norden

The *Nordic Economic Policy Review* is published by the Nordic Council of Ministers and addresses policy issues in a way that is useful for informed non-specialists as well as for professional economists. All articles are commissioned from leading professional economists and are subject to peer review prior to publication.
The review appears twice a year. The journal is distributed free of charge to the members of the Nordic economic associations. The easiest way to subscribe to the NEPR is therefore to become a member of one of these associations, i.e.,

Denmark: Nationaløkonomisk Forening
Finland: Taloustieteellinen Yhdistys
Norway: Samfunnsøkonomene
Sweden: Nationalekonomiska Föreningen.

For institutional subscriptions, please contact nepr@iies.su.se.

# Nordic Economic Policy Review

Number 2 / 2011

Productivity and competitiveness

Editors: Jakob B. Madsen and Anders Sørensen

TemaNord 2011:570

**Nordic Economic Policy Review Number 2 / 2011**
Productivity and competitiveness

TemaNord 2011:570


978-92-893-2295-9
Print: Arco Grafisk A/S, Skive
Cover: Pub.Unit/NCM
Layout: Pub.Unit/NCM
Copies: 5030

Printed on environmentally friendly paper
This publication can be ordered on www.norden.org/order. Other Nordic publications are available at www.norden.org/publications

Printed in Denmark

**Nordic Council of Ministers**
Ved Stranden 18
DK-1061 København K
Phone (+45) 3396 0200
Fax (+45) 3396 0202
**www.norden.org**

**Nordic co-operation**

*Nordic cooperation* is one of the world's most extensive forms of regional collaboration, involving Denmark, Finland, Iceland, Norway, Sweden, and three autonomous areas: the Faroe Islands, Greenland, and Åland.

*Nordic cooperation* has firm traditions in politics, economics, and culture. It plays an important role in European and international collaboration, and aims at creating a strong Nordic community in a strong Europe.

*Nordic cooperation* seeks to safeguard Nordic and regional interests and principles in the global community. Common Nordic values help the region solidify its position as one of the world's most innovative and competitive places.

# Content

# The widening productivity gap between the EU and the US: An introduction to the conference on productivity and competitiveness

Jakob B. Madsen* and Anders Sørensen**

## *Summary*

This paper introduces the papers presented at the Nordic Conference on Competitiveness and Productivity and discusses them in relation to the widening productivity gap between Europe and the US since 1995. It is shown that productivity growth is driven by R&D, human capital and efficiency of production such as human resource management, government regulation, competition and openness. It is concluded that potential explanations for the widening productivity gap are low investment, inefficient use of information technology, compressed wage structure and deregulation of the otherwise regulated labor market in Europe.

*Keywords:* Endogenous growth, productivity gap, convergence.
*JEL classification numbers:* O1, O2, O4.

---

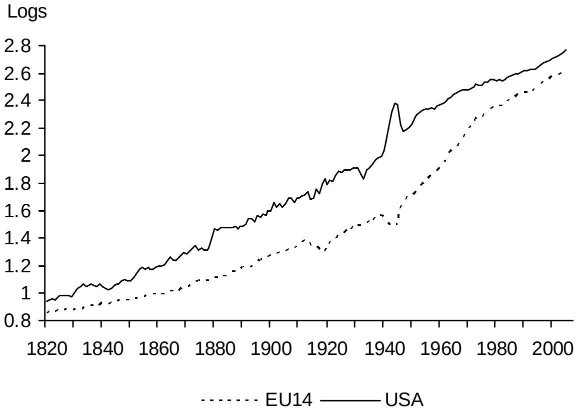
* Department of Economics, Monash University, jakob.madsen@buseco.monash.edu.au.

** Corresponding author, Department of Economics, Copenhagen Business School, as.eco@cbs.dk.

Over the last two centuries, trend productivity has been growing between one and two percent in the US and the EU14 countries, where the EU14 countries consist of the 15 EU members in 2004 excluding Luxembourg (see Figure 1 for a country sample). Productivity is here measured as total factor productivity (TFP), where labor is measured as employment multiplied by annual hours worked and income and capital stock are measured in purchasing power parities. Figure 1 shows that the US had been the productivity leader over EU14 during the last two centuries. This does not mean that the US was ahead of all EU14 countries during the entire period. The UK was more productive than the US in the nineteenth century; however, the US took the lead in the early twentieth century.

The productivity gap between the US and EU14 widened over the period from 1820 to 1945 due to stronger increases in educational attainment and research intensity in the US than in EU14 (Madsen, 2010). Thereafter, the gap narrowed substantially up to mid 1995 due to technology spillovers through imports, a stronger increase in research intensity and educational attainment in EU14 than in the US, and catch-up to the technology frontier (Madsen, 2007, 2008).

**Figure 1. Total factor productivity**




Source*:* See Madsen (2010).

Note*:* The EU14 countries consist of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden and the UK. The EU average is an unweighted average.

The productivity gap between the US and EU14 has widened since 1995, a question that this special issue is concerned about. The recent divergence raises the question as to why the catching-up process has stalled, why the productivity performance in Europe has been relatively poor since the mid 1990's and whether the gap will continue to widen. Basically, the gap will continue to widen if structural and institutional factors in EU15 have permanently lowered the incentives to do R&D, if the R&D productivity is slowing down in EU15 relative to the US, and if EU15 cease to take advantage of the technology that is produced in the US. The papers in this special issue examine these topics from a microeconomic and macroeconomic perspective; both of which give important insight into factors affecting productivity.

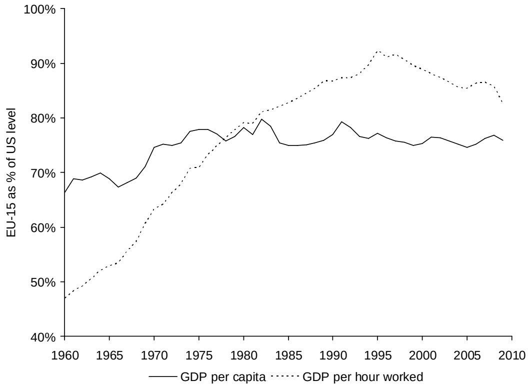
In this paper, we examine different factors that affect the absolute and relative productivities among the OECD countries and briefly discuss which factors highlighted in the contributed papers in this special issue suggest as potential answers. The next section sets the scene of the productivity path in the US and EU15 and investigates macro issues related to this (Section 3). The micro issues are discussed in Section 4 and Section 5 concludes the paper.

## 1. The productivity puzzle

Figure 2, which is from van Ark, shows the path of relative productivities between the US and EU15. The figure shows, in terms of productivity per hour worked, that EU15 was catching up with the US during the period 1950 to 1995; thereafter the relative productivities have diverged. In absolute terms, EU15 and the US both experienced high productivity rates during the period 1950-1973; thereafter, during the period 1973-1995, the growth rates slowed down considerably, particularly in the US. Van Ark finds that the labor productivity convergence in the period 1973-1995 was driven by convergence in capital intensity. The labor productivity growth rates recovered in the US after 1995 while they slowed down in the EU15 and, therefore, resulted in a reversal of the convergence that took place up to 1995. Van Ark argues that the slow labor productivity in Europe after 1995 may have been related to the loosening up of some labor market rigidities that have increased employment in EU15 more than in the US. Van

Ark shows that the recent decline in labor productivity growth in EU15 has been dominated by a marked productivity growth slowdown in trade, finance and business services.[1]

**Figure 2. Total economy GDP per hour worked and GDP per capita in EU15, 1960-2009 (relative to the United States)**




Source: The Conference Board Total Economy Database, January 2011, http://www.conference-board.org/ economics/database.cfm.

Note*:* EU15 refers to the 15 countries constituting the European Union before 2004 and includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom. Relative levels are based on purchasing power parities for GDP for 2005 from the OECD.

The finding of van Ark raises the question as to why productivity has slowed down in EU15 while, at the same time, it has picked up in the US since 1995 and resulted in a productivity divergence. Van Ark addresses this issue by decomposing the sources of growth and finds that the growth in total factor productivity has been the main factor behind the productivity slowdown in EU15 which, in turn, raises the issue of which factors have been responsible for this path in TFP. In the next section, it is shown that TFP is driven by R&D, human capital and efficiency of production such as management practice, government regulation and export exposure of companies.

[1] To gain further insight into the US-EU productivity puzzle, it would be obvious to study the role of sectors in the divergence of aggregate productivity. However, such an analysis is not possible due to the lack of an appropriate conversion factor for measuring productivity levels at the sector level; see Sørensen (2001) and Sørensen and Schjerning (2008).

# 2. Macroeconomic perspective on the productivity puzzle

## *2.1 TFP, R&D and growth*

Below it is argued that R&D is driving technological progress and, therefore, labor productivity growth along the balanced growth path. The knowledge stock denoted by *A* can be influenced by several factors beyond R&D such as human capital, resource allocation, regulations, management practices, openness of the economy, organizational efficiency etc; i.e. some of the factors on which we will focus below and that are discussed in the papers in this special issue on productivity and competitiveness. If these factors also affect the incentives to undertake R&D, the efficiency of undertaking R&D and the implementation of new technology, they can have permanent growth effects.

In the remainder of this section, we treat *A* as the technological state of the economy that can be influenced by R&D and human capital. R&D and human capital can have temporary as well as permanent growth effects depending on whether there are any scale effects in ideas production. To show this more explicitly, consider the following homogenous Cobb-Douglas production function:

$$Y = K^{\alpha}(AhL)^{(1-\alpha)}, \qquad (1)$$

where *Y* is real output, *A* is the technological state or level of TFP, *K* is capital stock, *L* is raw labor, *h* is human capital such as educational attainment, and $\alpha$ is the share of income going to capital.

Eq. (1) can be rewritten in terms of labor productivity growth rates as:

$$g_{Y/L} = \alpha g_{K/L} + (1-\alpha)(g_A + \mathrm{g_h}), \qquad (2)$$

where $g_{Y/L}$ is labor productivity growth, $g_{K/L}$ is growth in the capital-labor ratio, $g_A$ is technological progress, and $g_h$ is growth in educational attainment. From this equation, it follows that labor productivity growth is a positive function of capital deepening, technological progress and growth in educational attainment. However, since capital deepening in the long run is entirely driven by technical progress and human capital, which is usually measured by educational attainment and cannot grow forever, labor productivity growth is entirely driven by technological progress in the long run. Increasing investment and educational attainment, or the quality of teaching, can boost labor productivity temporarily but not in the long run. Furthermore, productivity can only keep growing in the long run if the technological progress is sustainable.

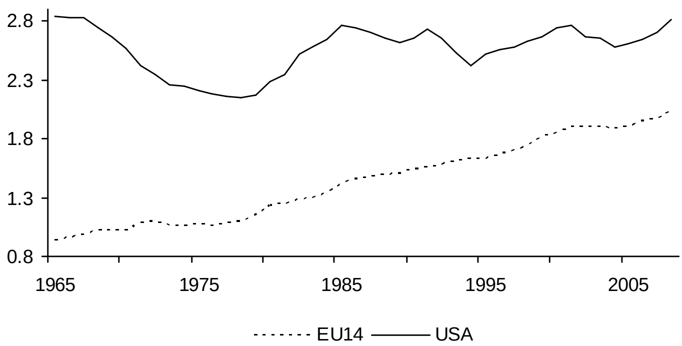
To see which factors determine technological progress, consider the following ideas production function:

$$\Delta A = \lambda \left( X / Q \right)^{\sigma} A^{\gamma}, \qquad 0 < \sigma \leq 1, \quad \phi \leq 1, \tag{3}$$

where $\lambda$ is a research productivity parameter, $Q$ is product variety, σ is a duplication parameter (0 if all innovations are duplications and 1 if there are no duplicating innovations), $X$ is the innovative activity such as R&D or patenting, $\Delta A$ is the number of new ideas generated, and $\gamma$ is returns to scale in knowledge. From Eq. (3), it follows that technological progress is driven by R&D. Whether the growth effects of R&D are permanent or temporary depends on the knowledge scale parameter, $\gamma$, which is one in first-generation growth models. First-generation endogenous growth models assume that growth is proportional to R&D. Since growth rates have not increased proportionally with the massive increase in R&D workers in the OECD countries over the past 50 years, first-generation endogenous growth models have been replaced by second-generation models. The two dominating second-generation endogenous growth models are semi-endogenous growth theory and Schumpeterian growth theory. Semi-endogenous growth theory assumes that $\gamma < 1$, so that R&D only has temporary growth effects. Schumpeterian growth models assume that $\gamma = 1$, which implies that as long as research intensity is positive the economy will continue to grow. However, the latter group of models applies a functional form that does not imply a long-run growth rate that increases proportionally with the number of R&D workers, which was the objection towards the first-generation endogenous growth models.

Empirical evidence is predominantly in favor of the Schumpeterian growth paradigm (Ha and Howitt, 2007; Madsen, 2008, 2010; Ang and Madsen, 2012), which implies that R&D has permanent growth effects and that productivity growth is positively related to research intensity. Madsen (2008, 2010) and Ang and Madsen (2012) find that growth is significantly positively related to research intensity. Figure 3 shows research intensity measured by economy-wide R&D expenditures divided by nominal GDP for the US and EU14, where EU14 is once more EU15 minus Luxembourg. The graph shows that the research intensity has fluctuated around 2.5 percent for the US while it has been markedly increasing in EU14 since 1965. The sharp decline in the US during the period from 1968 to 1978 to a large extent reflects declining R&D expenditures related to the NASA program. The convergence in R&D expenditures between the US and EU14 is likely to have played an important role for the productivity convergence between the two country groups over the period considered.

**Figure 3. R&D-GDP ratio**




Sources. See Madsen (2008).

Notes*:* See notes to Figure 1.

During the period 1965-2008, which is covered by Figure 3, research intensity has been higher in the US than in EU14 despite EU14 having experienced higher productivity growth rates than the US. This result is consistent with Madsen (2008) who fails to find a positive relationship between growth and research intensity in the cross-country dimension. This lack of correlation in the cross-section dimension is likely to reflect the fact that research productivity varies across nations and that productivity growth is influenced by factors other than R&D such as human capital, quality of teaching, human resources management, regulations, catch-up, and the openness of the economy.

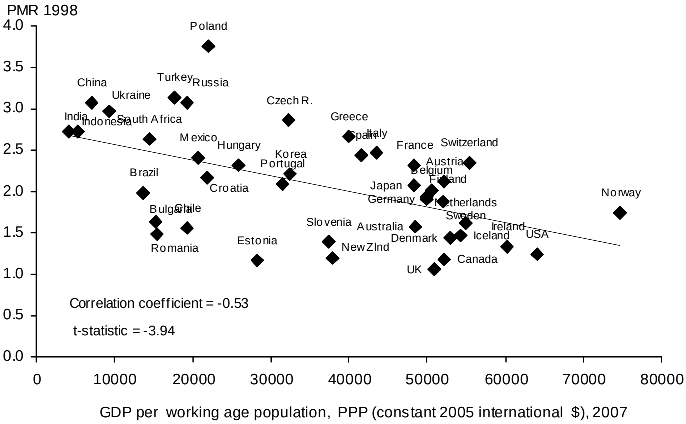
### *2.2 Regulation and resource allocation*

The reallocation of production factors from low to high productive increases economic growth. This positive relationship has been established for a group of OECD countries in Sørensen (2007), where a reallocation of labor input from low- to high-wage sectors is found to result in higher aggregate productivity growth.

The results of anticompetitive measures can result in the failure to catch up to the frontier by altering the incentives to adopt the leading technologies available in the market and innovate among the existing firms. This can be done by reducing the rivalry among incumbents and by making the entry of new, innovative, firms difficult. Conversely, the opening up of markets and increased competitive pressures provide both opportunities and incentives for firms to upgrade their capital stocks, adopt new technologies and innovate to reach and possibly push out frontier production techniques. Furthermore, anticompetitive regulations may deter the reallocation of resources from less to more efficient firms by increasing adjustment costs such as entry and exit costs, and adjustment costs to reallocate factors of production such as capital and labor.

In their paper on regulation and resource allocation, Arnold, Nicoletti and Scarpetta examine whether regulations can affect productivity by preventing resources from flowing to the most productive uses. There are several channels through which anticompetitive product market regulation can affect productivity. Although product market regulation seeks to prevent the formation of monopolies, it often results in regulations drifting away from their original public interest aims, resulting in the protection of special interest groups. Furthermore, the costs may exceed their expected benefits and the designs do not often follow the evolvement of the industrial structure and product composition. Figure 4, which is identical to Figure 1 in Arnold, Nicoletti and Scarpetta, indicates that anticompetitive product regulation is associated with a lower income per capita across countries. Among the OECD countries, there has been a convergence towards a more pro-competitive stance during the period 1998-2008. There has also been a convergence between EU15 and the US, which suggests that this factor has not been responsible for the recent productivity divergence between the US and EU15.

**Figure 4. Anticompetitive product market regulation and GDP per capita**[a]




Source: Woelfl et al. ( 2010).

Note: [a] Based on a "simplified" OECD PMR indicator (see Woelfl et al., 2010). PMR measured in 1998 for OECD countries; 2008 for Chile, Estonia, Israel, Slovenia, Brazil and China; 2007 for Croatia, Indonesia, South Africa and Ukraine; 2006 for Bulgaria, India and Romania.

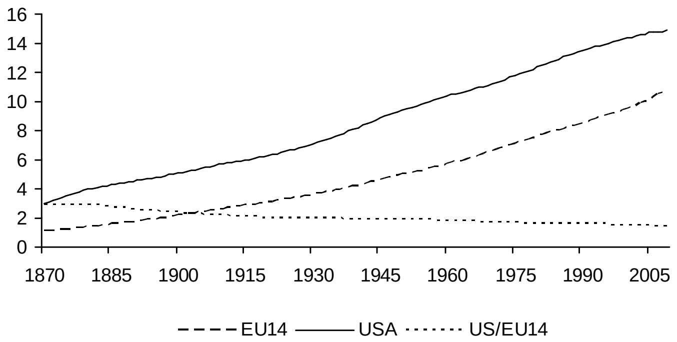
## *2.3 Human capital and growth*

Human capital is influential for income and growth by enhancing the quality of workers and potentially also ideas production. Human capital includes everything that affects the skills of individuals such as on-the-job training, educational attainment and cognitive skills. For practical purposes, human capital is predominantly measured as educational attainment among the population of working age, where educational attainment is the average number of years of schooling. Educational attainment can be measured using gross enrolment rates at different levels among different age cohorts, where gross enrolment rates are measured as the percentage of an age group that is enrolled in primary, secondary and tertiary education. Alternatively, educational attainment can be measured directly from population censuses that are usually carried out in five or ten year intervals.

De la Fuente summarizes the effects of educational attainment on productivity growth in the OECD countries and Spanish states based on his previous research. He uses census data for the OECD countries; however, the data are corrected for outliers and implausible numbers since there are quite a few implausible numbers in the data. De la Fuente finds that educational attainment has significantly positive productivity growth effects and that the social returns to education are higher than the social returns to fixed capital, which renders investment in human capital economically viable from society's point of view.

Figure 5 shows educational attainment among the working age population since 1870. Both the US and EU15 have experienced massive improvements in educational attainment, and it has been an important factor behind the industrialization (Madsen, 2010). Although de la Fuente finds productivity growth to be positively related to the *level* of human capital, Figure 5 suggests that there can probably not be a permanent positive relationship between productivity growth and human capital: otherwise there would have been a marked increase in labor productivity growth rates over the period 1870 to 2009. This has not been the case, however. Furthermore, it is unlikely that the pace of innovations will increase as the workforce becomes more educated. Better educated people are more likely to increase the efficiency of production and, therefore, the level of productivity than increasing the rate of innovation. Does that refute the findings of de la Fuente? Not necessarily since he considers the movement in productivity from one steady state to another. Thus, his analysis does not imply any permanent growth effects of educational attainment.

**Figure 5. Educational attainment**




Source*:* Madsen (2010).

To shed some light on the productivity puzzle, the relative level of educational attainment between the US and EU14 is represented by the dotted line in Figure 5. The decline in relative educational attainment during the entire period shows that EU14 has been catching up with US educational attainment and may well have been a factor that has contributed to the productivity convergence. However, since the convergence in educational attainment has continued since 1995, educational attainment has not contributed to the productivity divergence since 1995.

## 3. Microeconomic perspective on the productivity puzzle

### *3.1 Export and productivity*

The world has undergone three globalization waves since the First Industrial Revolution started in the mid eighteenth century in the UK. The first wave was in the latter part of the eighteenth century and was brought to a halt by the Napoleonic War, the second wave started during the Second Industrial Revolution and was brought to a halt by WWI and the last innovation wave started in the 1960's and has since then gained momentum in terms of international trade and capital movements. A recurring question is whether trade openness increases growth. There are several reasons as to why trade openness increases growth (see Madsen, 2009), the main reason being that it exposes firms to more competition and, as such, forces them to innovate and keep pace with the technology frontier. Furthermore, openness is associated with foreign direct investment (FDI) which tends to be technologically more advanced than domestic investment.

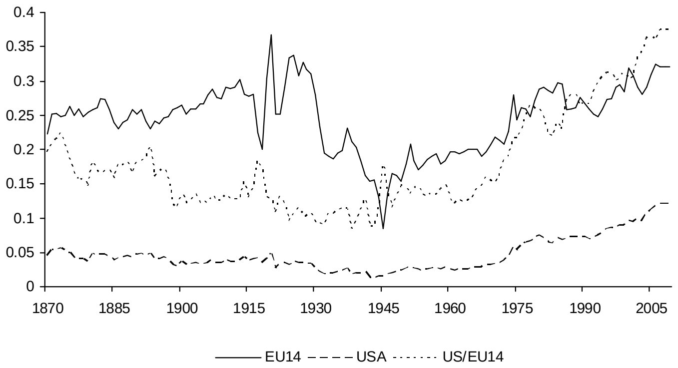
Wagner surveys the micro evidence on the relationship among productivity, export exposure and FDI. The evidence suggests that producers that are more exposed to the export market are more productive than producers that predominantly serve the domestic market. Foreign direct investors are, in turn, more productive than export producers, which further stresses the importance for producers of being exposed to foreign competition. The trouble with these empirical findings is how to deal with the two-way relationship between foreign exposure and productivity, as discussed by Wagner. More productive and sophisticated firms tend to invest abroad and have a high export exposure because they are sophisticated in the first place. However, if these firms are growing in an environment with high trade openness, they are likely to be more productive than in a closed economy.

In all events, there seem to be positive productivity effects associated with openness even when feed-back effects from growth to exports are allowed for. Thus, there are productivity gains associated with policies that seek to bring down foreign trade and investment barriers such as tariffs and embargoes on imports of certain products and restrictions of FDI and other capital flows between firms.

Has the path in export exposure contributed to the widening of the productivity gap between the US and Europe? Figure 6 shows openness over the past 140 years in the US and EU14. Openness is measured as imports divided by GDP. Ideally, openness should be measured as exports plus imports divided by nominal GDP. However, data on exports are not widely available back in time and imports tend to follow exports in the long run. The figure shows that openness peaked in EU14 in the mid 1920's and not recently as is often assumed. Openness declined for both country groups during the interwar period and was driven by the Depression and an increasing nationalism (Madsen, 2001). Openness has been increasing since WWII; particularly in the US. The ratio between openness in the US and EU14 has more than doubled since 1960 as the US economy has been transformed from being quite a closed economy to a more open economy, noting that the US economy will always be more closed than EU14 because of the sheer size of the country in terms of population.

Does the path in relative openness explain the productivity gap between the US and EU14 over time? Productivity-wise, the US was forging ahead of EU14 during the period in which the relative openness was declining (1870 to 1945). Openness converged quite quickly during the period 1960-1978, which was a period during which EU14 was quickly catching up with the US. However, the lack of a positive correlation between openness and productivity growth does not disprove a positive relationship; it may just be that openness is not that influential for productivity and that factors other than openness have overridden the positive productivity effects from trade. The US's recent increase in relative openness may have contributed to its productivity advance *vis-à-vis* EU14.

**Figure 6. Openness**




Sources*:* See Madsen (2009).

Notes*:* Openness is measured as imports of goods and services divided by nominal GDP.

### *3.2 Product and process innovations*

Hall reviews evidence on the relationship between innovations and productivity at industry and firm levels and discusses measurement issues related to innovation and TFP. Innovation is a broader concept than R&D. Where R&D is creative work undertaken on a systematic basis in order to increase the stock of knowledge and the use of this stock of knowledge to devise new applications, innovation is the introduction of new or significantly improved products or processes. According to these definitions, innovation for example also includes the development of products that are new to the market or to the firm but not necessarily to the world; such activities are not included in R&D.

Hall shows that the measurement of TFP is not trivial under imperfect competition and when the data in factor inputs are mismeasured. Furthermore, Hall makes a very important distinction between product and process innovation, where product innovation increases the firm's revenue while process innovation increases the firm's productivity under imperfect competition – a distinction that is difficult to make at the macro level. In her survey, Hall finds strong empirical evidence of positive labor productivity effects of product innovations and a somewhat more ambiguous impact of process innovation.

Does differences in innovation activities explain the widening in the productivity gap between the US and EU since 1995? According to Hall, the empirical results do not suggest that European firms underperform in relation to innovation. Innovation in European countries is as high or maybe higher than in other countries outside Europe. Thus, since growth is driven by innovations in the long run, innovations are unlikely to be an important source of the widening productivity gap.

