industries, 1997-2003 (I)	<ul style="list-style-type: none"> ● use Levinsohn and Petrin (2003) productivity measure ● fixed effect estimation ● include one lag 	<ul style="list-style-type: none"> ● positive effect of induced R&D on TFP ● effect of induced R&D is smaller than of scheduled R&D ● positive effect of environmental regulation on R&D ● positive direct effect of environmental regulation on TFP (larger for pollution control fees than for PACE)
Lanoie et al. (2008)	Δ TFP	Investment in pollution-control equipment to total input costs	Canadian manufacturing industries (Quebec), 1985-1994 (I)	<ul style="list-style-type: none"> ● up to three years of lagged regulatory variable ● GLS estimation 	<ul style="list-style-type: none"> ● negative contemporaneous effect ● positive effect of second and third lag (outweighs initial negative effect) ● effect stronger in sectors which face more international competition
Domazlicky and Weber (2004)	Δ Efficiency score derived from DEA	Emission into air, water, land or underground	Six American chemical industries, 1988-1993 (I)	<ul style="list-style-type: none"> ● compare efficiency score under the assumption of weak and free disposability of bad outputs 	<ul style="list-style-type: none"> ● positive effect on productivity growth ● adjusted and traditional efficiency measures are significantly different from each other ● no correlation of abatement costs and productivity growth found
Alpay et al. (2002)	Δ TFP	<ul style="list-style-type: none"> ● Pollution abatement costs (US) ● Frequency of reported plant inspections (Mexico) 	Mexican and US food sector, 1962-1994 (I)	<ul style="list-style-type: none"> ● use elasticities to calculate contribution of environmental regulation to productivity growth ● calculates TFP growth with and without abatement costs 	<ul style="list-style-type: none"> ● positive effect on productivity growth in Mexico, no significant effect in the United States ● negative effect on profitability in Mexico, no significant effect in the United States ● larger productivity growth rates in Mexico when environmental regulation is included in productivity calculation, the United States show slightly lower productivity growth rates
Jeon and Sickles (2004)	Δ Efficiency score derived from DEA	CO ₂ emissions	17 OECD and 11 Asian economies, 1980-1995 (M)	<ul style="list-style-type: none"> ● compares efficiency scores of three scenarios (free emission, no change of emission levels, partial reduction of emissions) 	<ul style="list-style-type: none"> ● adjusted TFP growth is lower than traditional for OECD countries whereas it is higher for ASEAN countries ● productivity growth is lower in constant emission scenario than in free emissions scenario for OECD and ASEAN economies ● productivity growth is higher in scenario of emission reduction in OECD and ASEAN economies

Table A.1. **Overview of empirical studies** (cont.)

Author and year	Dep. variable	Independent variable	Sample	Methodology	Result
Wu and Wang (2008)	Δ Efficiency score derived from DEA	CO ₂ emissions	17 APEC economies, 1980-2004 (M)	<ul style="list-style-type: none"> compares efficiency scores of three scenarios (free emission, no change of emission levels, partial reduction of emissions) fixed effect regression of dummy marking years of UNFCCC ratification on productivity growth 	<ul style="list-style-type: none"> productivity growth slightly higher in scenario of no change and reduction of emission levels than in free emission scenario negative effect of ratification of UNFCCC on productivity growth
Yörük and Zaim (2005)	Δ Efficiency score derived from DEA (CO ₂ , NOX and water pollutants)	UNFCCC protocol ratification	OECD economies, 1983-1998 (M)	<ul style="list-style-type: none"> compares traditional with adjusted productivity index (emission reduction scenario) fixed effect regression of dummy marking years of UNFCCC ratification on adjusted productivity growth 	<ul style="list-style-type: none"> adjusted productivity growth is significantly larger than traditional effect of NOX and water pollutants is largest significant positive effect of UNFCCC ratification on adjusted MFP growth (no effect on traditional MFP growth)
Innovation					
Johnstone and Labonne (2006)	Environmental R&D expenditure	Compiled environmental stringency indicator (survey data)	7 OECD manufacturing sectors, 2003 (F)	<ul style="list-style-type: none"> survey data probit estimation 	<ul style="list-style-type: none"> positive effect of environmental taxes on environmental R&D expenditure negative effect of technology-based standards
Arimura et al. (2007)	Environmental R&D expenditure	Compiled environmental stringency indicator (survey data)	7 OECD manufacturing sectors, 2003 (F)	<ul style="list-style-type: none"> survey data Tobit and bivariate probit estimation 	<ul style="list-style-type: none"> positive effect of perceived policy stringency on environmental R&D expenditure no support for stronger effect of flexible policy instruments
Lanoie et al. (2011)	Environmental R&D Environmental and business performance	Compiled environmental stringency indicator (survey data)	7 OECD manufacturing sectors, 2003 (F)	<ul style="list-style-type: none"> uses survey data (Heckman sample selection procedure) tests all versions of Porter Hypothesis in one common framework probit, 2SLS, instrumental variable probit estimation 	<ul style="list-style-type: none"> support for weak and narrow version of Porter Hypothesis, no support for strong version effect of environmental taxes is driven by stringency of taxes stringency of performance standards has larger impact than technology based standards direct effect on business performance is negative and larger than indirect positive effect
Jaffe and Palmer (1997)	Total R&D expenditure Number of successful patents	PACE	US manufacturing industries, 1975-1991 (I)	<ul style="list-style-type: none"> pooled model and fixed effects estimation including up to five lags 	<ul style="list-style-type: none"> positive effect on R&D expenditure of lagged PACE no effect of PACE on number of successful patents
Brunnermeier and Cohen (2003)	Environmental patents	PACE Number of inspections	146 US manufacturing industries, 1983-1992 (I)	<ul style="list-style-type: none"> different estimation models: fixed effect, Poisson count data model, negative binomial fixed and random effects 	<ul style="list-style-type: none"> positive effect of PACE on number of patents insignificant results for number of inspections environmental innovation is more likely to occur in industries which are exposed to international competition
Kneller and Manderson (2012)	R&D expenditure Total R&D activity	Environmental protection expenditures	UK manufacturing industries, 2000-2006 (I)	<ul style="list-style-type: none"> GMM estimation include up to two lags 	<ul style="list-style-type: none"> positive effect on environmental R&D and investment in environmental capital, no effect on overall R&D or total capital accumulation environmental R&D may crowd out non-environmental R&D no evidence of crowding out effect of environmental capital

Table A.1. **Overview of empirical studies** (cont.)