### *3.3 A personnel economics approach to productivity enhancement*

Lazear and Shaw show that there are large productivity variations among workers within the same firm and that the variance is greater among high-skilled jobs than among lower-skilled jobs. Furthermore, they find that even within very narrowly defined jobs, productivity varies significantly across workers within narrowly defined occupations. The variance of productivity that is explained by personal differences is 71 percent for windshield installers, while it is only 7 percent for teachers (measured as variations in student test scores). Given the large difference between star performers and low-performers, there is plenty of room for improvement in individual productivity within firms. The use of more innovative human resources (HR) practices can narrow the gap between high and low performers within and across firms. Personnel economics, founded by Lazear, shows how HR practices can raise productivity. The studies summarized by Lazear and Shaw show that the effect of personnel practices innovations can be as high as 50 percent of output.

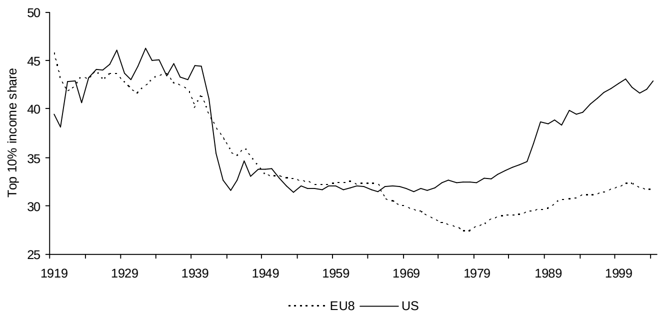
HR practices can raise productivity by giving workers the incentive, the opportunity, and the ability to do this, where incentive pay gives workers the incentive, team problem-solving gives them the opportunity, and training gives them the ability to raise their performance. Lazear (2000) finds that the introduction of piece rate pay among workers who drive trucks to customers' homes and install windshields in cars that have damaged windshields has substantial productivity effects. When the firm changes from hourly pay to payment by the amount of windshields installed, productivity rises by 26 percent because employees raise their effort on the job. Furthermore, Lazear finds that productivity rises by about 20 percent more because more productive workers are hired by the firm. Thus, the economy can gain a better match between workers' skills and firms' needs, to the extent to which the firm does not attract higher effort workers.

Personnel economics may give some insight into why the productivity gap between the US and EU15 has widened since 1995. Although no macro evidence on the contribution of management practices to productivity growth has thus far been produced, Lazear and Shaw suggest that there is indirect evidence supporting the view that human resource management has recently contributed to growth. First, productivity has risen over time in industries that are computer-using industries, and computer use is correlated with HR changes in the US. At the sectoral level, this argument is consistent with the findings of Stiroh (2002) and Dahl et al. (2011) who estimate the quantitative effect on labor productivity growth of IT after 1995 in the US and Europe, respectively. Stiroh finds an increase of 2 percent for US labor productivity growth in ICT-intensive industries post-1995, whereas Dahl et al. find an increase for European economies of around 40 percent of the IT-induced productivity effects in the US. In this sense, the difference in the utilization of IT between the two regions has contributed to the productivity divergence.

Second, time trends show that US firms have increased their use of innovative HR management practices which may not be the case in European firms. The view is confirmed by Bloom et al. (2011) who find that US multinationals operating in Europe obtained a higher productivity from IT than non-US multinationals. Moreover, it is found that the US advantage in IT is primarily due to its HR practices on promotions, pay, hiring and firing.

A third piece of evidence of whether personal economics can give some insight into the productivity puzzle is presented in Figure 7. In the words of Lazear and Shaw: "perhaps the most important point that comes from the compensation literature is that pay compression is the enemy of productivity". Based on this quote, we present the income inequality over the past century for the US and EU8 (see the note for Figure 7 for the country sample). Income inequality is measured as the income share among the top 10 percent income earners. The figure shows that income inequality has widened between EU8 and the US since the mid 1960's. However, the inequality gap between the US and EU8 first started to widen markedly at the end of the 1980's. The widening inequality may explain some of the productivity gap between the US and EU8 because the widening inequality in the US may, to some extent, reflect an increasingly efficient use of resources.

**Figure 7. Income inequality**




Source*:* Leigh (2007).

Notes*:* EU8 consists of Finland, France, Germany, Ireland, the Netherlands, Spain, Sweden and UK.

### *3.4 Teaching efficiency*

Using data on test scores for various subjects for approximately 1 400 public schools, Bogetoft and Wittrup examine the efficiency of these schools in terms of student grades while controlling for schooling resources used, peer group effects and socioeconomic background such as parents' educational background, income, employment, and ethnicity. Since some of these factors are highly influential for schooling grades, it is important to control for them. The results show that socioeconomic background explains 28 percent of the variation in test scores among students while 93 percent of the remaining variance are explained by individual ability and only 7 percent of the remaining variance are explained by the quality of the school.

Furthermore, Bogetoft and Wittrup's estimates show that schools can, on average, save 13 percent of their expenditure without sacrificing quality if their resource allocation follows best practice resource allocation. Finally, they show that the average score among students can be improved by ten percent without increasing the resource usage. The main contributors to this improvement are the size of the school and the allocation of teachers' time towards more teaching.

## 4. Concluding remarks and policy implications

The discussion in this issue suggests that the productivity performance in EU15 has not been impressive recently and that its productivity relative to the US has widened over the last 15 years. However, it is not at all clear which factors have been responsible for this development. Potential candidates, as discussed in this issue, are R&D, educational attainment, quality of schooling, human resources management, regulation and deregulation and export exposure. However, relative to the US, research intensity and educational attainment among the population of working age among the EU countries have improved and, as such, have not contributed to the widening of the productivity gap. The stronger reduction in the anticompetitive regulation among the EU15 countries than the US since 1998 (the first year at which data are available) should have alleviated the productivity growth decline in EU15 relative to the US.

Lazear and Shaw as well as van Ark suggest that an important explanation for the productivity paradox between the US and EU15 is related to information technology. Van Ark points to high levels of investment in information technology in the US in the second half of the 1990's, followed by a rapid productivity growth in the market services sector of the economy in the first half of the 2000's – two developments that did not take place in Europe. Lazear and Shaw suggest that there is indirect evidence supporting the view that complementarities between human resource management and information technology have recently contributed to growth, especially in the US. First, productivity has risen over time in computer-using industries, and computer use is correlated with HR changes in the US. Second, time trends show that US firms have increased their use of innovative HR management practices which may not be the case in European firms; a view that is supported by Bloom et al. (2011).

Another potential contributing explanation for the productivity puzzle suggested by Lazear and Shaw is the compressed wage structures in many European economies. In the words of Lazear and Shaw: "perhaps the most important point that comes from the compensation literature is that pay compression is the enemy of productivity". This is an avenue that may be fruitful to explore in greater detail to gain some insight into the poor productivity performance of European economies. As a final potential explanation, van Ark points towards the labor market as a potential source of the productivity puzzle. The deregulation of the otherwise very regulated labor market in Europe since the 1980's has probably increased the demand for labor, particularly of unskilled types and, as such, has counterbalanced some of the increasing capital intensity that has occurred as a result of technological progress since 1995 and, thus, has prevented labor productivity from increasing as much as it would otherwise have done.

## References

Ang, J.B. and Madsen, J.B. (2012), Can second-generation endogenous growth models explain the productivity trends in the Asian miracle economies?, forthcoming in Review of Economics and Statistics.

Bloom, N., Sadun, R. and Van Reenen, J. (2011), Americans do I.T. better: US multinationals and the productivity miracle, forthcoming in American Economic Review.

Dahl, C.M., Kongsted, H.C. and Sørensen, A. (2011), ICT and productivity growth in the 1990s: Panel data evidence on Europe, Empirical Economics 40, 141-164.

Ha, J. and Howitt, P. (2007), Accounting for trends in productivity and R&D: A Schumpeterian critique of semi-endogenous growth theory, Journal of Money, Credit and Banking 39, 733-774.

Lazear, E.P. (2000), Performance pay and productivity, American Economic Review 90, 1346-1361.

Leigh, A. ( 2007), How closely do top income shares track other measures of inequality?, Economic Journal 117, F589-603.

Madsen, J.B. (2001), Trade barriers and the collapse of world trade during the great depression, Southern Economic Journal 67, 848-868.

Madsen, J.B. (2007), Technology spillover through trade and TFP convergence: 135 Years of evidence for the OECD countries, Journal of International Economics 72, 464-480.

Madsen, J.B. (2008), Economic growth, TFP convergence and the world export of ideas: A century of evidence, Scandinavian Journal of Economics 110, 145-167.

Madsen, J.B. (2009), Trade barriers, openness, and economic growth, Southern Economic Journal 76, 397-418.

Madsen, J.B. (2010), The anatomy of growth in the OECD since 1870, Journal of Monetary Economics 57, 753-767.

Sørensen, A. (2001), Comparing apples to oranges: Productivity convergence and measurement across industries and countries: Comment, American Economic Review 91, 1160-1167.

Sørensen, A. (2007), Inter-industry wage differentials and allocative inefficiency, International Economic Journal 21, 1-26.

Sørensen, A. and Schjerning, B. (2008), Productivity measurement in manufacturing and the expenditure approach, Review of International Economics 16, 327-340.

Stiroh, K.J. (2002), Information technology and the U.S. productivity revival: What do the industry data say?, American Economic Review 92, 1559-1576.

Woelfl, A., Wanner, I., Roehn, O. and Nicoletti, G. (2010), Product market regulation: Extending the analysis beyond OECD countries, OECD Economics Department Working Paper 799, Paris.

# Up the hill and down again: A history of Europe's productivity gap relative to the United States, 1950-2009

Bart van Ark[*]


## *Summary*

This paper provides a historical perspective on the American-European productivity paradox since 1995. It reviews the growth and productivity performance in both regions since 1950, and investigates the reasons for the widening of the productivity gap since 1995. Using the EU-KLEMS Growth and Productivity Accounts, the paper shows that the productivity slowdown in Europe is mainly attributable to the slower emergence of the knowledge economy compared to the United States. This is observed in low growth contributions from investment in information and communication technology in Europe, the relatively small share of technology-producing industries, and slow multifactor productivity growth. The article emphasizes the role of market service sectors in accounting for the productivity growth divergence between the two regions.


*Keywords:* Productivity, economic growth, investment, innovation.
*JEL classification numbers:* O47, O52.

[*] The Conference Board and University of Groningen, bart.vanark@conference-board.org.

The benefits of the modern knowledge economy differ greatly between advanced economies. The EU-15, that is the 15 European Union countries that constituted the Union up to 2004 and that constitute the focus of this chapter, experienced a sharp slowdown in labour productivity growth (measured as GDP per hour of work) from an annual rate of 2.7 per cent during the period 1973-1995 to 1.5 per cent during the period 1995-2007. At the same time, labour productivity in the United States increased sharply from 1.3 per cent to 2.1 per cent between 1973-1995 and 1995-2007. While differences in the timing of business cycles in the United States and the European Union may have some effect on this comparison, they do not explain these divergent trend growth rates. Another explanation that has been pursued in the literature is the comparative variant of the productivity paradox, that is, the smaller productivity gains from new technologies, notably information and communication technology (ICT), and innovation, in Europe relative to the United States. The explanations have been studied in detail elsewhere (Gordon, 2004; Timmer and van Ark, 2005).

This paper places the American-European productivity paradox in a historical perspective. The slower labour productivity growth rates in Europe compared to the United States since 1995 reverse a long-term pattern of convergence. This article first reviews the growth and productivity performance in Europe since 1950, considering three periods characterized by different drivers of productivity. In the period 1950-1973, European growth was characterized by a traditional catch-up pattern based on the imitation and adaptation of foreign technology, coupled with strong investment and supporting institutions. However, the traditional postwar convergence process came to an end by the mid 1970's (Crafts and Toniolo, 1996; Eichengreen, 2007). Then, output and labour productivity growth in both Europe and the United States began to slow down in the period from 1973 to 1995. However, while the gap in output (and average per capita income) growth rates narrowed between the two regions, Europe's productivity growth remained much faster than that in the United States. During this time period, Europe experienced a strong decline in labour force participation and a fall in hours worked which, in turn, triggered a substitution of capital for labour bringing capital-labour ratios in some major European economies to levels well above those of the United States by the mid 1990's. US labour productivity growth accelerated from 1995 until around 2004, after which it began to slow down, whereas the rate of productivity growth in Europe fell throughout the period, with the exception of two brief positive spells during the peaks of the business cycle at the end of the 1990's and around 2006-2007. Finally, during the Great Recession in 2008/2009, the labour productivity growth rates in Europe and the US rapidly diverged, as the US saw a pickup in productivity growth as the labour market shrunk well beyond that in the European Union. Therefore, the EU saw a decline in productivity parallel to the contraction of the economy.

In Section 2 of the article, we focus on the European growth experience, especially in the period from 1995 to 2007, using a new and detailed database called the EU-KLEMS Growth and Productivity Accounts. The level of detail in this database makes a discussion of a number of developments during this period possible: changes in patterns of capital-labour substitution; the increasing importance of investment in information and communications technology; the use of more high-skilled labour; the different dynamics across sectors, like those producing information and communications technology, or manufacturing and services more generally; and the diversity of productivity experience across the countries of Europe.

We show that the slowdown in Europe since the mid-1990's is mainly attributable to the slower emergence of the knowledge economy as compared to the United States. In Section 3, we consider various explanations which are not mutually exclusive: for example, lower growth contributions from investment in information and communication technology in Europe, the relatively small share of technology-producing industries in Europe, and slower multifactor productivity growth (which can be viewed as a proxy for advances in technology and innovation). Underlying these explanations are issues related to the functioning of European labour markets and the high level of product market regulation in Europe. This article emphasizes the key role of market service sectors in accounting for the divergence in productivity growth between the two regions (see also Inklaar et al., 2008).

Finally, in the concluding section, we look at some of the policy implications for Europe to strengthen its productivity growth performance. The slowing growth and faltering emergence of the knowledge economy in Europe since the mid-1990's have led to an ambitious action program of the European Commission, called the "Lisbon Agenda", which was executed during the first decade of the 21st century. Its goal was to make Europe by 2010 "the most competitive and dynamic knowledge-based economy in the world." In 2010, this program was succeeded by a new growth strategy "Europe 2020" to become a smart, sustainable and inclusive economy. Both strategies have been focused on the importance of employment growth, innovation, especially through ambitious targets for research and development, as well as a focus on environmentally friendly growth strategies. Unfortunately, neither strategy has so far led to a reversal of Europe's downward productivity trend.

Although we do not believe that there is one silver bullet to revive growth, we argue that the future for European productivity growth will strongly depend on the performance of its services sector. The nations of Europe also need to find their own ways of adjusting to the opportunities and dislocations of the new information and communications technologies. Thus, within the broader growth and competitiveness agenda, we emphasize greater labour mobility and flexibility in service product markets within and across countries as being especially important.

## 1. European productivity: 1950-2009

Europe's growth performance relative to that of the United States since 1950 can be usefully divided into three periods: 1950-1973, 1973-1995, and 1995-2007. The comparative European experience in GDP per capita and in GDP per hour is illustrated in Figure 1. The measures are compared relative to the US levels and are adjusted for differences in relative price levels using the GDP-based purchasing power parities for 2005 from the OECD. We also added the latest years for which data are available, 2008 and 2009, on the basis of provisional national accounts estimates.

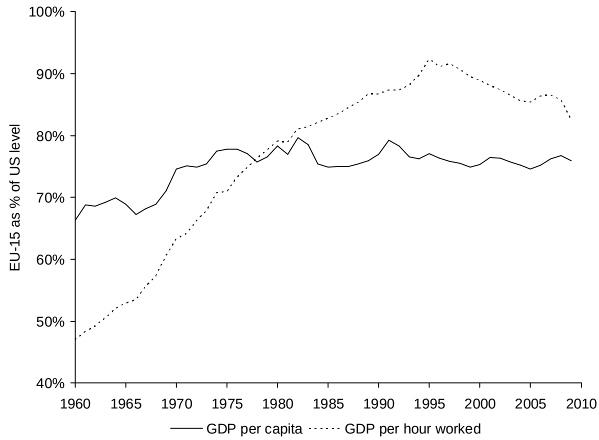
### *1.1 European catch-up: 1950-1973*

During the first period, from 1950 to 1973, rapid labour productivity growth in the European Union went together with catching-up in terms of per capita income levels with the United States. The reasons for this dual catching-up process during the 1950's and 1960's have been extensively discussed in the literature and can broadly be divided into three groups:

technology imitation, new institutions, and business governance and management models (for example, Boltho, 1982; Chandler, 1990; Crafts and Toniolo, 1996; Eichengreen, 2007).

Imitation of technology and incremental innovation allowed European countries to speed up growth and productivity quite rapidly following the Depression of the 1930's and the devastation of Europe's economies during World War II. Many European countries could draw upon their legacy as industrializing nations during the nineteenth and early twentieth century. Compared to other parts of the world, Europe after World War II already had a relatively well-educated population and a strong set of institutions for generating human capital and financial wealth, which allowed a rapid recovery of investment and absorption of new technologies developed elsewhere, notably in the United States.

**Figure 1. Total economy GDP per hour worked and GDP per capita in EU-15, 1960-2009 (relative to the United States)**




Source: The Conference Board Total Economy Database, January 2011, http://www.conference-board.org/data/economydatabase/.

Notes: EU-15 refers to the 15 countries constituting the European Union before 2004 and includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom. The EU has expanded to include ten new member states mainly in Central and Eastern Europe in 2004 and another two in 2007; the new members are not included here. Relative levels are based on purchasing power parities for GDP for 2005 from the OECD.

This process was strengthened by the emergence of a new set of institutions in the area of wage bargaining (Eichengreen, 2007). Although there were important differences between countries, essentially these arrangements involved limiting wage demands in exchange for a rapid redeployment of profits for investment. Through this arrangement, a consensus was developed between workers and capitalists that benefited both productivity and per capita income. In addition, European capital markets favoured the emergence of large “national champion” companies while at the same time (notably in Germany) supporting a strong system of small- and medium-sized enterprises. In several northwestern European countries, the education system tended to emphasize technical and vocational training. These characteristics of European institutions largely lasted until the end of the 1960’s, after which labour markets became increasingly tight, leading to substantially higher wage demands.

### *1.2 The productivity slowdown: 1973-1995*

The “golden age” of post-World War II growth came to an end rather abruptly in the early 1970’s, followed by a period of significantly slower growth lasting almost two decades on both continents (Maddison, 1987). Table 1 shows that while US GDP growth slowed down from 3.9 per cent on average per year in the period 1950-1973 to 2.9 per cent in the period 1973-1995, EU-15 growth slowed down substantially more from 4.9 per cent in the period 1950-1973 to only 2.2 per cent in the period 1973-1995. However, average growth rates of per capita income between the United States and the EU-15 became quite similar at 1.8 per cent (for the EU) and 1.9 per cent (for the US) between 1973 and 1995. Further details on the growth slowdown during this period are provided by Crafts and Toniolo (1996), Baily and Kirkegaard (2004) and Eichengreen (2007).

Looking back at Figure 1, one striking observation is that while per capita income in Europe hovered around between 70 to 80 per cent of the US level between 1973 and 1995, the labour productivity gap between Europe and the United States continued to narrow. Indeed, average annual labour productivity growth in the EU-15 was still more than twice as fast as in the United States, at 2.7 per cent in the EU-15 against 1.3 per cent in the United States from 1973 to 1995. Thus, the labour productivity gap virtually closed from more than 30 percentage points in 1973 to only 7.5 percentage points in 1995, as shown in Table 2. In some European countries, including Belgium, France, Germany, and the Netherlands, GDP per hour worked was even higher than the US level in 1995. In Europe, the combination of an unchanged gap in per capita income and a narrowing gap in labour productivity was related – by accounting identity – to a decline in labour force participation rates and a fall in working hours per person employed. Working hours per capita in the European Union countries declined from more than 10 per cent above the US level in 1973 to 84 per cent of the US level by 1995, as shown in Table 2.

**Table 1. Average annual growth rates of GDP, GDP per capita, and GDP per hour worked, EU-15 and United States, 1950-2007 (in per cent)**

| | GDP | GDP per capita | GDP per hour worked |
|---|---|---|---|
| ***1950-1973*** | | | |
| EU-15 | 4.9 | 4.2 | 4.9 |
| US | 3.9 | 2.5 | 2.6 |
| ***1973-1995*** | | | |
| EU-15 | 2.2 | 1.9 | 2.7 |
| US | 2.9 | 1.8 | 1.3 |
| ***1995-2007*** | | | |
| EU-15 | 2.4 | 2.0 | 1.5 |
| US | 3.2 | 2.1 | 2.1 |

Source: Calculations based on The Conference Board Total Economy Database, January 2011, at http://www.conference-board.org/data/economydatabase/.

Note: See Figure 1. The growth rates are presented as differences in the log of the levels of each variable instead of a percentage change in the actual level in order to facilitate aggregation to regional averages and a decomposition of growth sources.

A substantial literature has explored why Europe's labour market institutions have led to less work, in particular during the period 1973-1995. Blanchard (2004) stresses how the trade-off between preferences for leisure and work developed differently in Europe and the United States. Prescott (2004) estimates that income taxes account for virtually all the difference in labour participation rates across European countries. Nickell (1997) shows that besides high payroll taxes, other labour market issues, such as generous unemployment benefits, poor educational standards at the bottom, and high unionization with little coordination also play an important role in accounting for Europe's rise in unemployment since the mid 1970's. Europe's welfare state rapidly expanded in the 1970's, causing an increase in labour cost, a strong bias towards insiders in the labour market, and an increase in structural unemployment, in particular among youth and elderly workers.

One result of Europe's slowing growth in labour input was a rapid increase in capital intensity, as the rise in wages supported the substitution of capital for labour. Table 2 shows that Europe's capital stock per hour worked was at 75 per cent of the US level in 1973, but it was slightly ahead of the US level by 1995. Some European countries had a substantially higher capital stock per hour worked than the US in 1995, including Austria, Belgium, Finland, France, Germany, and the Netherlands. As a result, the high labour productivity levels in the European Union by the mid 1990's should be interpreted with caution. Economists draw a distinction between labour productivity, which can be measured by GDP per hour worked, and multifactor productivity, which relates to the level of output after accounting for labour as well as capital inputs. As will be argued in more detail below, even though Europe experienced a relatively strong growth in labour productivity, the growth in multifactor productivity was much lower. This indicates that Europe's higher labour productivity growth during this period may not have been so much the result of catch-up, access to superior technology, or even faster innovation, but can be largely attributable to accumulated labour market rigidities.

**Table 2. Levels of EU-15 relative to the United States, 1950-2007 (in per cent)**

| | 1950 | 1973 | 1995 | 2007 |
|---|---|---|---|---|
| GDP per capita | 51.5 | 75.4 | 77.1 | 76.2 |
| Hours worked per capita | 126.2 | 110.9 | 83.5 | 88.1 |
| GDP per hour worked | 40.8 | 68.0 | 92.4 | 86.6 |
| Capital input per hour worked* | | 75.3 | 103.6 | 103.0 |

Source*:* Calculations based on The Conference Board Total Economy Database, January 2011, at http://www.conference-board.org/data/economydatabase/

Note*:* Output and capital levels are converted by GDP purchasing power parities for 2005. * Measured as capital stock per hour worked.

### *1.3 Europe's falling behind: 1995-2007*

Since the mid 1990's, there has been a dramatic change in the patterns of productivity growth in Europe and the United States. In the United States, average annual labour productivity growth accelerated from 1.3 per cent during the period 1973-1995 to 2.1 per cent during 1995-2007. Comparing the same two time periods, annual labour productivity growth in the European Union declined from 2.7 to 1.5 per cent. By 2007, GDP per hour worked in the EU was more than 10 percentage points below the US level, while the capital intensity levels remained relatively high, above the US levels, suggesting that the adjustment productivity was mainly made through a slowdown in multifactor productivity growth (Table 2).

The slowdown in labour productivity may be related to the rapid growth in labour input in many European countries. During the late 1980's and 1990's, several European countries introduced labour market reforms and instigated active labour market interventions to bring long-term unemployed people to work and raise the participation rate. The slowdown in labour productivity growth and the decline in relative capital intensity in Europe since 1995 suggest the possibility that just as limited employment growth accompanied higher labour productivity in Europe in the 1973-1995 period, that pattern might have reversed itself in the more recent time period (Gordon, 2004). While in the short run, labour productivity growth might decline due to the dampening of real wage growth and the consequent reduction in the rate of substitution of capital for labour, it is unlikely that the elasticity of labour input on productivity would be large in the medium and long term. According to Blanchard (2004), an employment-productivity trade-off would only exist under the assumption of stagnant output growth, which is an unlikely assumption for the medium and long run. Indeed, despite slowing productivity growth, the European Union has not experienced a slowdown in GDP growth since 1995. A related argument that has been suggested for the productivity slowdown is that the rise in employment has raised the share of low-skilled workers in the workforce. However, there are no signs of a slowdown in the skill level of the labour force in Europe. On the contrary, the average skill-level of the employed labour force has continued to improve since the mid-1990's. Thus, the labour market is unlikely to be the only or even the main explanation for the slowdown in productivity growth.

When put into a comparative perspective, the productivity slowdown in Europe is all the more disappointing as US productivity growth has accelerated since the mid 1990's. The causes of the strong US productivity resurgence have been extensively discussed (see, for example, Jorgenson et al., 2008). In the mid 1990's, there was a burst of higher productivity in industries producing information and communications technology equipment, and a capital-deepening effect from investing in information and communications technology assets across the economy. In turn, these changes were driven by the rapid pace of innovation in information and communications technologies, fuelled by the precipitous and continuing fall in semiconductor prices. With some delay, arguably due to the necessary changes in production processes and organizational practices, there was also a multifactor productivity surge in industries using these new information and communications technologies – in particular in market services industries (Triplett and Bosworth, 2006).