Author and year	Dep. variable	Independent variable	Sample	Methodology	Result
Lanjouw and Mody (1996)	Patent counts	PACE	US, Japanese and German economies, 1971-1988 (M)	<ul style="list-style-type: none"> ● evaluate effect of pollution abatement capital expenditure on patent count with simple time-series correlation 	<ul style="list-style-type: none"> ● positive effect on patent count, but lagged by 1-2 years ● evidence is found that foreign regulations also influence domestic patent count
Popp (2006)	Environmental patents	SOX and NOX regulations	US, Japanese and German economies, 1967-2003 (M)	<ul style="list-style-type: none"> ● evaluates effect of domestic and foreign regulation on innovation with simple time-series correlation 	<ul style="list-style-type: none"> ● inventors respond to environmental regulation pressure in their own country but not to foreign environmental regulation
Johnstone et al. 2010a	Patent counts in renewable energy sectors	Renewable energy policy variables	25 OECD countries, 1978-2003 (M)	<ul style="list-style-type: none"> ● panel estimated with a negative binomial model, ● fixed effects are included, ● 3 of 6 policy variables are modelled with dummies (introduced or not) 	<ul style="list-style-type: none"> ● renewable energy policies have a significant effect on related patents, ● feed-in-tariffs have an additional positive effect on solar power patents, renewable energy certificates have a positive effect on wind energy patents.
Johnstone et al. 2010b	Environmental patent counts	Perceptions of environmental policy stringency, flexibility and predictability (WEF survey)	OECD countries, 2000-2007 (M)	<ul style="list-style-type: none"> ● panel estimated with a negative binomial model, ● due to high collinearity of the policy variables, orthogonal factors are extracted, ● no fixed effects are included 	<ul style="list-style-type: none"> ● policy stringency, flexibility and stability have a positive coefficient (weak Porter Hypothesis).
De Vries and Withagen (2005)	Environmental patents	Dummy variable for regulations	14 OECD economies, 1970-2000 (M)	<ul style="list-style-type: none"> ● instrumental variable approach ● fixed effect estimation 	<ul style="list-style-type: none"> ● large positive effect on patent count
Kalamova and Johnstone (2011)	Bilateral FDI flows	Perceived environmental policy stringency index (WEF)	27 OECD countries + 99 host countries, 2001-2007 (M)	<ul style="list-style-type: none"> ● panel estimation with controls for drivers of FDI 	<ul style="list-style-type: none"> ● positive effect of lax environmental stringency on FDI inflows ● non-linear effect: FDI inflows decrease after a certain threshold of laxity is reached
Investment and capital stock					
Gray and Shadbegian (1998)	Investment	PACE	American paper mills, 1979-1990 (F)	<ul style="list-style-type: none"> ● different regression specifications, including dummies for high pollution abatement investment years and a variable which divides total plant abatement investment over time 	<ul style="list-style-type: none"> ● plants with high abatement investment spend significantly less on productive capital – crowding out effect. For one dollar spent on environmental capital, investment into productive capital is reduced by 1.88. ● effect is 0.99 when adjusted for within firm (across plant) allocation.
Nelson et al. (1993)	Age of capital stock	<ul style="list-style-type: none"> ● Enforcement costs of environmental agency ● Value of pollution control facilities 	US electric utilities, 1969-1983 (F)	<ul style="list-style-type: none"> ● 3SLS estimation 	<ul style="list-style-type: none"> ● positive effect on age of capital stock ● increase in age of capital does not impact emission level

Note: Studies are listed according to the dependent variable of the analysis.

Abbreviations: DEA – Data Envelopment Analysis; FDI – Foreign Direct Investment; GLS – Generalised Least Squares; GMM – Generalised Methods of Moments; PACE – Pollution Abatement Costs and Expenditures; PAOC – Pollution Abatement Operating Costs; SFA – Stochastic Frontier Analysis; TFP – Total Factor Productivity; 3SLS – Three State Least Square.

In sample descriptions: (F) stands for firm/plant-level analysis, (I) for industry and (M) for macro/country-level analysis.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

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

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

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

OECD Journal

Economic Studies

Lessons from OECD forecasts during and after the financial crisis

Christine Lewis and Nigel Pain

The effect of the global financial crisis on OECD potential output

Patrice Ollivaud and David Turner

The future of health and long-term care spending

Christine de la Maisonneuve and Joaquim Oliveira Martins

Japan's challenging debt dynamics

Yvan Guillemette and Jan Stráský

Rescuing the Phillips curve: Making use of long-term unemployment in the measurement of the NAIRU

Elena Rusticelli

Foreign direct investment and reverse technology spillovers: The effect on total factor productivity

Edmund Amann and Swati Virmani

Environmental policies and productivity growth – a critical review of empirical findings

Tomasz Koźluk and Vera Zipperer

Consult this publication on line at http://dx.doi.org/10.1787/eco_studies-v2014-1-en.

This work is published on the OECD iLibrary, which gathers all OECD books, periodicals and statistical databases.
Visit www.oecd-ilibrary.org for more information.

Volume 2014

OECD publishing
www.oecd.org/publishing



ISSN 1995-2848
VOLUME 2014 (1 ISSUE)
13 2014 01 1 P



9 771995 284003