In Europe, the advent of the knowledge economy has been slower since the mid 1990s. In the next section, we exploit the EU-KLEMS database on industry-level growth accounts to develop a better view of how inputs and productivity have contributed to the change in the growth performance of European countries since 1995, in particular in comparison with the United States.

### *1.4 Divergence during the Great Recession: 2007-2009*

In 2008/2009, advanced economies were hit by the deepest recession since the 1930's. In 2008, GDP growth in the EU-15 slowed to 0.3 per cent and then fell dramatically to –4.3 per cent in 2009. The United States experienced a standstill in GDP growth in 2008, but contracted less severely than the EU-15, at –2.6 per cent, in 2009 (Table 3). Traditionally, productivity is pro-cyclical, which implies that in a downturn, labour productivity growth slows down or even declines since initially, output growth slips more than employment growth. Businesses typically hold onto their staff (labour hoarding) and equipment at least for a while to see how the economy will develop before laying off people or scrapping machines. Adjustments are usually made through lowering the capacity utilization and reducing the working hours of staff.

This typical pattern of pro-cyclicality in labour productivity can be observed for Europe's performance during the recession. The EU-15 showed an average slowdown in labour productivity growth of –0.7 per cent from 2007-2009. The US, however, showed an atypical increase in productivity of 3.2 per cent per year over the same period. As the two previous US recessions in 1990/1991 and 2001/2002 also exhibited counter-cyclicality going into the recession, various explanations have been put forward for the change in the relationship among output, productivity, and employment in the US. These range from labour-market-based explanations, pointing at increased flexibility in hiring and firing, technology-based explanations pointing at the role of ICT in continuing productivity increases during recessions, and explanations related to financial market incentives and executive compensation. The latter may have stimulated short-term gains in performance relative to a long-term erosion in the sources of growth available to US companies.

**Table 3. Growth rates of GDP, GDP per capita, total hours worked and GDP per hour worked, EU-15 and the United States, 2007-2009 (in per cent)**

| | GDP | GDP per capita | Total hours worked | GDP per hour worked |
|---|---|---|---|---|
| ***2007*** | | | | |
| EU-15 | 2.8 | 2.3 | 1.6 | 1.1 |
| US | 1.9 | 0.9 | 1.0 | 1.0 |
| ***2008*** | | | | |
| EU-15 | 0.3 | -0.2 | 0.4 | -0.1 |
| US | 0.0 | -0.9 | -0.8 | 0.8 |
| ***2009*** | | | | |
| EU-15 | -4.3 | -4.6 | -3.0 | -1.3 |
| US | -2.6 | -3.5 | -5.0 | 2.5 |
| ***2007-09*** | | | | |
| EU-15 | -2.0 | -2.4 | -1.3 | -0.7 |
| US | -1.3 | -2.2 | -2.9 | 1.6 |

Source*:* Calculations based on The Conference Board Total Economy Database, January 2011, at http://www.conference-board.org/data/economydatabase/.

While there may be no unique explanation for the continued labour productivity growth in the US vis-à-vis the typical pro-cyclicality in the European Union during the recession, it should also be pointed out that European countries did not show a unique pattern of productivity growth amongst themselves. For example, in Germany, labour productivity growth declined by 2.4 per cent from 2007 to 2009, as the government and businesses chose to use short-time working schemes and other instruments to dampen the threat of large layoffs. The total hours worked in Germany therefore only fell by 1.4 per cent per year. In contrast, in Spain, large structural labour market problems led to massive layoffs, in particular of temporary and migrant employees in tourism, construction and agriculture, causing a drop in hours worked of 6.3 per cent per year from 2007 to 2009, but a productivity improvement of 3.3 per cent.

Clearly there is no silver bullet to deal with productivity issues during recessions and, ultimately, the long-term strength of the economic structure of an economy, as measured by its industry composition and sources of growth determines its long-term growth potential.

## 2. Growth accounting for Europe and the United States

To assess the contribution of various inputs to GDP growth, we apply the neoclassical growth accounting framework pioneered by Solow (1957) and further developed by Jorgenson and associates (Jorgenson and Griliches, 1967; Jorgenson et al. 1987). Using this framework, measures of output growth can be decomposed into the contributions of inputs and productivity within a consistent accounting framework. This approach allows researchers to assess the relative importance of labour, capital, and intermediate inputs to growth, and to derive measures of multifactor productivity growth. The output contribution of an input is measured by the growth rate of the input, weighted by the income shares of that input. Under neoclassical assumptions, the income shares reflect the output elasticity of each input and, assuming constant returns to scale, they sum to one. The portion of output growth not attributable to inputs is the multifactor productivity residual. Multifactor productivity indicates the efficiency with which inputs are being used in the production process, and includes pure technological change, along with changes in returns to scale and in mark-ups. Multifactor productivity, as a residual measure, also includes measurement errors and the effects from unmeasured output and inputs, such as research and development and other intangible investments, including organizational improvements (Corrado et al. 2009; van Ark et al., 2009).

Our growth decompositions are based on the November 2009 release of the EU-KLEMS database. This database provides harmonised measures of output growth, productivity, employment creation, and capital formation at a detailed industry level for European Union member states, Japan, and the United States from 1980 to 2007. In particular, this database contains unique industry-level measures of the skill distribution of the work force and a detailed asset decomposition of investment in physical capital. Labour input reflects changes in hours worked, but also changes in labour composition in terms of age, gender, and educational qualifications over time. Physical capital is decomposed into six asset categories, of which three are information and communications capital – including information technology hardware, communication equipment, and software – and three are capital that does not involve information and communications technology – machinery and equipment, transport equipment, and nonresidential structures. Residential capital, which does not contribute to productivity gains in any direct way, is excluded from the analysis.

The EU-KLEMS database makes it possible to compare and analyse the contribution of high-skilled labour and information and communications technology capital for labour productivity growth at an industry level between countries. In this chapter, our focus is on the market economy, which means that we exclude health and education services, as well as public administration and defence from the analysis below. This exclusion implies a faster acceleration of output growth in both the European Union and the United States since 1995 than for the total economy reported in the previous section, but the difference in the pace of acceleration between the two regions does not change. Moreover, in the remainder of this discussion, the European Union only includes 10 countries, excluding Greece, Ireland, Luxembourg, Portugal, and Sweden from the original EU-15, because no industry-level accounts going back to 1980 were available for these five countries.

Table 4 provides a summary of the growth contributions of factor inputs and multifactor productivity to labour productivity growth in the market economy in the ten European Union countries and in the United States for the periods 1980-1995 and 1995-2007. When comparing the period before and after 1995, the annual growth rate of output in the European Union accelerates, and the growth differential relative to the United States drops from 1.2 percentage points (2.1 per cent in Europe versus 3.3 per cent in the United States) to 1.0 percentage point (2.5 per cent in Europe versus 3.5 per cent in the United States). As described in the previous section, total hours worked in the European Union grew rapidly after 1995, to some extent making up for the shortfall in the earlier period. In contrast, the growth in total hours worked slowed down substantially in the United States – in particular after 2000 – even though the average growth rate in hours was comparable to that of the European Union between 1995-2007. As a result, labour productivity growth in the US market economy increased significantly as compared to a large slowdown in Europe after 1995.

Table 4 shows that changes in labour composition contributed 0.2-0.3 percentage points to labour productivity growth both in the European Union and the United States during this entire time period. Even though this contribution is small, its positive sign implies that the process of transformation of the labour force to higher skills has proceeded at roughly equal rates in Europe and the United States, thus confirming the above observation that Europe has not raised its share of low-skill workers. Instead, the upward trend in the skill-content of the employees shows that newcomers on the labour market have on average had more schooling than the existing labour force.

Concerning the total contribution of capital deepening to labour productivity growth, measured by capital services per hour, Table 3 shows somewhat larger differences between the European Union and the United States as compared to labour composition. The contribution of capital deepening declined in Europe while rising in the United States between the two time periods. The growth contribution of information and communications technology (ICT) capital per working hour has been lower in Europe than in the United States and, since 1995, it has accelerated more slowly (Timmer and van Ark, 2005). This slower uptake in deepening of ICT capital is, in part, related to the overall decline in capital-labour ratios across Europe since the mid 1990's, as there has been a rapid growth in European employment.

The largest difference between the European Union and the United States shown in Table 4 is in the contribution of multifactor productivity growth. Whereas multifactor productivity growth in the United States accelerated by half a percentage point from 0.7 per cent in 1980-1995 to 1.2 per cent in 1995-2007, it fell by the same degree from 1.1 to 0.6 per cent between these two periods in the European Union. As a residual measure, multifactor productivity has multiple interpretations, but in some way, it does reflect the overall efficiency of the production process. Its reduced growth rate is therefore a major source of concern across Europe.

It should be stressed that the growth differential between the EU and the US was especially large between 1995 and 2004. The differences became significantly smaller after 2004 when Europe saw a slight acceleration in multifactor productivity growth in the market economy from 0.4 per cent (in 1995-2004) to 1.2 per cent (in 2004-2007) due to a cyclical peak, whereas US MFP growth began to slow from 1.4 per cent to 0.4 per cent between the two periods.

**Table 4. Contributions to growth of real output in the market economy, European Union and the United States, 1980-2007 (annual average growth rates, in percentage points)**

| | European Union* | | United States** | |
|---|---|---|---|---|
| | 1980-1995 | 1995-2007 | 1980-1995 | 1995-2007 |
| 1. Growth rate of market economy output | 2.1 | 2.5 | 3.3 | 3.5 |
| 2. Hours worked | -0.5 | 0.8 | 1.3 | 0.9 |
| 3. Labour productivity | 2.5 | 1.6 | 2.0 | 2.6 |
| ***Contributions from*** | | | | |
| 4. Labour composition | 0.3 | 0.2 | 0.2 | 0.3 |
| 5. Capital services per hour | 1.2 | 0.9 | 1.0 | 1.2 |
| 6. ICT capital per hour | 0.4 | 0.5 | 0.7 | 0.9 |
| 7. Non-ICT capital per hour | 0.8 | 0.4 | 0.3 | 0.3 |
| 8. Multifactor productivity | 1.1 | 0.6 | 0.7 | 1.2 |
| ***Contribution of the knowledge economy to labour productivity growth (4)+(6)+(8)*** | 1.8 | 1.3 | 1.7 | 2.4 |

Source*:* EU-KLEMS database, November 2009; see O'Mahony and Timmer (2009).

Note*:* * excludes 5 member states of EU-15: Greece, Ireland, Luxembourg, Portugal and Sweden; Data for the European Union refers to ten countries: Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Spain, and the United Kingdom. ** based on US old standard industrial classification. "ICT" is information and communications technology.

When looking at the growth accounts from the perspective of the emerging knowledge economy, one might focus on the summed contributions of three factors: direct effects of investments in information and communication technology; changes in the labour composition mostly driven by a greater demand for skilled workers; and multifactor productivity growth. MFP growth might – as indicated above – include the impact of intangible investments such as organizational changes related to the use of information technology. Table 4 shows that the combined contribution of these three factors to labour productivity growth declined by 0.5 percentage points in Europe between the two time periods, from 1.8 percentage points in 1980-1995 to 1.3 percentage points in 1995-2007. In contrast, in the US economy, the contribution of these three knowledge economy components increased from 1.7 percentage points in 1980-1995 to 2.4 percentage points in 1995-2007.

There is a large variation in labour productivity growth across European countries. Similar to the rows in Table 4, the first column of Table 5 shows the growth rate of output for 10 European countries over the 1995-2007 time period. The second and third columns divide that growth in output into changes in hours worked and changes in output per hour, or labour productivity. Columns 4-7 divide the growth in labour productivity into the contributions from four factors: changes in labour composition; investments in information and communication technology capital; other types of physical capital; and multifactor productivity.

One key observation to be drawn from this table is that the main difference in labour productivity growth between individual European economies and the United States is to be found in multifactor productivity, not in differences in the intensity of the production factors. Indeed, the bottom row shows that the standard deviation for multifactor productivity growth across the set of countries is by far the largest, ranging from –0.6 per cent in Spain to + 2.8 per cent in Finland. By way of illustration, the difference in the contribution of capital deepening in information and communications technologies between a high investor like the United States and a low investor like Italy explains 0.5 percentage points of a labour productivity growth difference of 2 percentage points between those two countries during 1995-2007. The remaining 1.5 percentage point difference is (more than) accounted for by the differences in multifactor productivity growth. Differences in multifactor productivity also seem to have driven the divergence in labour productivity between European countries. In Belgium, multifactor productivity growth has been close to zero per cent per year and in Denmark, Italy, and Spain, it is even negative. Only Finland exceeds the US growth rate of multifactor growth in the market economy, and Finland is a special case that will be discussed in more detail in the next section.

How should we explain the large differences in multifactor productivity growth across countries? In the next section, a division of the market economy measures by industry focuses the attention on the performance of the market services sector.

**Table 5. Contributions to growth of real output in the market economy, EU economies and the United States, 1995-2007 (annual average growth rates, in percentage points)**

| | Growth rate of output | Output contribution from | | Labour productivity contributions from | | | | Labour productivity contribution of the knowledge economy |
|---|---|---|---|---|---|---|---|---|
| | | Hours worked | Labour productivity | Labour composition | ICT capital per hour | Non-ICT capital per hour | MFP growth | |
| | 1 = 2+3 | 2 | 3=4+5+6+7 | 4 | 5 | 6 | 7 | 4+5+7 |
| Austria | 2.8 | 0.6 | 2.2 | 0.1 | 0.5 | 0.0 | 1.5 | 2.2 |
| Belgium | 2.5 | 0.8 | 1.7 | 0.2 | 0.9 | 0.4 | 0.1 | 1.3 |
| Denmark | 2.3 | 1.3 | 1.0 | 0.1 | 1.0 | 0.1 | -0.1 | 0.9 |
| Finland | 4.6 | 1.3 | 3.3 | 0.1 | 0.5 | –0.1 | 2.8 | 3.5 |
| France | 2.5 | 0.5 | 2.0 | 0.3 | 0.3 | 0.4 | 0.9 | 1.6 |
| Germany | 1.4 | -0.3 | 1.7 | 0.0 | 0.5 | 0.5 | 0.7 | 1.2 |
| Italy* | 1.5 | 1.1 | 0.4 | 0.1 | 0.2 | 0.4 | -0.4 | 0.0 |
| Netherlands | 3.1 | 1.0 | 2.1 | 0.4 | 0.5 | 0.0 | 1.1 | 2.0 |
| Spain | 3.7 | 3.0 | 0.6 | 0.4 | 0.4 | 0.5 | -0.6 | 0.1 |
| United Kingdom | 3.2 | 0.6 | 2.6 | 0.4 | 0.8 | 0.4 | 1.0 | 2.2 |
| European Union** | 2.5 | 0.8 | 1.6 | 0.2 | 0.5 | 0.4 | 0.6 | 1.3 |
| United States*** | 3.5 | 0.9 | 2.6 | 0.3 | 0.9 | 0.3 | 1.2 | 2.4 |
| standard deviation**** | 0.9 | 0.8 | 0.9 | 0.1 | 0.3 | 0.2 | 1.0 | 1.0 |

Source: Calculations based on the EU-KLEMS database, November 2009; see O'Mahony and Timmer (2009).

Note: "ICT" is information and communications technology. "MFP" is multifactor productivity. * Data for Italy excludes agriculture and private households. ** Data for the European Union excludes five member states of EU-15: Greece, Ireland, Luxembourg, Portugal and the numbers may not sum exactly due to rounding. *** based on old standard industrial classification for the United States **** Standard deviation for EU countries and the United States.

## 3. Structural change and sectoral productivity growth

During the postwar period, Europe has experienced a large shift of production and employment from manufacturing and other goods-producing industries (such as agriculture and mining) towards services. Market services include a wide variety of activities, ranging from trade and transportation services, to financial and business services, and also hotels, restaurants, and personal services. Over the period 1980-2007, the share of labour input going to manufacturing has typically declined by one-third or more in most countries. Market services now account for almost half of the market economy employment in all countries and the share of total labour hours going to market services is not much lower in Europe than in the United States. While there are differences across European countries, even in Germany, a country in which manufacturing traditionally plays an important role, the number of hours worked in market services is now more than 2.5 times larger than in manufacturing.

The growing importance of market services is the result of a number of interacting forces (Schettkatt and Yocarini, 2006). A higher per capita income leads to a higher demand for services. There is also an increasing marketization of traditional household production activities, including services like dining outside the home, cleaning, and care assistance. Finally, many manufacturing firms are outsourcing service-related activities, including business services, trade, and transport activities. Whatever the underlying causes of the shift from manufacturing to services, it has important implications for productivity growth. Traditionally, manufacturing activities have been regarded as the locus of innovation and technological change and thus, being the central source of productivity growth. For example, more productive manufacturing was the key to post-World War II growth in Europe through a combination of economies of scale, capital intensification, and incremental innovation. More recently, rapid technological change in computer and semiconductor manufacturing seemingly reinforces the predominance of innovation in the manufacturing sector. In contrast, the increasing weight of services in output was thought to slow aggregate productivity growth. Baumol (1967) called this the “cost disease of the service sector.” The diagnosis of the disease argues that productivity improvements in services are less likely than in goods-producing industries because most services are inherently labour-intensive, making it difficult to substitute capital for labour in service industries. Although Baumol originally mainly referred to services activities like education, health, and public services, it was widely believed to hold for many other services sectors as well. This hypothesis has recently been disputed in the literature (for example, Triplett and Bosworth, 2006) and, as the following discussion will show, cannot be supported by the evidence from the EU-KLEMS data.

**Table 6. Major sector contribution to average annual labour productivity growth in the market economy, 1995-2007 (annual average growth rates, in percentage points)**

| | Market economy 1=2+3+4+5 | Contributions from | | | |
|---|---|---|---|---|---|
| | | ICT production 2 | Goods production 3 | Market services 4 | Reallocation* 5 |
| Austria | 2.2 | 0.3 | 1.7 | 0.2 | -0.1 |
| Belgium | 1.7 | 0.3 | 0.9 | 0.6 | -0.1 |
| Denmark | 1.0 | 0.3 | 0.4 | 0.4 | -0.1 |
| Finland | 3.3 | 1.7 | 1.3 | 0.5 | -0.1 |
| France | 2.0 | 0.4 | 0.8 | 0.7 | 0.0 |
| Germany | 1.7 | 0.5 | 0.9 | 0.4 | 0.0 |
| Italy | 0.4 | 0.2 | 0.2 | 0.0 | -0.1 |
| Netherlands | 2.1 | 0.4 | 0.6 | 1.2 | -0.2 |
| Spain | 0.6 | 0.1 | 0.2 | 0.3 | -0.1 |
| UK | 2.6 | 0.5 | 0.7 | 1.6 | -0.2 |
| EU** | 1.6 | 0.4 | 0.7 | 0.6 | -0.2 |
| US*** | 2.6 | 0.8 | 0.3 | 1.8 | -0.2 |

Source: Calculations based on EU-KLEMS database, November 2009; see O'Mahony and Timmer (2009).

Note: The reallocation effect in the last column refers to labour productivity effects of reallocations of labour between sectors. The European Union aggregate refers to ten countries in the table. Information and communications technology production includes manufacturing of electrical machinery and post and telecommunications services. Goods production includes agriculture, mining, manufacturing (excluding electrical machinery), construction, and utilities. Market services include distribution services; financial and business services, excluding real estate; and personal services. The numbers may not sum exactly due to rounding.

To evaluate the effect of structural changes on productivity growth, we need to look at the contributions of individual sectors to the aggregate economy. Table 6 shows overall labour productivity growth for the market economy split into contributions from labour productivity growth in the ICT production sector (including production of electrical machinery and telecommunication services), goods production (including agriculture, mining, manufacturing other than electrical machinery, utilities, and construction), and the market services sector (including trade, hotels and restaurants, transport services, financial and business services, and social and personal services), each weighted by its share of value added, along with an adjustment in the final column for the reallocation of hours between industries with different productivity growth rates or levels.

Table 6 shows that slow productivity growth in market services is not a universal truth, even among advanced countries with large service sectors. First, productivity growth in market services has been much faster in the United States than in Europe. At an average annual labour productivity growth rate of 1.2 per cent, market services only contributed 0.6 percentage points to labour productivity growth in Europe from 1995-2007. In contrast, labour productivity in market services increased at 3.0 per cent in the United States, contributing 1.8 percentage points to US productivity growth. Second, two countries in Europe – the Netherlands and the United Kingdom – also showed a rapid productivity growth in market services. Market services in the United Kingdom contributed almost as much to aggregate labour productivity growth as in the United States, mainly due to a strong performance in trade and business services industries. Incidentally, market services also appear to exhibit a rapid productivity growth in other Anglo-Saxon economies, such as Australia and Canada (Inklaaret al., 2007). In contrast, Italy and Spain show almost zero contributions from market services to aggregate labour productivity growth. Previous studies on the growth differential between Europe and the United States also stressed the differentiating role of market services (O'Mahony and van Ark, 2003; Losch, 2006; Inklaar et al., 2008).

The importance of market services for the productivity growth gap between Europe and the United States dwarfs the differences for other major sectors. Even though the United States has a somewhat bigger share in ICT–producing sectors, the productivity growth rates in these sectors are not dramatically different. As a result, the effect on the aggregate growth differential is only 0.4 percentage points (0.8 per cent in the United States as compared to 0.4 per cent in Europe). Goods production seems to be somewhat more important for aggregate productivity growth in Europe than in the US. The contribution from labour productivity growth in goods production in Europe is about the same as that of market services, despite the relative size of the former of only one-third of market services value added. For example, in France and Germany, manufacturing industries like machinery and car production are still important sources of productivity growth. In Spain and Italy, lackluster performance is not only due to slow growth in market services, but also in manufacturing, as traditional labour-intensive sectors have faced a particularly tough challenge from increasing low-wage competition from Eastern Europe and China.

**Table 7. Contributions of sectors to average annual labour productivity growth in market services, 1980-2007 (in percentage points)**

| | European Union* | | United States** | |
|---|---|---|---|---|
| | **1980-1995** | **1995-2007** | **1980-1995** | **1995-2007** |
| Market services labour productivity | 1.4 | 1.2 | 1.6 | 3.0 |
| ***Contributions from:*** | | | | |
| *Distribution services contribution* | 1.1 | 0.7 | 1.2 | 1.3 |
| of which from factor intensity growth | 0.4 | 0.4 | 0.4 | 0.5 |
| of which from multifactor productivity growth | 0.7 | 0.3 | 0.8 | 0.8 |
| *Finance and business services contribution* | 0.1 | 0.3 | 0.0 | 1.3 |
| of which from factor intensity growth | 0.5 | 0.4 | 0.8 | 1.0 |
| of which from multifactor productivity growth | -0.4 | -0.1 | -0.8 | 0.3 |
| *Personal services contribution* | -0.1 | -0.1 | 0.2 | 0.2 |
| of which from factor intensity growth | 0.1 | 0.1 | 0.1 | 0.1 |
| of which from multifactor productivity growth | -0.2 | -0.2 | 0.1 | 0.1 |
| *Contribution from labour reallocation* | 0.3 | 0.2 | 0.2 | 0.2 |

Source: Calculations based on EU-KLEMS database, November 2009; see O'Mahony and Timmer, 2009).

Note: European Union aggregate refers to 10 countries, as listed in Table 5. Factor intensity relates to the total contribution from changes in labour composition and in capital deepening of information and communications technology (ICT) and non-information and communications technology (non-ICT) assets. The reallocation effect refers to the impact of changes in the distribution of labour input between industries on labour productivity growth in market services. The numbers may not add up due to rounding.

A more in-depth focus on market services reveals that cross-Atlantic growth differences were especially large in distributive trade and financial and business services. In Table 7, we focus on the contribution of three major groups of market services industries – namely distributive trade (including retail and wholesale trade, and transport services); financial and business services; and personal services (including hotels and restaurants, and personal, community, and social services) – to labour productivity growth in aggregate market services. In Europe, the distribution sector contributed 0.7 percentage points to average annual labour productivity growth in market services from 1995 to 2007, as compared to 1.3 percentage points in the United States. In finance and business services, the gap was even bigger, at a 0.3 percentage point contribution in Europe relative to 1.3 percentage points in the United States. Drilling more deeply into the data, it turns out that for both sectors, multifactor productivity, and not factor intensity, was the key to the productivity growth differential between Europe and the United States. Differences in "factor intensity", which include the total contribution from changes in labour composition and deepening of all types of capital, appear very small. The fuelling of US multifactor productivity growth from trade, finance, and business services is confirmed in studies by Jorgenson et al. (2005) and Triplett and Bosworth (2006).

Recently, the release of a comprehensive revision of the GDP by industry data for the United States by the Bureau of Economic Analysis has led to significant changes in the output estimates for the United States, especially in the services sector. This raises the question whether the productivity advantage for services will still be observed in the latest data. A comparison of the data suggests that the revision of labour productivity growth rates for all market services, taken together, is small for the past decade, that is, 2.8 per cent pre-revision as compared to 2.6 per cent post-revision from 1998-2007 (EU KLEMS, November 2009; BEA, 2010). However, for some industries, the adjustments are larger, especially for the retail industry which is an important driver of the US productivity advantage. Moreover, it is generally well known that measurement problems in the finance and insurance sectors could possibly affect the productivity comparisons for that sector as well. Therefore, we analyse the dynamics of the productivity story for those two industries in more detail below.

### *3.1 The productivity dynamics of the retail sector*

Because multifactor productivity growth represents a multitude of factors which are not explicitly measured in a growth accounts framework, it is useful to look at what might lie behind this growth. While the factors may differ across sectors, the example of the retail sector may serve as an illustration of the complex interactions among productivity, investment, and regulations. Over the past 25 years, the retail sector has undergone a substantial transformation due to benefits from the increased use of information and communications technology, commonly referred to as the "lean retailing system" (Abernathy et al., 1999). The retail industry has changed from a low-tech industry where workers mainly shift boxes from the producer to the consumer depending on availability in stock, into an industry whose main activity is trading information by matching the production of goods and services to customer demand on a continuous basis. Various studies, including McKinsey Global Institute (2002), Baily and Kirkegaard (2004), Gordon (2004), and McGuckin et al. (2005), have discussed the reasons for superior performance in the US retail industry relative to Europe.

While there is significant evidence of a faster rise in ICT capital in the US retail sector as compared to Europe, the productivity impact of the greater use of barcode scanners, communication equipment, inventory tracking devices, transaction processing software, and similar equipment may be understated when solely focusing on the contribution of investment as directly measured in growth accounts. The use of ICT also provides indirect benefits for growth as measured by multifactor productivity by increasing the potential for other kinds of innovation. These innovation effects should in part be realized through "softer" innovations, such as the invention of new retail formats, service protocols, labour scheduling systems, and optimized marketing campaigns (McKinsey Global Institute, 2002).

Others have emphasized the role of "big box" formats, as most notably exemplified by the emergence of Wal-Mart, as the engine of productivity growth in US retailing (Basker, 2007). From this perspective, Europe's lagging behind in productivity is due to more restrictive regulations like store-opening hours; to land zoning and labour markets; and to cultural differences that inhibit a rapid increase in the market share of new large-scale retail formats. These new large-scale retail formats have been a main driver of growth in the United States, both because of increased competitive pressures on incumbent firms and the higher productivity levels of new entrants (Foster et al., 2006). In addition, deregulation in upstream industries such as trucking in the 1980s was necessary for the lean retailing model to work, because it allowed for more efficient ordering and shipping schedules.

The recent comprehensive revision of the GDP by industry data for the United States has led to a significant downward revision of retail output growth from 3.7 per cent to 2.3 per cent from 1998 to 2007. As a result, labour productivity has been revised from 4.2 to 1.8 per cent. The much slower growth of retail output is partly (one-third of the total adjustment) the result of improved source data and partly (the other two-thirds) the result of a methodological improvement, representing a move to directly measuring the price of retail margins rather than assuming a relation to the sales price. A reclassification of sub-industries between wholesaling and retailing may also have played a role as wholesaling productivity was revised upward. It would be too early, however, to conclude that these methodological changes will have altered the picture of the US productivity advantage in retail over Europe. A recent paper by Inklaar and Timmer (2008) has provided comparative estimates of productivity in the retail industry, using various methodologies including the margin approach now used by the US Bureau of Economic Analysis. Their paper shows that irrespective of the methodology applied, the US productivity advantage in the retail sector compared to Europe is maintained.

### *3.2 The productivity dynamics of the financial services sector*

As a result of the financial crisis which started in 2008, there have been largely anecdotal claims that the financial services sector has perversely driven US output and productivity growth during the previous decade or so. The reasoning has been that a relatively strong emphasis on measuring the rapid increase in the monetized value of financial instruments, especially in securities trading, has inflated output and productivity growth. In practice, however, measured labour productivity growth in financial services has been fairly moderate in both the US and Europe at between 2.5 and 3 per cent. And while productivity growth accelerated in Europe between 2004-2007, it did not in the United States. Part of this is related to the slow growth in insurance output which offset a faster growth elsewhere in the financial sector. In fact, MFP growth in financial services has improved in both regions since 2004.

There is considerable empirical evidence that a better-developed financial system enhances economic growth. Financial institutions can help select the most profitable investment projects. However, at the same time, a more efficient financial system may not stimulate growth beyond a certain level of financial development. In other words, advanced economies may not gain much in terms of growth benefits from a more efficient financial system. Therefore, it remains questionable whether the estimates of productivity growth in the financial sector correctly reflect the performance of the financial sector (Inklaar and Koetter, 2008). Current statistical practices do not seem to be up to the task of adequately measuring output and productivity. Moreover, the measurement practices differ between European countries and the United States.

A recent study by Inklaar and Wang (2010) shows that US bank output series in the national accounts is based on a "constant services per loan/transaction" basis, while in European countries, it is usually based on a "constant services per euro" basis. When comparing the number of transactions in deposits in the US with deflated account balances in European countries, it seems that European growth in this part of the finance sector is downwardly biased. In contrast, house-price deflated mortgages in the US, assuming constant services per mortgage loans, suggest a downward bias relative to the CPI-deflated outstanding mortgage balances in Europe. However, both activities are only a small part of aggregate finance and insurance, and a more comprehensive analysis is needed before a definite judgment on the comparative productivity performance of the financial sector can be reached.

By the same token, the benefits of financial innovation are often called upon in discussions of regulation in the financial sector, but their quantitative importance is generally unknown. Before we are able to measure these factors, there is a greater need for more clarity on what banks and other financial institutions actually provide: is it mostly transaction services; or does it also involve other intermediation services, such as information on for example risk profiles, and to what extent is bearing risk a "productive service"?

## 4. A policy perspective on European productivity growth

Since the mid 1990's, the European Union has experienced a slowdown in productivity growth, at a time when productivity growth in the United States accelerated significantly. The resurgence of productivity growth in the United States appears to have been a combination of high levels of investment in rapidly progressing information and communications technology in the second half of the 1990's, followed by rapid productivity growth in the market services sector of the economy in the first half of the 2000's. Conversely, the productivity slowdown in European countries is largely the result of slower multifactor productivity growth in market services, particularly in trade, finance, and business services. This pattern holds true for Europe as a whole and also for many individual European countries.

If multifactor productivity growth is indeed the metric to maximize, the main question for government policy is how to improve MFP growth. This is a surprisingly difficult question to answer. Even though many countries support actions by institutions to strengthen productivity growth through the implementation of practical measures related to training, shopfloor practices and other organizational innovation, a clear-cut "productivity policy" does not exist. This is due to the fact that productivity growth is an inherent result of a broad range of macroeconomic and structural policies.

One complicating factor for Europe is the projected slowdown in labour input growth during the 2010-2020 period, which is the result of the rapid ageing of the population and the limited immigration of skilled workers. This calls for an even larger emphasis on productivity, meaning that Europe needs to find mechanisms to exploit innovations to achieve greater multifactor productivity growth, especially in services. Unfortunately, the traditional catch-up and convergence model of the 1950's and 1960's may not help Europe get back on track. Since Europe had reached the productivity frontier by the mid 1990's, it may now require a new model of innovation and technological change to make better use of a country's own innovative capabilities (Acemoglu et al., 2006). Arguably, innovations in services are more difficult to diffuse between firms compared to "hard" technologies based in manufacturing. The greater emphasis on human resources, organizational change, and other intangible investments is strongly specific to individual firms. Moreover, as the firm internalizes most of the benefits of such changes, the legitimization for government support such as research and development and innovation subsidies to support "technology" transfer in services is limited. Service activities also tend to be less standardized and more customized than manufacturing production; they depend strongly on the interaction with the consumer and are therefore more embedded in national and cultural institutions. In this situation, the spillover of technologies across firms and nations becomes much more difficult. Recent work by Bloom and Van Reenen (2007) links corporate management practices to productivity. They find significant cross-country differences in corporate management practice, with US firms being better managed than European firms on average, as well as significant within-country differences with a long tail of badly managed firms. In other words, a simple "copying" of practices from other countries – or even from other firms within the same country – is not the most likely way for European service companies to attain greater productivity growth.

The main role for innovation policy in the services sector therefore limits itself to supporting learning processes that can lead to improvements in business practices. In particular, these include practices with regard to the investment in, and management of, intangible capital, including human resources practices and organizational innovations, but also branding and marketing (Corrado et al., 2009). Creating optimal conditions for market exchange or innovation cooperation between service firms and their clients and suppliers is an important vehicle to diffuse best practices. Innovators that depend strongly on their clients may benefit most from the fostering of an innovation culture which also impacts on users and, in the longer run, they profit from investment in education.

But the success of any innovation policy in services, generating a positive impact on productivity growth, largely depends on a complementary approach towards more flexible labour, product, and capital markets that allow resources for innovation to flow to their most productive uses. Crafts (2006) discusses the increasing evidence that restrictive product market regulations, in particular those limiting new entry, hinder technology transfer and have a negative impact on productivity, although most studies relate only to manufacturing industries. The diversity in productivity growth across European countries shows that some countries have been addressing these issues relatively successfully, while others have not. Even though most European countries have begun to make changes to institutional arrangements that increase the flexibility and competitiveness in labour and product markets, such changes vary greatly across countries. The changes that have occurred depend on the size and maturity of the industry, industry concentration, the nature of the education system, the availability of capital for startups, the sophistication of the consumer, and the characteristics of the legislative framework. More research is needed to understand the determinants of the differences in country experiences regarding innovation and regulations, in particular in services industries.

Finally, many service industries in Europe could benefit from a truly single market across Europe, where competition can be strengthened and scale advantages may be realized. Naturally, the European "single market" program has aimed at removing the barriers to free movement of capital, labour, and goods since the 1980's, but the effect on the services industry is generally seen as limited. The present drive in Europe towards a greater openness of service product markets across the European Union may have the potential of increasing productivity growth across Europe in the coming decade.

## Note

This paper is an updated and extended version of my article, with Mary O'Mahony and Marcel Timmer, "The productivity gap between Europe and the United States: Trends and causes", *Journal of Economic Perspectives*, vol 22 (1), Winter 2008, pp. 25-44. The current version includes industry level estimates from the EU-KLEMS growth accounting databases which are updated to 2007, preliminary estimates of the impact of the crisis on the comparative growth performance of the European Union and the US from 2007-2009, and an indicative assessment of the impact of recent revisions in the US National Income and Product Accounts on the estimates presented in this study. I am grateful to Robert Inklaar for his contribution to several parts of this paper, and to Vivian Chen, Ben Cheng and Reitze Gouma for updates on the latest estimates. The research for this paper is based on part of the EU-KLEMS project on Growth and Productivity in the European Union. This project was supported by the European Commission, Research Directorate General as part of the 6th Framework Programme, Priority 8, "Policy Support and Anticipating Scientific and Technological Needs".

## References

Abernathy, F.H., Dunlop, J.T., Hammond, J.H. and Weil, D. (eds.) (1999), A Stitch in Time: Lean Retailing and the Transformation of Manufacturing – Lessons from the Apparel and Textile Industries, Oxford University Press, Oxford.

Acemoglu, D., Aghion, P. and Zilibotti, F. (2006), Distance to frontier, selection, and economic growth, Journal of the European Economic Association 4, 37-74.

van Ark, B., Hao, J.X., Corrado, C. and Hulten, C. (2009), Measuring intangible capital and its contribution to economic growth in Europe, EIB Papers, European Investment Bank 14, 62-99.

Baily, M.N. and Kirkegaard, J.F. (2004), Transforming the European Economy, Institute for International Economics, Washington, DC.

Basker, E. (2007), The causes and consequences of Wal-Mart's growth, Journal of Economic Perspectives 21, 177-198.

Baumol, W.J. (1967), Macroeconomics of unbalanced growth: The anatomy of urban crisis, American Economic Review 57, 415-426.

Blanchard, O. (2004), The economic future of Europe, Journal of Economic Perspectives 18, 3-26.

Bloom, N. and Van Reenen, J. (2007), Measuring and explaining management practices across firms and countries, Quarterly Journal of Economics 122, 1351-1408.

Boltho, A. (1982), The European Economy: Growth and Crisis, Oxford University Press, Oxford.

Bureau of Economic Analysis (2010), Advance GDP by Industry Statistics for 2009 and Revised Statistics for 1998-2008, Washington D.C.

Chandler, A.D. (1990), Scale and Scope. The Dynamics of Industrial Capitalism, Harvard University Press, Cambridge MA.

Corrado, C., Hulten, C. and Sichel, D. (2009), Intangible capital and US economic growth, The Review of Income and Wealth 55, 661-685.

Crafts, N. (2006), Regulation and productivity performance, Oxford Review of Economic Policy 22, 186-202.

Crafts, N. and Toniolo, G. (1996), Economic Growth in Europe since 1945, Cambridge University Press, Cambridge.

Eichengreen, B. (2007), The European Economy Since (1945): Coordinated Capitalism and Beyond, Princeton University Press, Princeton.

Foster, L., Haltiwanger, J. and Krizan, C.J. (2006), Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s, Review of Economics and Statistics 88, 748-758.

Gordon, R.J. (2004), Why was Europe left at the station when America's productivity locomotive departed?, CEPR Discussion Paper 4416.

Inklaar, R. and Koetter, M. (2008), Financial dependence and industry growth in Europe: Better banks and higher productivity, GGDC Research Memorandum, GD-100, University of Groningen.

Inklaar, R. and Timmer, M.P. (2008), Accounting for growth in retail trade: An international productivity comparison, Journal of Productivity Analysis 29, 23-31.

Inklaar, R., Timmer, M.P. and van Ark, B. (2007), Mind the gap! International comparisons of productivity in services and goods production, German Economic Review 8, 281-307.

Inklaar, R., Timmer, M.P. and van Ark, B. (2008), Market services productivity across Europe and the U.S, Economic Policy 53, 141-194.

Inklaar, R. and Wang, J.C. (2010), Real output of bank services: What counts is what banks do, not what they own, mimeo, Groeningen University.

Jorgenson, D.W., Gollop, F.M. and Fraumeni, B.M. (1987), Productivity and US Economic Growth, Harvard University Press, Cambridge, MA.
Jorgenson, D.W., Ho, M.S. and Stiroh, K.J. (2005), Information Technology and the American Growth Resurgence, MIT Press, Cambridge, MA.
Jorgenson, D.W., Ho, M.S. and Stiroh, K.J. (2008), A retrospective look at the US productivity growth resurgence, Journal of Economic Perspectives 22, 3-24.
Jorgenson, D.W. and Griliches, Z. (1967), The explanation of productivity change, Review of Economic Studies 34, 249-283.
Losch, M. (2006), Deepening the Lisbon Agenda: Studies on Productivity, Services and Technologies, Austrian Federal Ministry of Economics and Labour, Vienna.
Maddison, A. (1987), Growth and slowdown in advanced capitalist economies: Techniques of quantitative assessment, Journal of Economic Literature 25, 649-698.
McGuckin, Robert, H., Spiegelman, M. and van Ark, B. (2005), The retail revolution: Can Europe match the US productivity performance?, The Conference Board, Research Report R-1358, New York.
McKinsey Global Institute (2002), Reaching Higher Productivity Growth in France and Germany – Retail Trade Sector, McKinsey Global Institute, Washington, DC.
Nickell, S. (1997), Unemployment and labour market rigidities: Europe versus North America, Journal of Economic Perspectives 11, 55-74.
O’Mahony, M. and van Ark, B. (eds.) (2003), EU Productivity and Competitiveness: An Industry Perspective: Can Europe Resume the Catching-up Process?, Office for Official Publications of the European Communities, Luxembourg.
O’Mahony, M. and Timmer, M.P. (2009), Output, input and productivity measures at the industry level: The EU KLEMS database, Economic Journal, Royal Economic Society 119, F374-F403.
Prescott, E.C. (2004), Why do Americans work so much more than Europeans?, Federal Reserve Bank of Minneapolis Quarterly Review 28, 2-13.
Schettkat, R. and Yocarini, L. (2006), The shift to services employment: A review of the literature, Structural Change and Economic Dynamics 17, 127-147.
Solow, R. (1957), Technical change and the aggregate production function, Review of Economics and Statistics 39, 212-220.
Timmer, M.P. and van Ark, B. (2005), Does information and communication technology drive EU-US productivity growth differentials?, Oxford Economic Papers 57, 693-716.
Timmer, M.P., Inklaar, R., O’Mahony, M. and van Ark, B. (2010), Economic Growth in Europe, A Comparative Industry Perspective, forthcoming, Cambridge University Press, Cambridge.
Triplett, J.E. and Bosworth, B.P. (2006), “Baumol’s disease” has been cured: IT and multifactor productivity in US services industries, in D.W. Jansen (ed.), The New Economy and Beyond: Past, Present and Future, Elgar, Cheltenham.

# Comment on van Ark: Up the hill and down again: A history of Europe's productivity gap relative to the United States, 1950-2009

Matti Pohjola[*]

Bart van Ark's paper is an insightful historical analysis of the fact that slower labour productivity growth in Europe as compared to the United States since the mid-1990s reversed the long-term pattern of convergence. It attributes this "American-European productivity paradox" to Europe's weaker ability to benefit from the knowledge economy and the market service sector, explains it by labour and product market regulation and calls for policy measures that would improve the performance of the market service sector. I will here concentrate on four issues which I find relevant for policy-makers, especially in the Nordic countries.

First, the analysis provides an excellent example of how detailed industry-level data on output and factors of production can be used to account for the sources of output and productivity growth at the industry and aggregate levels. The EU-KLEMS cross-country database, developed by van Ark and his colleagues and applied in this paper, can be used as a model for creating similar national databases as has been done by, for example, Statistics Finland. Given that it is nowadays fashionable to rethink and go beyond GDP as a measure of welfare, the analysis clearly shows how valuable it would be to first integrate national income accounts with EU-KLEMS-type productivity accounts. This would help us understand the

[*] Aalto University School of Economics, matti.pohjola@aalto.fi.

role that technology plays in creating the growth of income and, ultimately, of welfare.

Second, to assess the economic performance of the European Union, it is useful to analyse the EU countries as a single unit. But, given that much of policy-making still takes place at the national level, it is also interesting to consider the causes for the observed large variation in productivity growth across European countries. This is done by measuring the contributions of market services to productivity growth and showing that some countries (i.e. the Netherlands and the United Kingdom) have done almost as well as the United States while the others have fallen behind. From the Nordic perspective, it is, however, interesting to observe that the best performing country in terms of both labour and multi-factor productivity growth is Finland, which is treated as a special case in the paper (see van Ark's Table 5). Sweden is not included in the analysis but the EU-KLEMS database reveals that its productivity performance has been equally good and for the same reason as for Finland: the strong performance of ICT goods production. Moreover, Denmark is the country where the contribution from ICT capital has been the largest (see Table 5). The Nordic countries, especially Finland and Sweden, have been able to leverage the ICT revolution to gain global competitiveness. Understanding the reasons would be necessary for predicting productivity growth in the future, especially now when ICT good production is being outsourced to low-wage countries and when ICT services have come to dominate ICT equipment as the source of value.

Third, the paper measures the knowledge economy by summing up the productivity contributions arising from investment in ICT, changes in labour composition and growth in multi-factor productivity. This approach can be criticized for overestimating the impact because it includes ICT capital deepening. The social savings approach, pioneered by Robert Fogel, is different to growth accounting in that it does not attribute any portion of the capital deepening effect to the new technology, arguing that it is attributable to saving. Because the impact of ICT capital deepening has been larger in the United States than in Europe, omitting it from the analysis would narrow, although not totally eliminate, the gap in the contribution of knowledge to labour productivity growth. I very much agree with the view that multi-factor productivity is one of the key performance measures when testing the dynamism of an economic system. But it is not a measure of welfare and, consequently, should not be considered as the metric to be maximized but rather as a means of obtaining higher welfare.

My fourth and final point is the one on which I disagree. I do not think the findings of this paper can be interpreted to support the view that labour and product market regulation is the cause of the slow productivity growth in Europe. The reason is that the growth accounting method applied in the analysis is based on the explicit assumption that all markets are competitive. It is, for example, argued in Section 2 that the increase in Europe's capital intensity in 1973-1995 can largely be attributable to accumulated labour market rigidities. But it can also be attributed to Europe's high savings rates in this and the preceding period. Moreover, there is no correlation between union density rates and labour productivity growth rates in the data used in this paper. In fact, in 2000, the union density rate was the highest in the best performing country, namely Finland. In my view, the policy-makers should take a broader perspective based on the maximization of welfare rather than of multi-factor productivity when making decisions regarding, say, labour market institutions. The welfare gap between Europe and the United States is much smaller than the productivity gap when income equality and leisure are included in the measurement of welfare. In such comparisons, the Nordic and some other European countries often outperform the Anglo-Saxon ones.

# Regulation, resource reallocation and productivity growth*

Jens Matthias Arnold**, Giuseppe Nicoletti*** and Stefano Scarpetta****

## *Summary*

We review theory and evidence on the links between anti-competitive product market regulations, the efficiency in resource allocation and productivity growth. We show large differences of regulations across countries and industries over time. We argue that such differences have played an important role in driving resource allocation and productivity outcomes. Countries and industries where regulatory burdens are lighter have experienced higher GDP per capita and productivity growth rates. Moreover, where such burdens are lighter, reallocation of resources towards the highest productivity firms is stronger. Since the impacts of inappropriate regulations on productivity performance are quantitatively important, reforming such regulations can provide a significant boost to potential growth in OECD economies.

*Keywords:* Productivity, product market regulation, allocative efficiency.
*JEL classification numbers*: D24, E23, K23, L11, L51.

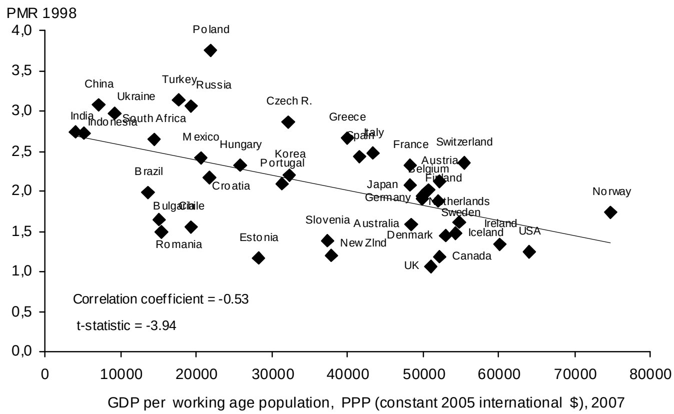
---

* The view contained in this article are those of the authors and do not necessarily reflect those of the OECD or its member countries.

** OECD, jens.arnold@oecd.org.
*** OECD, giuseppe.nicoletti@oecd.org.
**** OECD, stefano.scarpetta@oecd.org.

The analysis of differences in economic performance across countries largely deals with the role played by market rigidities in curbing incentives to innovate and in preventing resources from flowing to the most productive uses. In some cases, rigidities can be directly related to the nature of some economic activities, but they are often induced by inappropriate policies or institutions. This paper focuses on the role of one particular set of policy-induced rigidities, those that are related to regulations that curb product market competition, where competitive forces would be advantageous for society. There is widespread anecdotal evidence that, in countries where policies unduly curb competition, performance is subpar. As an example, Figure 1 suggests a negative and significant correlation between GDP per capita and the OECD summary indicator of anticompetitive product market regulations across a number of OECD and emerging economies. Indeed, countries with more stringent and anti-competitive product market regulations (according to the OECD synthetic indicator) were also those with a relatively lower GDP per capita, and *vice versa*. Needless to say, this is only illustrative because there are many other factors beyond regulations that determine a country's economic performance. Figure 2, however, also shows a negative correlation between Multi-Factor Productivity (MFP) growth and the stringency of product market regulations: countries that had pro-competitive regulations seem to have been more able than others to accelerate productivity growth over the past quarter century. And differences in productivity and productivity growth are the main determinant of cross-country gaps in levels and growth rates of GDP per capita. These simple correlations are sufficiently tight to merit further investigation: to what extent are they driven by the adverse effect of anti-competitive regulations on the ability to efficiently allocate resources and on the incentives to continuously improve efficiency (e.g. via innovation), which are at the heart of the growth process in market economies?

**Figure 1. Anticompetitive product market regulation and GDP per capita[a]**



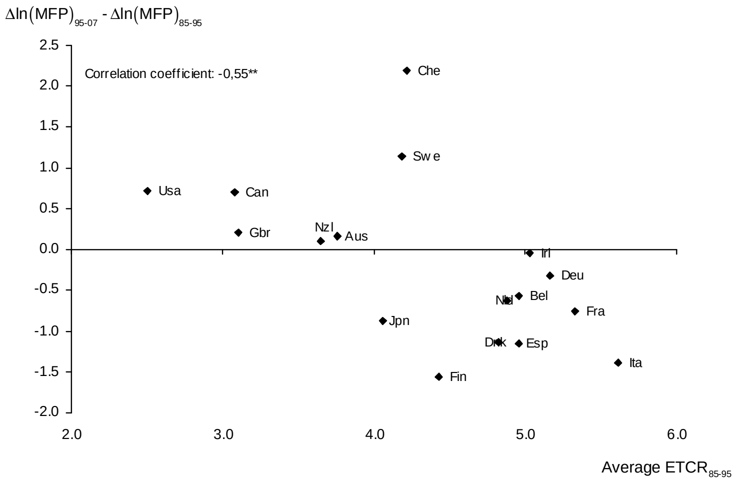

Source: Woelfl et al. ( 2010).

Note: [a] Based on a "simplified" OECD product market regulation (PMR) indicator (see Woelfl et al., 2010). PMR measured in 1998 for OECD countries; 2008 for Chile, Estonia, Israel, Slovenia, Brazil, Russia and China; 2007 for Croatia, Indonesia, South Africa and Ukraine; 2006 for Bulgaria, India and Romania.

In this paper, we address these issues by looking at the link between regulations, resource reallocation and productivity from different angles, i.e. aggregate, sectoral and firm-level. We survey some theory and evidence and, based on existing empirical research, we provide estimates of the extent to which regulations can affect productivity, checking whether this effect is economically relevant. Whenever possible, we discuss how the estimated effects of regulation on performance differ depending on the levels of development, industry characteristics and the relative efficiency of firms in terms of their dynamism or distance from the technological frontier. Indeed, heterogeneous performance (across countries, industries and firms) is a key feature of market economies and the influence of regulation on productivity is likely to differ across countries, industries and firms with different characteristics.

Throughout the paper, we focus for consistency on measures of product market policies that are provided by the OECD. These measures are based on laws and regulations that unduly curb competition and cover both general-purpose and sector-specific areas, such as administrative burdens on start ups and access to networks, respectively. They point to differences in the stringency of regulation that could potentially provide an explanation for differences in productivity developments. We also take into account intersectoral linkages, namely the possibility that sector-specific anticompetitive regulations can have an impact on performance beyond the regulated sector itself, due to the fact that regulated sectors are often important providers of intermediate inputs to other sectors.

**Figure 2. Productivity acceleration and regulation[a]**




Source: OECD Productivity Database and OECD International Regulation Database.

Note: [a] MFP is a multifactor productivity index. Regulation (ETCR) covers historical developments in anticompetitive provisions and industry settings in electricity, transport and communication industries.

The paper is organized as follows. First, we provide a short review of the main channels through which anticompetitive regulations can be expected to affect performance, focusing on their effect on technology adoption, innovation and the allocation of resources to the most productive firms as well as on intersectoral linkages. Second, we illustrate how regulations differ across countries and how they have changed over the past quarter century, pointing out the pervasive regulatory burdens that inappropriate sectoral regulations can impose on the economy as a whole. Third, we look at the cross-country evidence on the regulation-performance nexus, drawing on aggregate, industry-level and firm-level data. We start the analysis with a look at some recent evidence on the correlation between GDP per capita growth and regulation. Then, we turn to industry-level evidence. We show how cross-country productivity growth dispersion and average productivity growth performance can be related to regulation, with a focus on the divide between relatively "deregulated" English-speaking countries and relatively more regulated continental European countries. Finally, we report results relating the efficiency in the allocation of resources across firms and, in particular, the ability of the most dynamic firms to sustain high productivity growth rates, to the underlying regulatory environment.

## 1. How does regulation affect productivity?

Product market regulations, like other regulations, generally address public interest concerns about market failures, including monopoly conditions, externalities and asymmetric information. In this context, product market regulation can promote competition in certain industries by ensuring that market power in natural monopoly segments is not used abusively and by providing the correct incentives to market participants. However, regulatory frameworks may be flawed by several (possibly concurring) factors. Some regulations may drift away from their original public interest aims, resulting in the protection of special interest groups. Second, regulations (and their implementation) sometimes involve costs that exceed their expected benefits, leading to so-called "government failure". Third, technical progress, the evolution of demand and progress in regulatory techniques can make the design of regulations obsolete.

Inappropriate regulations can affect the productivity performance of an economy in many ways. Given the multiple channels and the potentially conflicting effects, it is hard to provide a single and exhaustive taxonomy of the regulation-productivity linkages.[1] The focus in this paper is on regulations that curb market competition (henceforth "anticompetitive regulations").[2] Our analysis is therefore related to the large and growing literature on the effects of competition on growth (see Aghion and Griffith, 2005 for a survey). Recent models of endogenous growth often include the feature that, with technology flows unfettered across countries, productivity growth in follower countries or industries depends on both the ability to catch up by adopting leading technologies available on the market and the

[1] For two recent attempts, see Griffith and Harrison (2004) and Crafts (2006).

[2] In other words, we concentrate on ways in which ill-designed regulations can harm productivity. We do not discuss the potential benefits of appropriate regulations for productivity.

ability to innovate, with the importance of innovation increasing as the country or industry gets closer to the world technology frontier (Aghion and Howitt, 1998; Acemoglu et al., 2006).

According to this line of research, anticompetitive regulations influence the productivity of existing firms by altering the incentives for technology adoption and investment in innovation. They can do so by reducing the rivalry among incumbents and by making the entry of new innovative firms difficult. Conversely, the opening up of markets and increased competitive pressures provide both opportunities and incentives for firms to upgrade their capital stocks, adopt new technologies and innovate to reach and possibly push out frontier production techniques.[3] While the empirical evidence is mixed, recent cross-country and micro-economic studies suggest that these effects are significant, especially where the absorptive capacity is high.[4]

The links between anticompetitive regulations and productivity are likely to be influenced by the level of economic development of each country and by the characteristics of both firms and industries within each country. On the one hand, a strand of research has highlighted that the effects of regulations on productivity differ across countries, firms and industries depending on their proximity to, or their distance from, best practice production techniques. On the other hand, another strand of research emphasizes the importance of anticompetitive regulations for the process of reallocation of resources from less to more efficient firms, which underpins the aggregate growth of market economies.

---

[3] The role of regulatory barriers and monopoly rights in curbing or preventing technology adoption has been illustrated by Parente and Prescott (e.g.,1994, 1999). Other models have focused on the role of new technologically advanced entrants. These may give incumbents the incentives to upgrade their capital through imitation. Aside from pure imitation, affiliates of foreign multinationals may also provide incumbents with positive externalities, such as exposure to foreign high-technology intermediate inputs (Rodríguez-Clare, 1996), learning spillovers from multinationals to their domestic suppliers (Javorcik, 2004) and skill spillovers for the host-country labour force (Fosfuri et al., 2001).

[4] For instance, evidence suggests that an increase in the presence of foreign affiliates is likely to be associated with higher levels of multifactor productivity. This evidence was surveyed by Keller (2004) and Görg and Greenaway (2002). For studies finding positive spillovers, see for instance Haskel et al. (2007), Griffith et al. (2006), Javorcik (2004) and Arnold, Javorcik and Mattoo (2011). Recently, the attention has focused on the precise channels through which these spillovers occur (see, for instance, Crespi et al., 2007).

### 1.1 Regulation, productivity and distance to best practice

At the aggregate level, the potentially different effect of anticompetitive regulations on growth depending on the stage of a country's development is just one element of the debate around "appropriate institutions" for growth (Acemoglu et al. 2006; Aghion and Howitt, 2006). The idea here is that regulations that encourage market openness and entry of new firms (domestic or foreign) can have differential effects on performance depending on whether growth is mainly fuelled by innovation or capital accumulation and technology adoption (e.g. via imitation), with the latter partly determined by the ability of a country to absorb and adapt to foreign technology. If the absorption and innovation capacity is low, as would happen in many developing countries, openness and entry may not have the same positive incentive effects that they usually have in more advanced countries. Thus, the adverse effect of anticompetitive regulations on growth would be expected to be stronger for countries that have a higher level of productivity and GDP per capita.

At the industry level, the effects of anticompetitive regulations can also differ depending on the industry propensity to use certain types of technologies. For instance, anticompetitive regulations may slow down the take up of new general-purpose technologies, such as information and communication technologies. This is because with low competitive pressures, the incentives to invest in such technologies so as to increase productivity and retain market shares may be lower than in more competitive markets. Poschke (2010) shows that the reduction in such incentives due to regulatory barriers to entry can explain a good deal of the productivity differences between the United States and Europe, once technology choice at entry of new firms is accounted for. Moreover, regulatory burdens can make the necessary within and cross-firm adjustments to new production techniques more costly than where such regulations are lighter (for instance by protecting the rents of providers of high technology intermediate inputs). Anticompetitive regulations, including border barriers, can also hinder the diffusion process, not least by preventing the prices of new general-purpose technologies from falling as rapidly as in the global market.

Most importantly, at the firm level the impact of anticompetitive regulations productivity can depend on the characteristics of incumbents, new entrants and exiting firms, particularly their position relative to frontier production techniques (Askenazy et al., 2008). In the aggregate, this can imply a non-linear link between regulation and productivity that depends on the overall degree of firm heterogeneity in regulated markets. In some cases, the relationship between aggregate innovation (and productivity) and competitive pressures can be hump-shaped, with too little or too much competition being harmful for innovative efforts (Aghion et al., 2005). For instance, the incentive effect of competition on incumbents' innovative activities is likely to be stronger for firms whose cost structure is close to that of their innovating rivals than for firms that have a large technological gap to fill (Aghion et al., 2004; Aghion et al., 2006). For firms that are far enough from the world frontier, the "Schumpeterian" discouragement effect due to an increase in entry (which can reflect competition in a market) can be strong enough to deter any innovation activity.

### *1.2 Regulation, productivity and resource reallocation*

Regulation can also affect aggregate productivity growth by making reallocation of resources across heterogeneous firms less efficiency-enhancing. There is a sizeable heterogeneity in firms' characteristics and productivity performance even in narrowly-defined industries, and a larger heterogeneity in relatively newer industries characterised by faster technological progress (see e.g. Caves, 1998; Bartelsman and Doms, 2000; Bartelsman et al., 2004). These heterogeneity patterns are often associated with the idea that firms, whether new entrants or incumbents, are continuously evolving and experimenting with new ideas and technologies (broadly defined to include the use of advanced technologies but also organizational structures) in order to gain market shares or simply survive.[5] Research based on firm-level data suggests that all market economies are characterized by a continuous process of reallocation of resources across such heterogeneous firms and that this process plays a major role for aggregate productivity and output growth (e.g. Olley and Pakes, 1996; Foster et al., 2002; Griliches and Regev, 1995; Bartelsman et al., 2004, 2009a). Resource reallocation

[5] Different theoretical models and growing empirical evidence support the idea that firms – both incumbents and new firms – are engaged in a continuous process of "experimentation" in which they choose whether to enter or stay in the market, and whether or not to expand and adopt new technologies that may have higher potentials but also run greater risks (see e.g. Sutton, 1997, Pakes and Ericson, 1998 and Geroski, 1995, for surveys). Indeed, entering a new market always involves significant uncertainties, especially if this is associated with the adoption of a new, potentially more productive but also more uncertain, technology.

is driven by incumbent firms adapting to market and technological changes, but also by firm dynamics – the entry of new firms, their expansion in the initial years of life and the exit of obsolete units. Firm turnover is a particularly important vehicle for the implementation of new innovations in industries characterised by faster technological progress,[6] where technology adoption often requires (more than in other industries) significant changes in the organization of production and skill composition.[7] Many of the new firms that enter the market fail in the initial years of life, but those that survive tend to grow, often at a higher pace than incumbent firms (see e.g. Geroski, 1995; Sutton, 1997; Bartelsman et al., 2004, 2009a). Interestingly, while the magnitude of firm turnover is fairly similar across countries, the characteristics of entrants and exiters, their growth performance and overall contributions to technological adoption and, ultimately, to productivity growth vary considerably (Foster et al., 2002; Bartelsman et al., 2004, 2009a; Griffith et al., 2006).

A growing body of empirical research has been relating differences in the contribution of resource reallocation to productivity growth to differences in policies and institutions that shape the business environment. The list of policy and institutional factors that are likely to promote experimentation and efficient resource allocation across sectors and firms is long. A substantial literature has examined the impact of credit constraints on firm dynamics and technology adoption (e.g. Rajan and Zingales, 1998; Beck et al., 2004; Klapper et al., 2006; Aghion et al., 2007). A more limited number of studies have looked at the role of labour market regulations in influencing labour reallocation and the adaptability of firms to technological shocks (Haltiwanger et al., 2006; Micco and Pagés, 2006). More recently, the focus has increasingly been on regulations in the product market, especially those that affect the intensity of competitive pressures.

Anticompetitive regulations are likely to influence the incentives for new firms to enter a given market, as well as for incumbents to engage in

---

[6] Bartelsman et al. (2004) as well as Bartelsman et al. (2005) indeed found that the entry of new firms plays a stronger role in boosting aggregate productivity in high-tech industries as compared to medium and low-tech industries.

[7] Newcomers may have a comparative advantage in implementing them relative to existing firms in as much as they do not have to incur any adjustment costs. The wider range of technology options available to entrant firms but also the greater uncertainty concerning the business plan explain the observed greater variance in the performance of young businesses compared to older incumbents.

experimentation and the associated reallocation of resources. Such regulations can hinder the reallocation of resources across firms with different productivities. A number of theoretical studies have tried to account for firm heterogeneity and modelled distortions to entry and exit as well as reallocation. For example, Bernard et al. (2003) and Melitz (2003) highlight the role of border barriers affecting the degree of competition in the product market. Building on models by Melitz and Ottaviano (2008) and Del Gatto et al. (2006), Corcos et al. (2007) find that lifting behind-the-border barriers may be even more important for productivity. In their models with heterogeneous firms, easing trade barriers generates a reallocation of resources in favour of more productive firms. The exit of low-productivity firms and the expansion in the domestic and foreign markets of more productive firms lead to an increase in aggregate productivity growth. Bergoeing et al. (2002) also allow for idiosyncratic differences in firm productivity and focus on the effect of a productivity shock on aggregate productivity when there are government-induced frictions in the reallocation of resources. Their simulations suggest that such frictions lengthen the period in which output is below potential. A few additional studies have further developed models with adjustment frictions that prevent resources from immediately being allocated to the most productive firms (see e.g. Restuccia and Rogerson, 2007; Hsieh and Klenow, 2009; Bartelsman et al., 2009b). Static and dynamic frictions partly depend on market characteristics and technological factors but are also clearly related to inappropriate product market regulations. In particular, frictions may represent the costs of adjustment – either in the form of entry and exit costs, or adjustment costs to reallocate factors of production such as capital and labour.[8] In these models as well, both policy-induced entry costs and regulations that raise the adjustment costs to technological shocks reduce aggregate productivity.

As stressed by Bartelsman et al. (2009b), inappropriate regulations may affect the reallocation dynamics on different margins in a variety of ways. For example, high start-up costs are likely to reduce firm turnover and potentially lead to a less efficient allocation of resources, but those firms that finally enter the market may have a higher productivity than otherwise

[8] The latter might involve a range of costs including the search and matching frictions that have been the focus of much of the recent literature on the dynamics of the labour market (see e.g. Davis et al., 1996; Restuccia and Rogerson, 2007; Hsieh and Klenow, 2009).

due to a tighter selection at entry. In turn, the average productivity of incumbents and exiting businesses will be lower. Similarly, certain market distortions might weaken the selection process at entry and exit leading to less systematic differences between entering, exiting and incumbent businesses. There is also an important time dimension: market conditions that promote experimentation and trial and error processes may be associated with more risk and uncertainty in the short run, leading to a lower immediate contribution from entry to productivity, but a higher long-run contribution once the trial and error process of experimenting firms has worked its way out (through learning and selection effects).

### *1.3 Intersectoral linkages*

Via the channels highlighted above, regulations that hinder competition can affect productivity not only in each regulated industry but also on other industries through intersectoral linkages. Lack of competitive pressures in a sector can generate trickle-down effects on other sectors by raising the costs, lowering the quality or reducing the availability of intermediate inputs, particularly in the case of services inputs where import competition is limited. Recent research has explored the indirect effects that barriers to competition in (upstream) sectors may have on the efficiency of resource allocation and the productivity performance in other (downstream) sectors (Bourlès et al., 2010; Barone and Cingano, 2011, Arnold, Javorcik and Mattoo, 2011).

The main idea is that upstream regulation generates market power for intermediate goods providers, which is used to extract rents and restrict access to key markets for downstream firms, reducing the opportunities and incentives for productivity improvements. Based on a variant of the innovation model by Aghion et al. (1997), Bourlès et al. (2010) show that anticompetitive upstream regulations can reduce competitive pressures in downstream markets by increasing the cost of finding an intermediate supplier. For instance, lack of competition in an upstream sector can generate barriers to entry that also curb competition in downstream sectors: tight licensing requirements in retail trade or transport can narrow the distribution channels for downstream firms and overly restrictive regulation in banking can reduce the range of available sources of financing for downstream firms, thereby curbing new entry and firm growth. Moreover, the incentives to improve efficiency downstream are also reduced by the ability of upstream firms to appropriate a share of the rents that downstream firms would earn from such improvements. This is because, if the markets for intermediate inputs are imperfect, downstream firms may have to negotiate with (and can be held up by) suppliers. In a similar vein, but based on a model of industry interdependence and international specialization, Barone and Cingano (2011) show that regulations restricting competition in upstream sectors for which import competition is weak (e.g. services) affect the cost and/or quality of products used as intermediate inputs in downstream industries or firms. This imposes unnecessary costs of adjustment to downstream firms wishing to improve efficiency and biases industry specialization away from industries that are intensive in the regulated inputs.[9] Resource allocation across industries and aggregate productivity growth are obviously also affected.

## 2. Tracking differences in regulation across countries, industries and over time

### *2.1 Measuring regulation*

Studying the quantitative effects of regulation on productivity requires measuring regulation in a relevant, consistent and comparable way across countries, industries and time. In the context of this study, relevance means only considering regulations that have an impact on competitive outcomes in markets, industries and countries. Consistency and comparability can be reached in a variety of ways. For instance, Griffith et al. (2004) and Aghion et al. (2006) have recently used EU data on anti-monopoly cases and the implementation of the Single Market Programme to address the potential policy determinants of competition, while Buccirossi et al. (2009) have used variability in competition law provisions and enforcement rules in a subset of OECD countries. In this paper, we focus on indicators of anti-

[9] Indeed, in theoretical models of industry interdependence, the under-development of markets for non-tradable inputs has been shown to constrain (or even prevent) the diffusion of input intensive technologies, thus affecting the patterns of resource allocation and international specialization (Rodriguez-Clare, 1996; Okuno-Fujiwara, 1988). Barone and Cingano's work is related to the growing literature on the relevance of institutions for resource allocation and comparative advantages (see, references therein).

competitive product market regulations drawn from the OECD international product market regulation database.[10] These indicators measure to what extent competition and firm choices are restricted where there are no *a priori* reasons for government interference, or where regulatory goals could plausibly be achieved by less coercive means. They are based on detailed information on laws, rules and industry settings (e.g. the extent of vertical integration or monopoly power), and cover both general purpose regulations (such as administrative burdens on startups) and sectoral regulations in energy (gas and electricity), transport (rail, road and air), communication (post, fixed and cellular telecommunications), retail distribution, professional services and banking, with country and time coverage varying across industries. This information covers both domestic anticompetitive regulations and industry-specific FDI restrictions in all OECD countries as well as in the BRIICS.[11]

The main advantages of using these indicators in empirical analysis is that they can be held to be exogenous to productivity developments and that they are directly related to underlying policies, a feature that business survey data do not have.[12] Another advantage is that, since they are composite constructs based on detailed information on specific policies, they address multicollinearity problems in estimation. At the same time, they make it possible to focus on the specific aspects of policies that are considered to be relevant for productivity. For instance, most of the analysis reported below deals with barriers to entry (including administrative burdens), sometimes explicitly distinguishing between border and non-border policies that affect these barriers. Yet another advantage of the OECD

[10] The data and underlying documentation are publicly available at www.oecd.org/eco/pmr. The most recent observations are currently for 2007/2008.

[11] The basic regulatory data includes: economy-wide indicators for all OECD countries and several non-OECD ones for 1998, 2003 and 2007; indicators for energy, transport and communication that cover most OECD countries over the 1975-2007 period (several non-OECD countries are also covered for the most recent period); indicators for retail distribution and professional services that cover most OECD countries and several non-OECD countries for 1998, 2003 and 2007; the indicator for banking that covers 30 OECD countries for 2003. The indicator of FDI restrictions covers a larger set of sectors over the 1981-2007 period.

[12] Naturally, endogeneity cannot be completely ruled out if, for instance, policies are affected by productivity outcomes through political economy channels. On the relative advantages of policy-based and survey-based composite indicators, see Nicoletti and Pryor (2006).

indicators is that they vary over countries, industries and time, though full time variability is limited to a subset of non-manufacturing industries.[13]

The OECD indicators are also used to summarize the potential burden of non-manufacturing regulations imposed on all business sectors via inter-sectoral linkages. This is particularly important because the non-manufacturing sector is undoubtedly the most regulated and sheltered part of the economy, while few explicit barriers to competition remain in markets for manufactured goods of OECD economies. However, as discussed above, even low-regulated industries may suffer from regulation-induced inefficiencies in non-manufacturing because all industries are heavy intermediate consumers of non-manufacturing inputs. Sectoral "regulation impact" (RI) indicators of the indirect burden of anti-competitive regulation in upstream non-manufacturing industries for downstream industries (including the regulated non-manufacturing industries themselves) are calculated for each country using information from input-output tables:[14]

$$RI_{kt} = \sum_{j} \left( NMR_{jt} + FDI_{jt} \right) \bullet w_{jk} \qquad 0 < w_{jk} < 1$$

where the variable $NMR_{jt}$ is an indicator of domestic anti-competitive regulation in non-manufacturing sector $j$ at time $t$, $FDI_{jt}$ is an indicator of FDI restrictions in non-manufacturing sector $j$ at time $t$, and weight $w_{jk}$ is the total input requirement of sector $k$ for intermediate inputs from non-

[13] Griffith et al. (2006) formulate a number of criticisms concerning the OECD indicators, the most compelling being that their time dimension is limited to a subset of non-manufacturing sectors that they do not think are sufficiently representative of economy-wide regulatory developments. Conway and Nicoletti (2006) show that the OECD indicator of non-manufacturing regulation is closely correlated, both across countries and over time, with a popular time-series indicator of economy-wide business regulation, the Economic Freedom of the World index by Gwartney and Lawson (2003). This is not surprising since most OECD product market reforms have been implemented in the non-manufacturing industries over the past decades.

[14] The resulting RI covers 39 sectors that use the outputs of these non-manufacturing industries as intermediate inputs for the 1975-2007 period. Given that some sectoral indicators (retail, professional services and banking) have a limited time coverage, we use their 2003 value to compute the regulation impact indicators. But the empirical results reported in the next section do not change if values for 1998, 2003 and 2007 are instead used, with interpolation between periods. This technique for calculating the regulation impact indicators has also been used by Faini et al. (2006) and Barone and Cingano (2011).

manufacturing sector *j*.[15] These RI indicators allow tracking the "trickle down" effects of inappropriate regulations in non-manufacturing industries on productivity in all sectors of the economy.[16]

### *2.2 Regulation: Cross-country patterns and historical developments*

Figure 3 shows cross-country patterns and the evolution of economy-wide product market regulation and FDI restrictions across non-manufacturing sectors. It suggests that, overall, regulatory approaches have converged across OECD countries over the past two decades towards a more pro-competitive stance. Looking at specific non-manufacturing sectors, convergence has taken place in particular in energy, transport and communication as well as in border barriers to FDI (for the latter see Panel B of Figure 3), while the available time-series data for retail trade and business services point to persistent differences in the regulatory stance across countries in these sectors. Despite convergence in many areas and sectors, differences in regulation persisted at the end of the period, suggesting that competitive pressures still differ considerably across both countries and sectors.

The figure also suggests that, in the most recent period for which data are available, regulations often tended to remain more adverse to competition in emerging economies than in OECD countries, though not necessarily in all sectors. Unfortunately, historical data are lacking and it is not possible to use the OECD indicators for tracking whether emerging economies have been converging in regulatory practices towards more advanced economies.

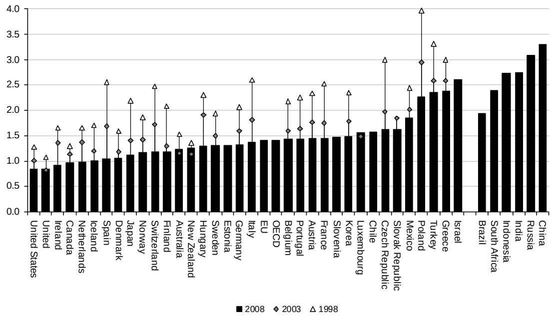
---

[15] The (harmonised) input-output data for OECD countries, and therefore the $w_{jk}$, exist at the 2-digit (ISIC rev3) level, implying that the NMR must also be calculated at this level of sectoral aggregation. See Arnold et al. (2008) for further details.

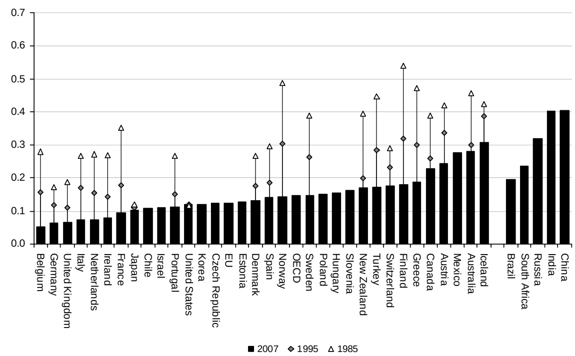
[16] All OECD indicators take continuous values on a scale going from least to most restrictive of competition. A detailed description of the indicators of economy-wide regulation can be found in Woelfl et al. (2009, 2010) while a detailed description of domestic non-manufacturing regulation and the trickle down indicators of "regulation impact" is provided in Conway and Nicoletti (2006). Indicators of FDI restrictions are described in Golub (2003) and Golub and Koyama (2006). The indicator of domestic anticompetitive regulations in banking is described in de Serres et al. (2006).

**Figure 3. Economy-wide product market regulation and border barriers**

**a. Overall regulation**




**b. FDI restrictions in business sectors**




Source: OECD International Regulation Database.

Note: OECD indicator, scale 0-6 from least to most restrictive

These patterns raise a number of issues that are relevant from both a research and policy point of view. First, to what extent is the more restrictive stance in a number of countries, including the emerging economies, slowing down their GDP per capita and productivity growth rates? Second, can differences in competitive pressures across industries that are induced by different regulatory approaches explain the wide cross-country and cross-industry dispersion of productivity growth rates observed in the OECD area? Third, to what extent can regulations that curb such pressures and generate unnecessary burdens for businesses hinder reallocation towards the most efficient firms? Fourth, do these regulations affect all countries, industries and firms equally, irrespective of their technological characteristics, dynamism and distance to best practice? We now turn to the cross-country empirical evidence on these issues.

## 3. Evidence on regulation and productivity

A growing number of recent empirical studies have focused on the effects of product market conditions on growth in productivity and GDP per capita. Some studies have focused on the impact of product market conditions on capital accumulation (Alesina et al., 2005) and its asset composition (Gust and Marquez, 2004; Conway et al., 2006; Bloom et al., 2010) as well as on their effects on innovation (Aghion et al., 2005; Aghion and Griffith, 2005). Here we focus on those cross-country studies that have directly related measures of anticompetitive regulation to GDP per capita and productivity growth. The review does not have the ambition to be exhaustive.

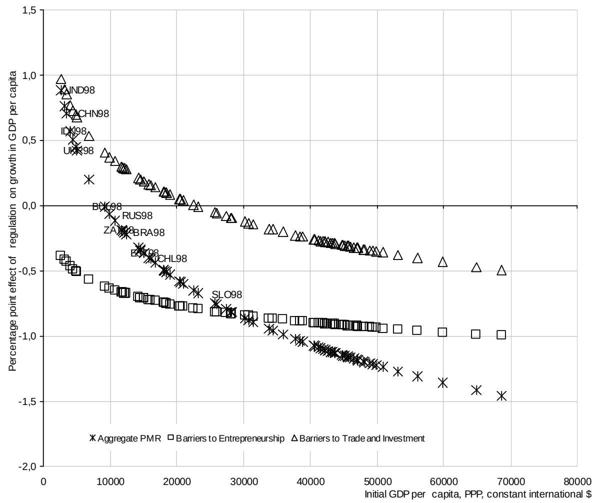
### *3.1 Some aggregate evidence*

Empirical research linking anticompetitive regulations to aggregate growth has found negative effects on GDP per capita, GDP per worker or multi-factor productivity (MFP) growth, but the results are not always robust and consistent across studies. These studies have taken empirical approaches based either on static cross-country growth regressions à la Barro and Sala-i-Martin or dynamic panel regressions. Static models have been estimated with either a fixed number of explanatory variables (in addition to regulation) or with methods that allow to identify the variables that are most likely to affect growth among a vast number of possible factors, including regulation (so-called Bayesian Model Averaging – BMA). Studies also differ in terms of the sample of countries used. As shown in Babetskii and Campos (2007), differences in methodology and sample coverage can significantly affect the size (and sometimes the sign) of the growth effects of changes in institutional variables.

A few recent studies illustrate well the fragility of aggregate findings. Using a BMA methodology and focusing on GDP per capita, Woelfl et al. (2010) find that easing anticompetitive regulations by an amount equivalent to moving from the regulatory stance of Brazil to that of the average OECD country could yield a 0.3 per cent higher average annual rate of per capita GDP. Boulhol et al. (2008) previously found similar results based on simple dynamic panel regressions. However, the statistical significance of results from both these studies is relatively weak. Using a more complex dynamic approach based on Bloom et al. (2002), Bouis et al. (2011) also find that anticompetitive regulations curb GDP per capita via their effect on MFP, but they are unable to sharply distinguish the influence of regulation from that of other institutional variables within their estimation framework. Finally, specifically focusing on MFP in the context of static cross-country growth regressions, Aghion et al. (2009) also find adverse effects of market rigidities (expressed as a combination of labour and product market regulation) on aggregate performance. A common feature of all these studies is that, among the various kinds of regulations that were tested, barriers to entry and entrepreneurship are found to be those having the most significant and damaging effects on performance.

One possible reason for the lack of robustness of results from aggregate studies is that the effects of regulation, and of different kinds of regulations, may vary with levels of development (Aghion and Howitt, 2006). This implies a "composition" effect that blurs the link between regulation and performance when this non-linearity is not accounted for in estimation. Some estimates from dynamic panels with thresholds (or simple dummies) differentiating among effects of regulations across income levels suggest that anticompetitive regulations may have particularly adverse effects on more advanced countries, while having lesser negative effects and even positive ones at low levels of development. For instance, Figure 4 shows how the effects of different kinds of regulations vary across countries with different initial GDP per capita levels according to the panel estimates of Woelfl et al. (2010). Negative effects of overall anticompetitive regulations (PMR) begin to be observed at GDP per capita levels just above those of Bulgaria or South Africa in 1998, with certain barriers to trade and investment still having positive effects even at higher income levels. As already mentioned, barriers to entrepreneurship have nonetheless uniformly negative effects on growth in all countries independent of GDP per capita levels. But the effects of all types of regulations on growth are estimated to become increasingly adverse as income levels rise, and particularly steeply so for those regulations that affect international openness. Aghion et al. (2009) found similar threshold effects of market rigidities on aggregate MFP growth, with rigidities decreasing growth only in countries with income close to the level of the United States. No such threshold effects were found, however, in the dynamic panel estimated by Bouis et al. (2011) suggesting that the jury is still out concerning the relevance of such effects for policy analysis and recommendations.

**Figure 4. The impact of regulation on growth at different levels of initial GDP per capita**




Source: Woelfl et al. (2010).

## *3.2 Regulation and industry-level productivity*

To begin exploring the link between regulation and industry-level productivity, Figure 5a shows the cross-industry distribution of labour productivity growth rates over the 1995-2005 period in two groups of countries for which we have consistent data: three relatively "deregulated" English-speaking countries – the United States, the United Kingdom and Ireland – and four relatively "restrictive" large European countries Germany, France, Italy and Spain. The figures focus on trend productivity growth rates to abstract from short-term fluctuations.[17] Moreover, the growth rates have been purged of idiosyncratic effects across countries and industries to make it possible to pool the productivity data in a meaningful way.[18] Therefore, values on the horizontal axis are not directly interpretable, while their dispersion (overall and across industries) is.

Several features emerge from Figure 5a. For both groups of countries, the overall distribution is skewed to the left, indicating prevalence of weak productivity growth rates, but has a long right tail, suggesting cases of high productivity growth. Interestingly, the right tail of fast growing industries is longer and thicker in English-speaking countries than in continental EU countries that have a higher concentration among relatively more slowly growing industries. As a consequence, English-speaking countries tend to have a higher median productivity growth than continental EU countries (as shown by the distance between the vertical lines).

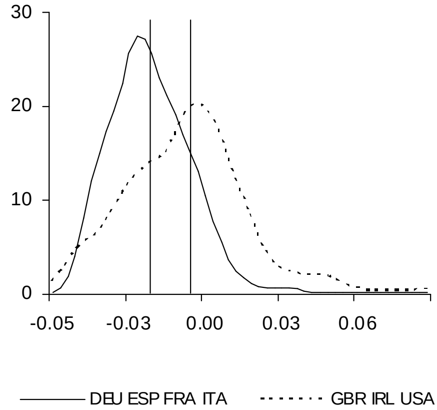
In the light of our previous discussion, it is natural to relate these differences in the distribution of productivity growth to underlying product market regulations that are more or less prone to help sustain efficiency improvements within each industry. As a first check on this conjecture, Figure 5b replicates the productivity growth distributions pooling together all countries, but now distinguishing between high- and low-regulated cases (each observation being for a country/sector/year, again purged of idiosyncratic factors). Low and high-regulated cases are defined as those falling within the first and fifth quintiles, respectively, of the distribution of the OECD regulation impact indicator. As explained above, using these indicators makes it possible to account for both the direct effects of anti-competitive regulations in each industry and the indirect effects via inter-sectoral linkages.

[17] To eliminate the cyclical component, a Hodrick-Prescott filter has been applied to the series using all available years.

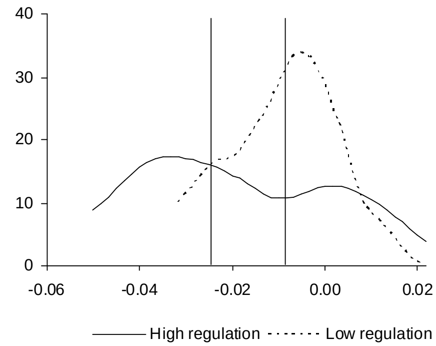
[18] In other words, the figure shows the distribution of the residual of a regression of productivity growth rates on country and sector dummies after applying a Hodrick-Prescott filter and eliminating outliers (top and bottom percentile of the distribution). The resulting distributions are based on country-industry-year observations.

**Figure 5. Labour productivity growth distributions across countries, industries and time, 1995-2005[a]**

**a. High and low regulation countries**




**b. High and low regulation industries[b]**




Source: Authors' calculations based on EU-KLEMS, March 2007, and OECD International Regulation Database.

Notes: [a] Trend productivity growth, Hodrick-Prescott (HP) filtered and purged of country and industry means. Agriculture, forestry, fishing, mining and construction are excluded, as are public administration, education and health sectors. [b] Observations are classified into low or high regulation if they fall in the first and last quintiles, respectively, of the distribution of the regulation impact indicator. The vertical bars represent sub-sample medians. These indicators reflect the trickle down effects of anti-competitive regulation in non-manufacturing sectors on industries that use the output of these sectors as intermediate inputs into the production process.

The figure suggests that regulation plays a role in shaping the distribution of productivity growth rates. Where regulation encourages competition and does not impose any excessive costs to businesses, the density of high productivity growth rates is thicker and median productivity growth is higher than where regulations are restrictive and costly. Moreover, the dispersion of productivity growth rates is much higher in highly-regulated countries, with observations displaying relatively low productivity growth being much more frequent.

The wide industry-level dispersion of productivity growth rates is a potentially important source of identification for econometric studies of the regulation-productivity link. A large number of such industry-level studies have been implemented over the past decade, mostly relying on dynamic panel data analyses (e.g., Scarpetta and Tressel, 2002; Nicoletti and Scarpetta, 2003; Conway et al., 2006; Griffith et al., 2006; Inklaar et al., 2008; Buccirossi et al., 2009; Bourlès et al., 2010) within the general framework proposed by Aghion and Howitt (2006).[19] In this framework, sectoral productivity growth in a given country depends on the ability to keep pace with growth in the country with the highest level of productivity (the leader) by either innovating or taking advantage of the best technology available. Productivity growth depends on both knowledge spillovers from the leader's innovation drive and the speed at which the productivity gap is closing due to, for instance, technology diffusion and adoption. In turn, the effect of anticompetitive regulation on productivity growth in follower countries is assumed to depend on the size of the sectoral productivity gap.[20]

While the basic estimation framework is similar, the various studies differ in data and coverage, control variables and, especially, in the measurement of product market policies, with the most recent studies focusing on the indirect burdens imposed by (upstream) non-manufacturing regulations on all (downstream) business sectors (see Arnold et al., 2008 for a survey). The match between the industry productivity dimension and the industry-level regulation impact indicators constructed by the OECD, as well as their time-series variability, has proved to be particularly useful for the estimation.

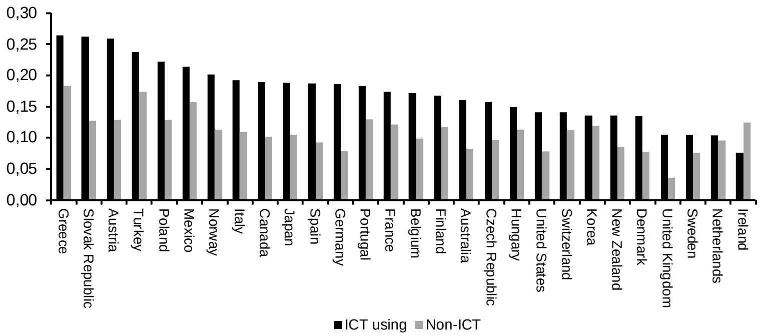
[19] The regression models used by these authors are variants of that proposed by Griffith et al. (2004) to test the effect of R&D expenditure on productivity growth. Bouis et al. (2011) have used a similar model to study the effects of regulation on aggregate MFP.

[20] All these studies include country- and sector-fixed effects. However, due to the presence of the interaction term between the productivity gap and regulation, the source of identification of the regulation effects is variability across all dimensions of the panel: time, industries and countries.

Given the differences in data and specification, the results from industry-level studies are not easily comparable. However, a number of common conclusions emerge:

- In all studies, regulations that restrict competition are found to curb labour productivity or MFP growth significantly, even though the point estimates vary.
- Studies that obtain separate estimates for different sectors (Conway et al., 2006; Inklaar et al., 2008) tend to find stronger negative effects in ICT-intensive industries.
- Regulations that appear to be most damaging for sectoral productivity growth are barriers to entry, consistent with the results found in aggregate growth regressions (see above).
- Studies that account for regulatory burdens implied by intersectoral linkages (Conway et al., 2006; Inklaar et al., 2008; Bourlès et al., 2010) find these burdens to provide an important explanation of the dispersion in productivity growth rates across countries, industries and over time.

**Figure 6. The burden of non-manufacturing regulation on ICT and non ICT sectors, 2003[a]**




Source: OECD International Regulation Database.

Notes: Scale normalised to 0-1 from least to most burdensome. [a] The figure shows the regulation impact indicator, which reflects the burden of anti-competitive regulation in non-manufacturing sectors on industries that use the output of these sectors as intermediate inputs into the production process.

Focusing on labour productivity, Conway et al. (2006) show that regulatory burdens have been particularly harmful to productivity improvements in ICT-intensive sectors, largely because they slowed down the catch-up process to best practice productivity. Conway and Nicoletti (2007) estimate the productivity growth "deficit" that would be suffered by countries whose anticompetitive regulations would hinder such a catch-up following a global positive productivity shock such as that experienced in the OECD area during the diffusion of ICT. In all countries, the detrimental effect of anti-competitive regulation, again expressed by the regulation impact indicator, is larger in ICT-intensive sectors given that the regulatory burden is estimated to be higher in these sectors in comparison to non-ICT intensive sectors (Figure 6). The estimated gap in productivity catch-up in ICT-intensive sectors is particularly sizeable in Austria, Greece, Italy, Germany, Norway and Belgium, all of which remain 30 per cent to 40 per cent below potential five years after the initial shock.

While virtually all industry-level studies of the regulation-productivity link on average find adverse effects of anti-competitive policies on growth, there is less agreement on whether these effects are uniform across countries (or sectors) independent of their distance to the technological frontier. Among the studies that have conditioned the effects of regulation on distance to frontier, Conway et al. (2006) and Nicoletti and Scarpetta (2003) find that regulation tends to slow down productivity growth more strongly in countries (or sectors) that are further away from global best practice productivity. They ascribe this result to the tendency of weak competitive pressures and burdens implied by regulation to lower incentives and opportunities and increase the costs of adopting best practice production technologies and methods. Average developments in industry productivity would thereby suffer from weak growth in the most efficient firms and a low contribution of firm turnover to efficiency improvements.[21]

Recent studies (Bourlès et al., 2010) suggest, however, that anticompetitive regulations in up-stream industries tend to have a more damaging effect on the multifactor productivity growth of sectors sufficiently close to the global productivity frontier. This is consistent with the neo-Schumpeterian view that lack of competition is particularly harmful where the "escape competition" effect benefiting an innovating firm is the strongest – that is in a situation of neck-and-neck rivalry among firms (Aghion et

[21] Since Conway et al. (2006) focus on labour productivity, the greater harm to productivity growth caused by anticompetitive regulations for countries and sectors that are further away from the global frontier can also be ascribed to the tendency of such regulations to curb capital formation (Alesina et al., 2005) and ICT investment (Gust and Marquez, 2004).

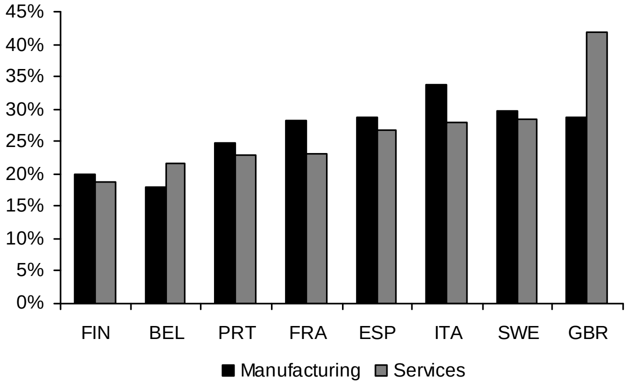
al., 1997; Aghion and Howitt, 2006). Interestingly, Bourlès et al. also find that this "closeness to frontier" effect vanishes in the most recent period (1995-2007) characterised by increased integration of global markets and the widest diffusion of ICT technologies. In other words, countries (and sectors) in their sample uniformly suffered from anticompetitive regulations in the most resent post independent of whether they are close to or well behind the frontier. Over the whole estimation period, regulation is found to curb productivity for more than 85 per cent of the observations, while significantly increasing it for a small share of them (3 per cent) namely for firms whose MFP levels are less than half of those of the global technology leader. However, over 1995-2007 regulation is estimated to have curbed MFP growth for virtually all observations in the sample. Using the average level of the productivity gap and the average level of regulation, regulation is estimated to curb annual MFP growth by around 1 percentage point over the whole period and by around 1.7 percentage points more recently.

### *3.3 Regulation and firm-level reallocation and productivity*

As discussed above, industry-level productivity growth hides a widespread heterogeneity in firms' performance within each industry and a continuous process of reallocation across them, through the entry of new firms and the exit of obsolete ones and the reallocation of factor inputs among continuers. All industries display persistent productivity dispersion, pointing to a (more or less) wide heterogeneity in the performance of firms. In this context, a natural question is whether market forces tend to reallocate resources towards firms with higher efficiency levels. A simple way of assessing the importance of reallocation for productivity is to ask the question – are resources efficiently allocated in a sector/country in the cross section of firms at a given point in time? To answer this question, we focus on multi-factor productivity, which is the appropriate measure of firm-level efficiency in the use of inputs, and we use the simple cross-sectional decomposition of multi-factor productivity levels for a sector at a point in time developed by Olley and Pakes (1996). Aggregate productivity (P) is decomposed into two terms involving the un-weighted average of firm-level multifactor productivity (AP) plus a cross term that captures *alloca-*

*tive efficiency* since it reflects the extent to which firms with greater efficiency have a greater market share (WP).[22]

**Figure 7. Contribution of resource allocation to sectoral MFP levels (early 2000s)**




Source: Authors' calculations based on Amadeus database.

Notes: Based on Olley-Pakes productivity decomposition. The data reported in the figure represent the share of the total MFP level that is due to an efficient allocation of resources. The degree of efficiency in resource allocation is measured by the cross-term of the Olley and Pakes decomposition (WP, see main text), relative to the overall average productivity level (P).

This decomposition essentially involves comparing un-weighted average productivity to weighted average productivity. To minimise the measurement problems involved in comparing these productivity levels across sectors or countries, we focus on the relative contribution of allocative efficiency to the observed overall average productivity level. This requires comparing productivity levels of firms in the same industry and countries thus ensuring that most measurement problems are controlled for.[23] Figure 7 presents the estimated indicator of efficiency (OP=WP/P) in the allocation of resources in a sample of EU countries for which we have

[22] Formally, the decomposition is given by: $P_t = (1/N_t)\sum_i P_{it} + \sum_i \Delta\theta_{it}\Delta P_{it} = AP_t + WP_t$ where $N$ is the number of businesses in the sector and $\Delta$ is the operator that represents the cross-sectional deviation of the firm-level measure of productivity ($P$) or business market share ($\theta$) from the industry simple average.

[23] Since no cross-country or cross-sector comparability issues arise in the Olley-Pakes decomposition, we take the standard approach of estimating, for each sector and country, a production function in logarithmic form and take the residual, i.e. the part of output that is not explained by factor inputs, as a measure of MFP.

consistent firm-level data from the Amadeus database over the early 2000s. It focuses on manufacturing and business services separately and for each of the two broad sectors, a weighted average of 2-digit industry level OP cross terms is used.

The OP decomposition suggests that, in all countries, allocative efficiency accounts for a significant fraction of the overall observed MFP levels: between 20-40 per cent of the observed productivity levels can be ascribed to the actual allocation of resources compared to a situation in which resources would be randomly allocated across firms in each sector. However, there are also differences across the two broad sectors and across countries. The United Kingdom stands with the highest degree of allocative efficiency in services, almost 15 percentage points above that of the second highest country in the service sector.[24]

To shed some further light on allocative efficiency, Figure 8 plots average firm growth by the quartile of the firm-level productivity distribution. The quartiles divide firms according to their MFP (relative to the median of the sector and country for which the production function was estimated, on average over 1998-2004). Thus, the top quartile represents the 25 per cent most productive firms in each industry. Firm growth is measured in terms of real value added, also averaged over 1998-2004, and normalised by the country/sector average (which is set equal to 1 in the figure).[25] Naturally, this is a partial analysis that does not consider dynamic processes – i.e. some of the low-productivity firms may be new ventures that are involved in a learning-by-doing process and catching up with the efficiency of more mature businesses, while some of the high productive businesses may have less scope for further expansion. Bearing this caveat in mind, the figure suggests that in all countries but one (Spain), more productive firms indeed experience a higher growth than their lower productivity counterparts. However, the difference in the growth of low- and high-productivity firms varies significantly across countries.[26] This confirms our finding

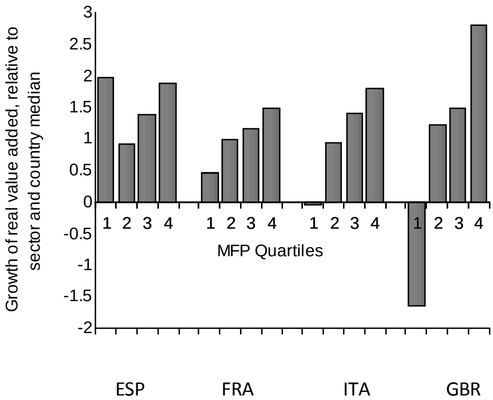
[24] A similar analysis is presented in Bartelsman et al. (2004) with reference to labour productivity in manufacturing in the 1990s in a sample of OECD countries that also includes the United States.

[25] In other words, a value of 3 for the highest quartile in the United Kingdom means that these firms grew on average three times faster than their peers in the same sector/country cell. To minimize endogeneity problems, the growth in firm value added is calculated as the average of $t+2$ and $t+3$ relative to the period of productivity measurement.

[26] While in Spain there is no clear relationship between productivity levels and expansion, in France the most productive 25 per cent have an average growth that is twice as high as the 25 per

based on the cross-sectional OP productivity decomposition, namely that some countries are better able to channel resources towards high-productivity firms, thereby encouraging them to grow rapidly and strongly contributing to the overall productivity performance.

**Figure 8. Do better firms grow faster?[a]**




Source: Authors' calculations based on the Amadeus database.

Notes: Value-added growth by quartiles of the MFP distribution of firms. [a] The figure presents the average real value added growth of the four quartiles of the MFP (relative to the median of the sector and country for which it was estimated) distribution of firms in each country. Firm level real value added growth is normalised by country/sector average to improve the comparability.

Two questions emerge at this point: Why have some countries been more able than others to reallocate resources towards fast growing firms? What are the mechanisms through which inappropriate regulations might affect reallocation across sectors and firms? A first step toward answering these questions is to correlate our OP indicator of allocative efficiency across countries, sectors and time with the OECD indicators of the regulatory burden imposed by non-manufacturing regulation on all sectors of the economy (Table 1). In other words, we investigate whether there is an association between anticompetitive regulations (affecting both upstream and downstream sectors) and the efficiency of the reallocation process within each industry.[27]

cent least productive firms, while in Italy it is three times as high and in the United Kingdom it is five times as high.

[27] We use a fixed-effect specification where, in addition to our regulation impact indicator, we include a full set of time-varying country-specific and sector-specific effects. The sample includes

The results for the overall business sector point to a negative effect of regulatory burdens on the efficiency of resource allocation. However, breaking down the sample into manufacturing and services suggests that the negative effect of regulation originates from services. This is not surprising, since cross-country differences in the regulatory environment and regulatory reforms over the past decade mostly concerned the services sector. Interestingly, if we split the industry sample between ICT-intensive and non-ICT intensive sectors, we find that regulatory burdens affect more strongly the ICT-intensive sectors, where such burdens are often highest (see Figure 6 above). In other words, in those sectors where there was more heterogeneity in firm performance because of greater experimentation and learning by doing around this new general purpose technology, regulations that restricted competition and entry of new firms have had a strong negative effect on the ability of the market to channel resources towards firms with the best performance. This illustrates one channel through which restrictive regulations that impinge on ICT-intensive sectors may have curbed the ability of some countries to fully benefit from the diffusion of new technologies over the past decade, as suggested by Conway et al. (2006) based on industry-level data.

**Table 1. Product market regulation and allocative efficiency**

| **Dependent variable: Olley-Pakes indicator** | **Business sector** | **Manufacturing only** | **Services only** | **ICT using sectors** | **Non-ICT using sectors** |
|---|---|---|---|---|---|
| Regulation Impact Indicator | -0.33***<br>(0.10) | 0.54<br>(1.44) | -0.37**<br>(0.16) | -0.30**<br>(0.14) | -0.26<br>(0.17) |
| Country-year FE | Yes | Yes | Yes | Yes | Yes |
| Industry FE | Yes | Yes | Yes | Yes | Yes |
| N | 894 | 703 | 191 | 417 | 477 |
| R2 | 0.20 | 0.21 | 0.19 | 0.21 | 0.20 |

Source: Author's calculations.

Notes: Standard Errors in parentheses. *, **, *** indicate statistical significance at the 10, 5 and 1 per cent levels, respectively. Agriculture, forestry, fishing, mining and construction are excluded, as are public administration, education and health sectors. ICT-intensive sectors include both ICT-producing and ICT-using sectors.

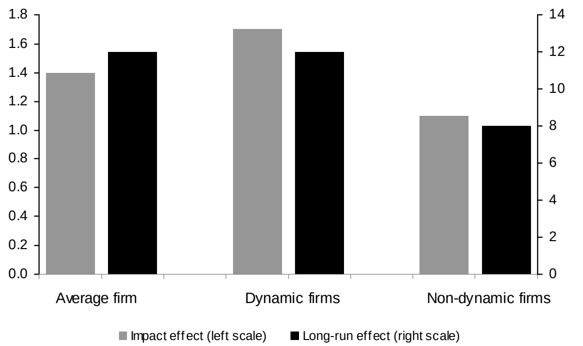
a set of OECD countries for which the Amadeus database has a good coverage of firms: Austria, Belgium, Finland, France, Germany, Italy, Portugal, Spain, Sweden, United Kingdom; the period is 1998-2004.

Further light on the link between regulation, reallocation and productivity growth can be provided by formal econometric analysis using firm-level data. This makes it possible to explore the effects of inappropriate regulations on firm-level productivity while accounting for heterogeneity in firm characteristics. Limits in the availability of comparable firm-level data have so far restricted the number of cross-country empirical studies of this kind. Most available studies have therefore focused on firm-level panels in individual countries. Moreover, most firm-level studies of the competition-performance nexus have used measures of competition based on market outcomes, such as entry rates, markups, market shares or concentration indices (Nickell, 1996; Blundell et al., 1999; Aghion et al., 2004, 2005, 2006; Forlani, 2011). Here we report results from three recent multi-country firm-level studies that have explicitly focused on the role of barriers to entry imposed by regulation.

Klapper et al. (2006) looked at the effect of entry regulations, as measured by the World Bank Doing Business indicators (World Bank, 2004), on entry rates, the size of entrants and their labour productivity growth rates in a two-year (1998-1999) panel of European firms covered by the Amadeus database. They note that depending on their design, entry regulations can play the alternative roles of screening the most efficient firms or protecting inefficient incumbents. They test which of these roles has been predominant using a difference-in-difference approach. They find evidence that regulations curb entry, increase the average size of firms at entry and lower the labour productivity growth of incumbents, strongly suggesting that these regulations are sheltering them from competitive pressures. The implications for resource reallocation are clear: inappropriate entry regulations tend to hamper the efficiency-enhancing role of firm demographics, distort the size distribution of firms and negatively affect aggregate productivity by lowering the incentives to improve efficiency in existing firms.

Daveri et al. (2010) focus on the direct effects of entry regulations on MFP growth of service sector firms in Italy and France over the 1995-2007 period. They measure regulations with detailed service sector information provided by the OECD for retail distribution, transport, communications and the professional services. They proceed in two steps: first they estimate the impact of entry restrictions on the market power of incumbents in these regulated sectors (as measured by mark-ups) and then relate this indicator of market power to the MFP growth of incumbents in the same sectors. They find that regulations curb firm-level productivity growth in regulated industries *via* a higher mark-up, namely regulations weaken competitive pressures and weaker competitive pressures slow down efficiency improvements.

**Figure 9. Percentage increase in MFP from easing anticompetitive regulation[a]**
**One standard deviation reduction in domestic and border entry barriers[b]**




Source*:* Table 1 columns 1 and 3 in Arnold, Nicoletti and Scarpetta (2011).

Notes: [a] Dynamic firms are defined as firms that catch up to the global frontier in their respective sector. [b] One standard deviation change is equivalent to moving from the level of regulation prevailing in Greece (one of the most regulated countries according to the OECD regulation impact indicators) to a situation of best practice regulation corresponding to adopting the least anticompetitive regulations observed in the OECD area in all sectors.

Arnold, Nicoletti and Scarpetta (2011) take a broader approach to investigate the impact of entry regulations on MFP growth. They take into account both the direct effects of regulations on firms' productivity in regulated non-manufacturing sectors (the "upstream" sectors) and the indirect effects of such regulations on firms in other ("downstream") sectors via intersectoral linkages, using the OECD indicators of regulation impact. They also account for firm heterogeneity by distinguishing between "dynamic" firms that catch up rapidly to the global frontier (for their sector) and firms that do not (the "non-dynamic" firms).[28] Their main results are summarised in Figure 9. Anticompetitive regulations are found to curb the productivity growth of all firms, dynamic and non-dynamic in both upstream and downstream sectors. On average, a substantial easing of such

[28] Arnold, Nicoletti and Scarpetta (2011) also distinguish between two sources of entry restrictions, domestic and border barriers (proxied by FDI restrictions). They find that these barriers are more harmful for dynamic firms that more rapidly approach the global frontier.

regulations is estimated to increase the productivity growth by over 1 per cent, implying an increase of over 10 per cent in the level of multifactor productivity in the long run. Interestingly, the estimated increase is significantly stronger for dynamic firms. Hence, regulation may have negative effects on the efficiency of resource reallocation by disproportionately hitting those firms that are driving improvements in aggregate productivity.

## 4. Concluding remarks

In this paper, we discuss theory and evidence that relate differences in the efficiency of resource reallocation and productivity performance across countries to anticompetitive product market regulations. We provide evidence that such regulations differ across countries and industries and have changed over time. Drawing on recent empirical studies, we find that regulations are of importance for performance.

We highlight three main sets of results:

There is solid evidence that the pace and depth of product market reforms are important for understanding both productivity and resource allocation outcomes. Countries and industries where direct and indirect regulatory burdens are lighter have generally experienced the highest GDP per capita and productivity growth rates in the studies we have surveyed.

Evidence at the firm level suggests that where regulatory burdens are lighter, the reallocation of resources towards the highest productivity firms is stronger. Moreover, firm-level productivity growth is also curbed by anticompetitive regulations.

The implications of inappropriate regulations for productivity performance are estimated to be quantitatively important. Therefore, reforming such regulations can provide a significant boost to potential growth in OECD economies.

The adverse effects of anticompetitive regulation on performance are often found to be non linear, with their intensity depending on the characteristics of countries, industries and firms. Some studies find the effects to be more severe for industries closer to international best practice and/or using more intensively new information technologies and for firms that are more dynamic. However, there is no consensus on the extent and direction of such differential effects and further research is needed to elucidate the interaction of regulation with levels of development and the heterogeneity of industries and firms.

## References

Acemoglu, D., Aghion, P. and Zilibotti, F. (2006), Distance to frontier, selection and economic growth, Journal of the European Economic Association 4, 37-74.

Aghion, P., Askenazy, P., Cette, G., Dromel, N. and Bourlès, R. (2009), Education, market rigidities and growth, Economics Letters 102, 62-65.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005), Competition and innovation: An inverted U relationship, Quarterly Journal of Economics 120, 701-728.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. and Prantl, S. (2004), Entry and productivity growth: Evidence from microlevel panel data, Journal of the European Economic Association 2, 265-276.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. and Prantl, S. (2006), The effects of entry on incumbent innovation and productivity, NBER Working Paper 12027.

Aghion, P., Fally, T. and Scarpetta, S. (2007), Credit constraints as a barrier to the entry and post-entry growth of firms, Economic Policy 22, 731-779.

Aghion, P. and Griffith, R. (2005), Competition and Growth, MIT Press, Cambridge MA.

Aghion, P., Harris, C., Howitt, P. and Vickers, J. (1997), Competition and growth with step-by-step innovation: An example, European Economic Review, Papers and Proceedings 41, 771-782.

Aghion, P. and Howitt, P. (1998), Endogenous Growth Theory, MIT Press, Cambridge MA.

Aghion, P. and Howitt, P. (2006), Joseph Schumpeter lecture: Appropriate growth policy: A unifying framework, Journal of the European Economic Association 4, 269-314.

Alesina, A., Ardagna, S., Nicoletti, G. and Schiantarelli, F. (2005), Regulation and investment, Journal of the European Economic Association 3, 791-825.

Arnold, J., Javorcik, B. and Mattoo, A. (2011), Does services liberalization benefit manufacturing firms? Evidence from the Czech Republic, in Journal of International Economics 85, 136-149.

Arnold, J., Nicoletti, G. and Scarpetta, S. (2008), Regulation, allocative efficiency and productivity in OECD countries: Industry and firm-level evidence, OECD Economics Department Working Papers 616, Paris.

Arnold, J., Nicoletti, G. and Scarpetta, S. (2011), Does anti-competitive regulation matter for productivity? Evidence from European firms, IZA Discussion Paper Series 5511, Bonn.

Askenazy P., Cahn, C. and Irac, D. (2008), Competition, R&D and the cost of innovation, PSE Working Papers 2008-32, PSE (Ecole normale supérieure), Paris.

Babetskii, I. and Campos, N.F. (2007), Does reform work? An econometric examination of the Reform-Growth Puzzle, IZA Discussion Papers 2638, Bonn.

Barone G. and Cingano, F. (2011), Service regulation and growth: Evidence from OECD countries, Economic Journal 121, 931-957.
Bartelsman, E.J. and Doms, M. (2000), Understanding productivity: Lessons from longitudinal microdata, Journal of Economic Literature 38, 569-595.
Bartelsman, E.J., Haltiwanger, J. and Scarpetta, S. (2004), Microeconomic evidence of creative destruction in industrial and developing countries, IZA Discussion Papers 1374, Bonn.
Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2009a), Measuring and analyzing cross-country differences in firm dynamics, in T. Dunne et al. (eds.), Producer Dynamics: New Evidence from Micro Data, NBER, Chicago University Press, Chicago.
Bartelsman, E.J., Haltiwanger, J. and Scarpetta, S. (2009b), Cross-Country differences in productivity: The role of allocation and selection, NBER Working Paper 15490.
Bartelsman, E.J., Scarpetta, S. and Schivardi, F. (2005), Comparative analysis of firm demographics and survival: Micro-level evidence for the OECD countries, Industrial and Corporate Change 14, 365-391.
Beck, T., Demirgüç-Kunt, A. and Maksimovic, V. (2004), Financial and legal constraints to firm growth: Does size matter?, Journal of Finance 60, 137-177.
Bergoeing, R., Kehoe, P.J., Kehoe, T.J. and Soto, R. (2002), Policy-driven productivity in Chile and Mexico in the 1980's and 1990's, American Economic Review 2, 16-21.
Bernard, A., Eaton, J., Jensen, J.B. and Kortum, S. (2003), Plans and productivity in international trade, American Economic Review 93, 1268-1290.
Bloom, D.E., Canning, D. and Sevilla, J. (2002) Technological diffusion, conditional convergence, and economic growth, NBER Working Paper 8713.
Bloom, N., Draca, M. Kretschmer, T. and Sadun, R. (2010), The economic impact of ICT, Final Report, Centre for Economic Performance, SMART N 2007/0020.
Blundell, R., Griffith, R. and van Reenen, J. (1999), Market share, market value and innovation in a panel of British manufacturing firms, Review of Economic Studies 66, 529-554.
Bouis, R., Duval, R. and Murtin, F. (2011), The policy and institutional drivers of economic growth across OECD and non-OECD economies: New evidence from growth regressions, OECD Economics Department Working Papers 843, Paris.
Boulhol, H., de Serres, A. and Molnar, M. (2008), The contribution of economic geography to GDP per capita, OECD Journal: Economic Studies 2008, 284-346.
Bourlès, R., Cette, G., Lopez, J., Mairesse, J. and Nicoletti, G. (2010), Do product market regulations in upstream sectors curb productivity growth? Panel data evidence for OECD countries, NBER Working Paper 16520.
Buccirossi, P., Ciari, L. Duso, T., Spagnolo, G. and Vitale, C. (2009), Competition policy and productivity growth: An empirical assessment, CEPR Discussion Papers 7470.
Caves, R.E. (1998), Industrial organization and new findings on the turnover and mobility of firms, Journal of Economic Literature 36, 1947-1982.
Conway, P. and Nicoletti, G. (2006), Product market regulation in the non-manufacturing sectors of OECD countries: Measurement and highlights, OECD Economics Department Working Papers 530, Paris.
Conway, P. and Nicoletti, G. (2007), Product market regulation and productivity convergence: OECD Evidence and implications for Canada, International Productivity Monitor nr 15.

Conway, P., de Rosa, D., Nicoletti, G. and Steiner, F. (2006), Product market regulation and productivity convergence, OECD Economic Studies 43, 39-76, Paris.
Corcos, G., Del Gatto, M., Mion, G. and Ottaviano, G. (2007), Productivity and firm selection: Intra- vs. international trade, Core Discussion Papers 2007/60, Université Catholique de Louvain.
Crafts, N. (2006), Regulation and productivity performance, Oxford Review of Economic Policy 22, 186-202.
Crespi, G., Criscuolo, C., Haskel, J. and Slaughter, M. (2007), Productivity growth, knowledge flows and spillovers, CEP Discussion Papers 785, Centre for Economic Performance, LSE.
Daveri F., Lecat, R. and Parisi, M.L. (2010), Service deregulation, competition and the performance of French and Italian firms, paper presented at the OECD-Banque de France conference on "Structural reforms, crisis exit strategies and growth", 9-10 December, http://www.banque-france.fr/gb/publications/telechar/seminaires/2010 /Daveri-Lecat-Parisi.pdf.
Davis, S., Haltiwanger, J. and Schuh, S. (1996), Job Creation and Destruction, MIT Press, Cambridge MA.
Del Gatto, M., Mion, G. and Ottaviano, G.I.P. (2006), Trade integration, firm selection and the costs of non-Europe, CEPR Discussion Papers 5730.
Faini, R., Haskell, J., Barba Navaretti, G., Scarpa, C. and C. Wey (2006), Contrasting Europe's decline: do product market reforms help?, in T. Boeri, M. Castanheira, R. Faini and V. Galasso (eds.) Structural Reforms without Prejudices. Oxford University Press.
Forlani, E. (2011), Competition in the service sector and the performances of manufacturing firms: Does liberalization matter?, CESifo Working Paper Series 2942, Munich.
Fosfuri, A., Motta, M. and Ronde, T. (2001), Foreign direct investment and spillovers through workers' mobility, Journal of International Economics 53, 205-222.
Foster, L., Haltiwanger, J. and Krizan, C.J. (2002), The link between aggregate and micro productivity growth: Evidence from retail trade, NBER Working Paper 9120.
Geroski, P. (1995), What do we know about entry?, International Journal of Industrial Organization 13, 421-440.
Golub, S. (2003), Measures of restrictions on inward foreign direct investment for OECD Countries, OECD Economic Studies 36, Paris.
Golub, S. and Koyama, T. (2006), OECD's FDI regulatory restrictiveness index: Revision and extension to more economies, OECD Economics Department Working Papers 525, Paris.
Görg, H. and Greenaway, D. (2002), Foreign direct investment and intra-industry spillovers: A review of the literature, GEP Research Paper 01/37, University of Nottingham.
Griffith, R. and Harrison, R. (2004), The link between product market reform and macro-economic performance, European Economy, Economic Papers 209.
Griffith, R., Harrison, R. and Simpson, H. (2006), Product market reform and innovation in the EU, CEPR Discussion Papers 5849.
Griffith, R., Redding, S. and Van Reenen, J. (2004), Mapping the two faces of R&D: Productivity growth in a panel of OECD industries, Review of Economics and Statistics 86, 883-895.

Griliches, Z. and Regev, H. (1995), Firm productivity in Israeli industry: 1979-1988, Journal of Econometrics 65, 175-203.
Gust, C. and Marquez, J. (2004), International comparisons of productivity growth: The role of information technology and regulatory practices, Labour Economics 11, 33-58.
Gwartney, J. and Lawson, R. (2003), The concept and measurement of economic freedom, European Journal of Political Economy 19, 405-430.
Haltiwanger, J., Scarpetta, S. and Schweiger, H. (2006), Assessing job flows across countries: The role of industry, Firm size and regulations, IZA Discussion Papers 2450, Bonn.
Haskel, J., Pereira, S. and Slaughter, M. (2007), Does inward foreign direct investment boost the productivity of domestic firms?, Review of Economics and Statistics 89, 482-496.
Hsieh, C-T. and Klenow, P. (2009), Misallocation and manufacturing TFP in China and India, Quarterly Journal of Economics 124, 1403-1448.
Inklaar R., Timmer, M.P. and van Ark, B. (2008), Market services productivity across Europe and the US, Economic Policy 23, 139-194.
Javorcik, B. (2004), Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages, American Economic Review 94, 605-627.
Keller, W. (2004), International technology diffusion, Journal of Economic Literature XLII, 752-782.
Klapper, L., Laeven, L. and Rajan, R. (2006), Entry regulation as a barrier to entrepreneurship, Journal of Financial Economics 82, 591-629.
Melitz, M. (2003), The impact of trade on intra-industry reallocations and aggregate industry productivity, Econometrica 71, 1695-1725.
Melitz, M. and Ottaviano, G. (2008), Market size, trade, and productivity, Review of Economic Studies 75, 295-316.
Micco, A. and Pagés, C. (2006), The economic effects of employment protection: Evidence from international industry-level data, IZA Discussion Papers 2433, Bonn.
Nickell, S. (1996), Competition and corporate performance, Journal of Political Economy 104, 724-746.
Nicoletti, G. and Scarpetta, S. (2003), Regulation, productivity and growth, Economic Policy 36, 11-72.
Nicoletti, G. and Pryor, F. (2006), Subjective and objective measures of governmental regulations in OECD nations, Journal of Economic Behavior and Organization 59, 433-449.
Okuno-Fujiwara, M. (1988), Interdependence of industries, coordination failure and strategic promotion of an industry, Journal of International Economics 25, 25-43.
Olley, G.S. and Pakes, A. (1996), The dynamics of productivity in the telecommunications equipment industry, Econometrica 64, 1263-1297.
Pakes, A. and Ericson, R. (1998), Empirical implications of alternative models of firm dynamics, Journal of Economic Theory 79, 1-45.
Parente, S.L. and Prescott, E.C. (1994), Barriers to technology adoption and development, Journal of Political Economy 102, 298-321.
Parente, S.L. and Prescott, E.C. (1999), Monopoly rights: A barrier to riches, American Economic Review 89, 1216-1233.

Poschke, M. (2010), The regulation of entry and aggregate productivity, Economic Journal 120, 1175-1200.

Rajan, R. and Zingales, L. (1998), Financial dependence and growth, American Economic Review 88, 559-586.

Restuccia, D. and Rogerson, R. (2007), Policy distortions and aggregate productivity with heterogeneous plants, NBER Working Papers 13018.

Rodriguez-Clare, A. (1996), Multinationals, linkages, and economic development, American Economic Review 86, 852-873.

Scarpetta, S. and Tressel, T. (2002), Productivity and convergence in a panel of OECD industries: Do regulations and institutions matter?, OECD Economics Department Working Papers 342, Paris.

Sutton, J. (1997), Gibrat's legacy, Journal of Economic Literature 35, 40-59.

de Serres, A., Kobayakawa, S., Slok, T. and Vartia, L. (2006), Regulation of financial systems and economic growth in OECD countries: An empirical analysis, OECD Economic Studies 43, 77-113.

Woelfl, A., Wanner, I., Kozluk, T. and Nicoletti, G. (2009), Ten years of product market reform in OECD countries – insights from a revised PMR indicator, OECD Economics Department Working Paper 695, Paris.

Woelfl, A., Wanner, I., Roehn, O. and Nicoletti, G. (2010), Product market regulation: Extending the analysis beyond OECD countries, OECD Economics Department Working Paper 799, Paris.

World Bank (2004), Doing Business in 2004, Washington, D.C.

# Comment on Arnold, Nicoletti and Scarpetta: Regulation, resource allocation and productivity growth

Mika Maliranta*

This paper addresses a topic that is both interesting and of utmost importance: how can policymakers raise the productivity level of a nation? Many agree with Paul Krugman's (1997) famous assertion that in the long run, productivity "is almost everything": it is the key driver of long-term economic growth and thereby, it is essential for the subjective well-being of citizens (see Sacks et al., 2010). In fact, multi-factor productivity (MFP), which measures the efficiency of total input use, is arguably a more ideal indicator of welfare than per capita output (Hulten, 2001). From a policy perspective, it is crucial to note that MFP can increase via two fundamentally different channels: 1) technical change (or productivity growth within plants/firms) and 2) the reallocation of inputs to those firms (and plants) that can use them more efficiently.

The authors provide an excellent account of the theoretical and empirical literature that emphasises the great importance of the (re)allocation process. It capably explains productivity growth over time – at least in some countries (OECD, 2003; Petrin et al., 2011) – and explains differences in productivity levels across countries (e.g., Bartelsman et al., 2009; Hsieh and Klenow, 2009). In Finland, for example, an increase in productivity-enhancing restructuring in the mid-1980's, especially between continuing plants, was essential for catching up with the international produc-

* ETLA, the Research Institute of the Finnish Economy and University of Jyväskylä, School of Business and Economics, mika.maliranta@etla.fi.

tivity frontier in the manufacturing sector that took place by the mid-1990's. Interestingly, intensified creative destruction coincided very closely with the deep-running deregulation of both the product and the financial markets since the mid-1980's (Honkapohja et al., 2009; Maliranta et al., 2010).

As pointed out by the authors, a large number of factors may hinder resources from being reallocated in a manner that is desirable from the point of view of society as a whole, at least in the long run. Clearly, there is a great demand for convincing empirical evidence regarding the policy measures by which a government may facilitate "creative destruction", not least of all because long-term gains (i.e., what constitutes the creative element) may materialize through indirect channels that are often hard to discern. This study pays special attention to one such mechanism: how regulations in upstream industries may curb productivity-enhancing restructuring in downstream industries.

Another reason why comprehensive analyses and discussions of this topic are badly needed is that restructuring mechanisms are likely to involve considerable challenges for some people. This may be the place where the growth-enhancing aspect of social security is likely to play a role. When well designed social welfare policies are effective, they cannot only facilitate the political acceptance of some painful tools for growth (like deregulation) but also encourage workers to seek higher-productivity jobs and firms to create those jobs in the midst of incessant creative destruction (Acemoglu and Shimer, 2000).

Indicating the motivation for the study, Figure 2 documents an interesting negative relationship between the stringency of product market regulations (*level*) in the years 1985-1995 and the *acceleration* of MFP growth after 1995. In fact, the correlation is even more pronounced when the MFP number for Finland is corrected from the erroneous −1.6 percent figure to the correct 0.3 percent figure. However, this correlation may be a temporary one. Although there are still considerable differences in the level of product market regulations across countries, it seems unlikely that MFP levels between countries will continue to diverge at an increasing rate in the future.[1] Interestingly, there seems to be a very strong *positive* correlation (r=0.62) between the *change* (between the years 1985-1995 and 1996-

[1] O'Mahoney and Timmer (2009) provide evidence that EU-15 countries have been lagging behind the U.S. since the mid-1990's in terms of MFP level.

2007) in product market regulation and the acceleration of MFP growth. This reflects the fact that those countries that have experienced decelerating MFP growth have typically also experienced the greatest deregulation (i.e., the greatest decrease in regulation) since the mid-1990's (Italy and Spain, for instance).[2]

In its empirical analysis, the study makes use of the popular static OP indicator of allocative efficiency (Olley and Pakes, 1996) and provides some very interesting results. However, the method has some well-known limitations. Its standard version cannot capture the separate contribution of entries and exits, which is unfortunate since the recent literature underlines the role of entries and exits (Bartelsman et al., 2009). Although entrants are typically small and tend to exhibit low productivity levels, they make a positive contribution to the allocation component of the standard Olley-Pakes method. Because the number of entrants is typically quite large, their contribution to the allocation component may be large (Maliranta and Määttänen, 2011). On the empirical side, it is interesting to note that countries like Spain and Italy seem to enjoy a high allocative efficiency (see Figure 7) given that they seem to have a relatively low MFP level in the market economy (O'Mahoney and Timmer, 2009).

In the future, it would be fascinating to use dynamic productivity decomposition to distinguish between the two types of reallocation: firm turnover via entries and exits and the reallocation of inputs between continuing firms/plants. A recent study by Inklaar et al. (2010) suggests that the role of intersectoral linkage (in this case, between banking and other private-sector industries) in the two main channels of industry productivity growth (i.e., technical change and reallocation between firms/plants) can be successfully analysed by dynamic productivity decomposition.

## References

Acemoglu, D. and Shimer, R. (2000), Productivity gains from unemployment insurance, European Economic Review 44, 1195-1224.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2009), Measuring and analyzing cross-country differences in firm dynamics, in T. Dunne et al. (eds.), Producer Dynamics, University of Chicago Press, Chicago.

[2] There does not seem to be any statistically significant relationship between the deregulation (between the years 1985-1995 and 1996-2007) and MFP growth (in the years 1996-2007).

Honkapohja, S., Koskela, E., Leibfritz, W. and Uusitalo, R. (2009), Economic Prosperity Recaptured. The Finnish Path from Crisis to Rapid Growth, MIT Press, Cambridge MA.

Hsieh, C-T. and Klenow, P.J. (2009), Misallocation and manufacturing TFP in China and India, Quarterly Journal of Economics 124, 1403-1448.

Hulten, C. (2001), Total factor productivity: A short biography, in C.R. Hulten et al. (eds.), New Developments in Productivity Analysis, University of Chicago Press (for NBER), Chicago.

Inklaar, R., Koetter, M. and Noth, F. (2010), Banking market competition, firm dynamics, and aggregate growth, manuscript, University of Groeningen.

Krugman, P. (1997), The Age of Diminished Expectations: U.S. Economic Policy in the 1990s, MIT Press, Cambridge MA.

Maliranta, M. and Määttänen, N. (2011), Allocation and industry productivity: Understanding the role of firm dynamics, manuscript, ETLA, Helsinki.

Maliranta, M., Rouvinen, P. and Ylä-Antilla, P. (2010), Finland's path to global productivity frontiers through creative destruction, International Productivity Monitor 20, 68-84.

O'Mahoney, M. and Timmer, M. (2009), Output, input and productivity measures at the industry level: The EU KLEMS Database, Economic Journal 119, F374-F403.

OECD (2003), The Sources of Economic Growth in OECD Countries, OECD, Paris.

Olley, G.S. and Pakes, A. (1996), The dynamics of productivity in the telecommunications equipment industry, Econometrica 64, 1263-1297.

Petrin, A., White, T.K. and Reiter, J.P. (2011), Impact of plant-level resource reallocations and technical progress on U.S. macroeconomic growth, NBER Working Paper 16700.

Sacks, D.W., Stevenson, B. and Wolfers, J. (2010), Subjective well-being, income, economic development and growth, CEPR Discussion Paper 8048.

# Human capital and productivity[*]

Angel de la Fuente[**]

*Summary*

This paper surveys the empirical literature on human capital and productivity and summarizes the results of my own work on the subject. On balance, the available evidence suggests that investment in education has a positive, significant and sizable effect on productivity growth. Moreover, according to my estimates, the social returns to investment in human capital are higher than those on physical capital in most EU countries and in many regions of Spain.

*Keywords:* Human capital, productivity, growth, measurement error.
*JEL classification numbers*: O40, I20, O30, C19.

---

[*] This paper has been prepared for a special issue of Nordic Economic Policy Review on productivity and competitiveness. It draws heavily on joint work with R. Doménech and other co-authors that has been partially financed by the European Commission, the OECD, the research department of BBVA and the Spanish Ministry of Science and Innovation (through grant no. ECO2008-04837/ECON and its predecessors). I would like to thank Pekka Ilmakunnas, Anders Sørensen, Jakob Madsen and other participants in the Helsinki NEPR conference for their useful comments and suggestions.

[**] Instituto de Análisis Económico (CSIC), angel.delafuente@iae.csic.es.

One of the most distinctive features of the "new" theories of economic growth has been the broadening of the concept of capital. While traditional neoclassical models almost exclusively focused on the accumulation of physical capital (equipment and structures), more recent contributions have attributed increasing importance to the accumulation of human capital and productive knowledge and to the interaction between these two factors. However, the empirical evidence has not always been consistent with the new theoretical models. In the case of human capital, in particular, a number of studies have produced discouraging results. Educational variables are often not significant or even enter with the "wrong" sign in growth regressions, particularly when these are estimated using differenced specifications or panel techniques. The accumulation of negative results in the literature during the second half of the nineties generated a growing skepticism about the role of schooling in the growth process and even led some authors (see in particular Pritchett, 2001) to seriously consider the reasons why educational investment may fail to contribute to productivity growth.

Many researchers in the area held on to more optimistic views, however. They (we) argued that the negative results found in certain studies could be explained by technical problems that have a great deal to do with the difficulty in correctly measuring human capital. This article provides a quick review of several strands of a literature that provides evidence in support of this hypothesis and a more detailed summary of my own work on the subject. The paper is organized as follows. Section 1 sketches the theoretical framework that has guided most studies on the contribution of education to economic growth, reviews the main empirical specifications used in the literature and briefly discusses some of its key results. Section 2 highlights some of the shortcomings of the cross-country schooling data sets most commonly used in the early empirical literature, discusses their implications for attempts to estimate the contribution of education to productivity growth and introduces a convenient indicator of data quality that can be used to quantify the information content of alternative schooling series and to estimate the size of the bias caused by measurement error. Section 3 summarizes the main findings of a series of papers I have written, mostly in collaboration with Rafael Doménech. In those papers, we construct new attainment series for 21 OECD countries and for the regions of Spain, develop measures of the information content of these and other schooling series and estimate a variety of growth specifications for both samples. Using these results, we have also constructed a set of meta-estimates of the coefficient of human capital in an aggregate Cobb-Douglas production function that corrects for the downward bias generated by measurement error. With this correction, we find that the contribution of investment in human capital to productivity growth is positive, quite sizable and implies rather respectable social returns that, for most territories in our two samples, compare quite favorably with those on physical capital and suggest that important externalities may be at work.

## 1. Human capital and economic growth: An overview of the literature

Theoretical models of human capital and growth are built around the hypothesis that the knowledge and skills embodied in humans directly raise productivity and increase an economy's ability to develop and adopt new technologies. In order to explore its implications and open the way for its empirical testing, this basic hypothesis is generally formalized in one of two (not mutually exclusive) ways. The simplest one involves introducing the stock of human capital (which will be denoted by $H$ throughout this paper) as an additional input in an otherwise standard aggregate production function linking national or regional output to the stocks of productive inputs (generally employment and physical capital) and to an index of technical efficiency or total factor productivity (TFP). The second possibility is to include $H$ in the model as a determinant of the rate of technological progress (that is, the rate of growth of TFP). This involves specifying a technical progress function that may include as additional arguments some indicator of investment in R&D and a measure of the "technological gap", that is, of the distance between each country's productive technology and the best practice frontier. I will refer to the first of these links between human capital and productivity as *level effects* (because the stock of human capital has a direct impact on the level of output) and to the second as *rate effects* (because $H$ affects the growth rate of output through TFP). Box 1 develops a simple model of growth with human capital that formalizes the preceding discussion and incorporates both effects.

*Box 1. A descriptive model of human capital and growth*

This box develops a simple model of growth and human capital that has two components: an aggregate production function and a technical progress function. The production function will be assumed to be of the Cobb-Douglas type:

$$Y_{it} = A_{it} K_{it}^{\alpha_k} H_{it}^{\alpha_h} L_{it}^{\alpha_l}, \qquad \text{(B.1)}$$

where $Y_{it}$ denotes the aggregate output of territory *i* at time *t*, $L_{it}$ is the level of employment, $K_{it}$ the stock of physical capital, $H_{it}$ the average stock of human capital per worker, generally measured by school attainment, and $A_{it}$ an index of technical efficiency or total factor productivity (TFP) which summarizes the current state of the technology and captures omitted factors such as geographical location, climate, institutions and endowments of natural resources. The coefficients $\alpha_i$ (with *i* = *k*, *h*, *l*) measure the elasticity of output with respect to the stocks of the different factors. An increase of 1 percent in the stock of human capital per worker, for instance, would increase output by $\alpha_h$ percent, holding constant the stocks of the other factors and the level of technical efficiency.

Under the standard assumption that (B.1) displays constant returns to scale in physical capital and labor while holding average attainment constant, (i.e. that $\alpha_k + \alpha_l = 1$), we can define a per capita production function that will relate average labor productivity to average schooling and to the stock of capital per worker. Letting $Q = Y/L$ denote output per worker and $Z = K/L$ the stock of capital per worker and dividing both sides of (B.1) by total employment, *L*, we have:

$$Q = AZ^{\alpha_k} H^{\alpha_h}. \qquad \text{(B.2)}$$

The technical progress function describes the determinants of the growth rate of total factor productivity. I will assume that country *i*'s TFP level can be written in the form:

$$A_{it} = B_t X_{it}, \qquad \text{(B.3)}$$

where $B_t$ denotes the world "technological frontier" (i.e. the maximum attainable level of efficiency in production given the current state of scientific and technological knowledge) and $X_{it} = A_{it}/B_t$ is (an inverse indicator of) the "technological gap" between country *i* and the world frontier. It will be assumed that $B_t$ grows at a constant and exogenous rate, *g*, and that the growth rate of $X_{it}$ is given by

$$\Delta x_{it} = \gamma_{io} - \lambda x_{it} + \gamma H_{it}, \tag{B.4}$$

where $x_{it}$ is the log of $X_{it}$ and $\gamma_{io}$ is a country-fixed effect that helps control for omitted variables such as R&D investment. Notice that this specification incorporates a technological diffusion or *catch-up* effect. If $\lambda > 0$, countries that are closer to the technological frontier will experience lower rates of TFP growth. As a result, relative TFP levels will tend to stabilize over time and their steady-state values will be partly determined by the level of schooling.

Some recent theoretical models suggest that the accumulation of human capital may give rise to important externalities that would justify corrective public interventions. The problem arises because some of the benefits of a more educated labor force will typically "leak out" and generate output gains that cannot be appropriated in the form of higher earnings by those undertaking the relevant investment, thereby driving a wedge between the private and social returns to education. Lucas (1988), for example, suggests that the average stock of human capital at the economy-wide level increases productivity at the firm level, holding the firm's own stock of human capital constant. It is also commonly assumed that the rate effects of human capital on technical progress include a large externality component because it is difficult to privately appropriate the full economic value of new ideas. Azariadis and Drazen (1990), and implicitly also Lucas (1988), stress that younger cohorts are likely to benefit from the knowledge and skills accumulated by their elders, thus generating potentially important intergenerational externalities that operate both at home and at school. The literature also suggests that human capital can generate more diffuse "civic" externalities, as an increase in the educational level of the population may help reduce crime rates or contribute to the development of more effective institutions.

### *1.1 From theory to data: Alternative approaches to empirical analysis*

Empirical studies of the effects of human capital on productivity (or more broadly, of the determinants of economic growth) have followed one of two alternative approaches. The first involves the specification and estimation of an *ad-hoc* equation relating growth in total or per capita output to a set of variables that are considered to be relevant on the basis of informal theoretical considerations. The second approach is based on the estimation of a structural relation between the level of output or its growth rate and the relevant explanatory variables that is derived from an explicit theoretical model built around an aggregate production function and, possibly, a technical progress function of the type described in Box 1. This basic framework for the "structural" analysis of the determinants of growth can give rise to a large number of empirical specifications. Some of the most common examples are discussed in Box 2. The production function can be directly estimated with the relevant variables expressed in levels or in growth rates when reliable data are available for the stocks of all relevant production inputs. Alternatively, its parameters can be recovered from other specifications (*convergence* and *steady state* equations) that are designed for estimation when only data on investment flows (rather than factor stocks) are available. These specifications can be derived from a production function by replacing factor stocks or their growth rates by convenient approximations in terms of investment rates using the procedure developed by Mankiw et al. (1992) within the framework of a generalized Solow model with several types of capital.

*Box 2. Some common empirical specifications*

For estimation purposes, it is generally convenient to work with the production function written in logarithms or in growth rates. Using lower case letters to denote logarithms, and the combination of lower case letters and the symbol "Δ" to denote growth rates, the production function given by equation (B.1) in Box 1 yields the two following specifications:

$$y_{it} = a_{it} + \alpha_k k_{it} + \alpha_h h_{it} + \alpha_l l_{it} + \varepsilon_{it}, \qquad \text{(B.5)}$$

$$\Delta y_{it} = \Delta a_{it} + \alpha_k \Delta k_{it} + \alpha_h \Delta h_{it} + \alpha_l \Delta l_{it} + \Delta\varepsilon_{it}, \qquad \text{(B.6)}$$

where $\varepsilon_{it}$ and $\Delta\varepsilon_{it}$ are stochastic disturbances.

One difficulty that arises at this point is that both these equations contain terms that are not directly observable (in particular the level of TFP, $a_{it}$, or its growth rate, $\Delta a_{it}$). To proceed with the estimation, it is necessary to make further assumptions about the behavior of these terms. Different assumptions will generate different econometric specifications. The simplest possibility is to assume that the rate of technical progress is constant over time and across countries, i.e. that $\Delta a_{it} = g$ for all *i* and *t*. In this case, *g* can be estimated as the regression constant in equation (B.6) and $a_{it}$ is replaced in equation (B.5) by $a_{io} + gt$, where $a_{io}$ and *g* give rise to country-specific constants and a common trend, respectively. An alternative and more sophisticated approach involves specifying $\Delta a_{it}$ as a function of other variables in equation (B.6). One possible specification is the one given by the technical progress function described by equations (B.3) and (B.4) in Box 1.

When data on factor stocks or their growth rates are not available (or are not considered reliable), a generalized Solow model can be used to approximate these variables in terms of observed investment rates. In such a model, the long-term equilibrium values of factor ratios are simple functions of investment rates, and the behavior of these ratios away from such an equilibrium can be approximated as a function of investment rates and initial output per worker. If we are willing to assume that most countries are reasonably close to their long-run equilibria, equation (B.5) can be replaced by an equation relating output per worker to investment rates in physical and human capital. Otherwise, the relevant equation will involve the growth rate of output and it will include initial output per worker as an additional regressor in order to pick up transitional dynamics along the adjustment to the long-run equilibrium. Two rather standard specifications of the resulting *steady state* and *convergence equations* (which do not allow for rate effects) would be

$$q_{it} = a_{it} + gt + \frac{\alpha_k}{1-\alpha_k\alpha_h}\ln\frac{s_{kit}}{\delta+g+n_{it}} + \frac{\alpha_h}{1-\alpha_k-\alpha_h}\ln\frac{s_{hit}}{\delta+g+n_{it}}, \qquad \text{(B.7)}$$

and

$$\Delta q_{it} = g + \beta\left(a_{io} + gt\right) + \beta\left(\frac{\alpha_k}{1-\alpha_k\alpha_h}\ln\frac{s_{kit}}{\delta+g+n_{it}} + \frac{\alpha_h}{1-\alpha_k-\alpha_h}\ln\frac{s_{hit}}{\delta+g+n_{it}}\right) - \beta q_{it}. \quad \text{(B.8)}$$

where $q$ is the log of output per worker, $s_k$ and $s_h$ stand for investment in physical and human capital measured as a fraction of GDP, $n$ for the rate of growth of employment or the labor force and $\delta$ for the rate of depreciation (which is assumed to be the same for both types of capital). Parameter $\beta$ measures the speed of convergence towards the long-run equilibrium or steady state and can be shown to be a function of the degree of returns to scale in both types of capital considered jointly and of the length of the period over which we are taking observations.

### *1.2 Empirical evidence: A bird's eye view*

A large number of empirical studies have analyzed the relationship between human capital and economic growth using the different specifications outlined above.[1] Early attempts in this direction, by and large, produced positive results that tended to confirm economists' traditionally optimistic views regarding the macroeconomic payoff to investment in education. Landau (1983), Baumol et al. (1989), Barro (1991) and Mankiw et al. (1992), among many others, find a variety of educational indicators to have the expected positive effect on output growth. During the second half of the nineties, however, a new round of empirical papers produced rather disappointing results on the effects of schooling on aggregate productivity. Unlike most previous studies, most of these papers used pooled quinquennial data and relied on either panel techniques or the use of differenced specifications to control for unobserved country heterogeneity. In this setting, educational variables were often found to be insignificant or even entered with the "wrong" sign in growth regressions (see, for instance, Benhabib and Spiegel, 1994, Islam, 1995, Caselli et al., 1996 and Pritchett, 2001).

While some researchers have been willing to take such counterintuitive results at face value, many others have been rather skeptical (see, for in-

[1] For a more detailed survey of the relevant literature, see Section 3 of the Appendix to de la Fuente and Ciccone (2003).

stance, Barro, 1997). These authors have tended to attribute negative results on schooling and growth to various econometric and specification problems and poor data quality. Measurement error, in particular, has been widely recognized to be a potentially important problem for two reasons. First, because the series of average years of schooling commonly used in the literature are likely to contain a great deal of noise and, second, because years of schooling can be expected to be a very imperfect measure of skills in any event. In addition, the first problem is likely to be particularly important in a panel setting, where parameter estimates rely heavily on the time-series variation of the data, because measurement error arising from changes in classification and data collection criteria tends to generate a great deal of spurious volatility in the schooling series that will make it difficult to identify its contribution to productivity growth.

Although it is too early for the issue to have been conclusively settled, my reading of the evidence accumulated over the last decade or so is optimistic. We have good reasons to believe that the negative results found in some of the previous literature can indeed be largely attributed to deficiencies in the human capital data used in earlier studies. Papers making use of improved data sets on attainment or econometrically allowing for measurement error strongly suggest that increases in average schooling do indeed have a substantial impact on productivity growth. The results are generally even stronger and sharper when direct measures of skill levels are used to proxy for human capital, suggesting that improvements in the quality of schooling can have an even larger effect on aggregate output than increases in its quantity.

The wave of negative results on the growth effects of education that occurred in the second half of the nineties is clearly associated with the introduction of panel data techniques. While early studies relied on cross-section data (working with a single observation per country that described average behavior over a period of several decades), studies in the second group have used several observations per country, taken over shorter periods, and have employed panel techniques or differenced specifications that basically eliminate the cross-section variation in the data before proceeding to the estimation. While theses estimation techniques have the important advantage that they control for unobservable differences across countries, they also have some disadvantages. Perhaps the main one is that they are more sensitive to measurement error in the data as errors tend to be greater in the time-series than in the cross-section dimension because they tend to cancel out when we work with averages over long periods. As I have already noted, this suggests that a possible explanation for the negative results obtained in panel data studies is related to the poor quality of the schooling data that have been used until recently in the growth literature. As we will see below, most of the earlier databases on international schooling levels contain large amounts of noise that can be traced back to various inconsistencies of the primary data used to construct them. The existence of this noise induces a downward bias in the estimation of the coefficients that measure the impact of human capital (that is, a tendency to underestimate their values) because it generates spurious variability in the stock of human capital that is not matched by proportional changes in the level of productivity.

A number of recent studies provide evidence that is consistent with this hypothesis. Starting with Krueger and Lindhal (K&L, 2001), some authors have constructed statistical indicators of the informational content of different attainment series (*reliability ratios*) that can be used to calculate the likely size of the attenuation bias and conclude that the value of this ratio is sufficiently low to explain the lack of significance of educational indicators in previous studies. Other authors, including Cohen and Soto (2007), de la Fuente and Doménech (D&D, 2001a, 2001b, 2006) and Barro and Lee (2010), have tried to improve the signal-to-noise ratio in the schooling series by exploiting new sources of information and introducing different corrections. They find that the results concerning the impact of education on growth improve considerably when these revised series are used. I will return to these issues in much greater detail in the next two sections.

Another interesting development is the use of cross-country data on direct measures of skill that may provide better proxies for the stock of human capital than years of schooling. While such data are still rather scarce, some recent papers suggest this to be likely to be a very fruitful line of research. Hanushek and several coauthors[2] construct indicators of labor force quality using mean country scores in a number of international student achievement tests in mathematics, science and reading, while Coulombe et al. (2004) use data drawn from IALS, an international study on the skill level of the adult population conducted by the OECD and Statis-

[2] See among others Hanushek and Kimko (2000) and Hanushek and Woessman (2008, 2009).

tics Canada. In both cases, the results of growth regressions point to even larger output effects than those obtained using revised attainment data. While not entirely free of problems, these estimates do suggest that the quality of education is likely to be at least as important as its quantity and that the return to improvements in schooling quality could be extraordinarily high, for not only are their expected benefits large but the relevant costs will generally also be much lower than those of increasing attainment for they do not involve a further sacrifice of student time and output.

## 2. Cross-country data on schooling: Problems and consequences

Most governments gather information on a number of educational indicators through population censuses, labor force surveys and specialized studies and surveys. Various international organizations collect these data and compile comparative statistics that provide easily accessible and (supposedly) homogeneous information for a large number of countries. The most comprehensive regular source of international educational statistics is UNESCO's *Statistical Yearbook*. This publication provides reasonably complete yearly time series on school enrollment rates by level of education for most countries in the world and contains some data on the educational attainment of the adult population, government expenditures on education, teacher/pupil ratios and other variables of interest.[3]

The UNESCO enrollment series have been used in a large number of empirical studies of the link between education and productivity. In many cases, this choice reflects the easy availability and broad coverage of these data rather than their theoretical suitability for the purpose of the study. Enrollment rates can probably be considered to be an acceptable, although imperfect, proxy for the flow of educational investment but they are not necessarily a good indicator of the existing stock of human capital since average educational attainment (which is often the more interesting varia-

[3] Other useful sources include the UN's Demographic Yearbook, which also reports educational attainment levels by age group and, in recent years, the OECD's annual report on education in its member countries (Education at a Glance), which contains a great deal of information about the inputs and outputs of the educational system.

ble from a theoretical point of view) responds to enrollment flows only gradually and with a very considerable lag.

In an attempt to remedy these shortcomings, a number of researchers have constructed data sets that attempt to directly measure the educational stock embodied in the population or labor force of large samples of countries during a period of several decades. These data sets have generally been constructed by combining the available data on attainment levels with the UNESCO enrollment figures to obtain series of average years of schooling and of the composition of the population or labor force by educational level. The best known early attempts in this line are the work of Kyriacou (1991), the first versions of the Barro and Lee data set (1993, 1996, 2000) and the series constructed by World Bank researchers (Lau, Jamison and Louat, 1991; Lau, Bhalla and Louat, 1991; Nehru et al. (NSD), 1995).

In de la Fuente and Doménech (D&D, 2006) we briefly review the methodology used in these studies and compare the different data sets to each other, focusing in particular on the OECD, where the quality of the available information should, in principle, be better than in developing countries. The analysis of the different series reveals very significant discrepancies among them in terms of the relative positions of many countries and implausible estimates or time profiles for at least some of them. Although the different studies generally coincide when comparisons are made across broad regions (e.g. the OECD vs. LDCs in various geographical areas), the discrepancies are very important when we focus on the group of industrialized economies. Another cause for concern is that existing estimates often display extremely large changes in attainment levels over periods as short as five years (particularly at the secondary and tertiary levels).

To a large extent, these problems have their origins in the deficiencies of the underlying primary data. As noted by Behrman and Rosenzweig (1994), there are good reasons to worry about the accuracy and consistency of UNESCO's data on both attainment levels and enrollment rates. Our analysis of the different schooling data sets confirms this diagnostic and suggests that many of the problems detected in these data can be traced back to shortcomings of the primary statistics, which do not seem to be consistent, across countries or over time, in their treatment of vocational and technical training and other courses of study, and reflect at times the number of people who have started a certain level of education and, at others, those who have completed it.

*2.1 Attenuation bias and a measure of data quality*

The poor quality of cross-country schooling data is a serious concern because it tends to obscure the relationship between the variables of interest and generates a tendency to underestimate the impact of human capital on productivity. To understand the origin of the *attenuation bias* caused by measurement error, assume the level of productivity, $Q$, to be a linear function of the stock of human capital, $H$, given by

$$Q = bH + u, \tag{1}$$

where $u$ is a random disturbance. Given this relationship, variations in the stock of human capital, $H$, will induce changes in $Q$, and the relative magnitude of the variations in these two variables will allow us to estimate the value of the coefficient $b$. Now, if $H$ is measured with error, that is, if what we observe is not $H$ itself but a noisy proxy for it, say

$$P = H + \varepsilon, \tag{2}$$

where $\varepsilon$ is a random measurement error term, then part of the apparent variation in the stock of human capital (over time and across countries) will be due to measurement error – that is, it will be noise rather than a true signal. Since such variations do not logically induce any response in $Q$, this variable will appear to be less sensitive to $H$ than it really is, thereby biasing the estimated value of $b$ toward zero.

In summary, attenuation bias arises because measurement error introduces "noise" that tends to hide the true relationship between the variables of interest. It can be shown that the size of the bias will be inversely related to the information content of the series, as measured by its *reliability ratio*, $r$. This indicator is defined as the ratio between the signal and the sum of the signal and noise contained in the data, that is,

$$r \equiv \frac{\operatorname{var} H}{\operatorname{var} P} = \frac{\operatorname{var} H}{\operatorname{var} H + \operatorname{var} \varepsilon}, \tag{3}$$

where *var H* measures the signal contained in the series (i.e. the true variation in human capital) and *var* $\varepsilon$ the noise that distorts it.[4] This ratio is very useful, first because it provides an indicator of the information content of each series, and second because the error in the estimation will be inversely proportional to its value. As a result, the reliability ratio can be used to correct the attenuation bias so as to obtain consistent estimators of the parameter of interest (i.e. estimators that are not biased in large samples).

Since $H$ and $\varepsilon$ are not observed separately, reliability ratios cannot be directly computed. They can, however, be estimated using a procedure developed by Krueger and Lindhal (2001) whenever several noisy proxies are available for the variable of interest. Box 3 describes this procedure and an extension of it developed by de la Fuente and Doménech (2006).

*Box 3. Estimating reliability ratios*

Let $P_1 = H + \varepsilon_1$ and $P_2 = H + \varepsilon_2$ be two alternative proxies for the stock of human capital, *H*. It is easy to check that if the error terms of the two series, $\varepsilon_1$ and $\varepsilon_2$, are not correlated with each other, the covariance between $P_1$ and $P_2$ can be used to estimate the variance of *H*, which is the only unknown magnitude in equation (3). It follows that, under this assumption, $r_1$ can be estimated as

$$\hat{r}_1 = \frac{\operatorname{cov}(P_1, P_2)}{\operatorname{var} P_1}, \tag{B.9}$$

which turns out to be the formula for the OLS estimator of the slope coefficient of a regression of $P_2$ on $P_1$. Hence, to estimate the reliability of $P_1$ we run a regression of the form $P_2 = c + r_1 P_1$.[5] Notice, however, that if the measurement errors of the two series are positively correlated $\left(E\varepsilon_1\varepsilon_2 > 0\right)$

[4] Notice that the denominator of the last expression given in (3) implicitly assumes that the measurement error term, $\varepsilon$, is not correlated with *H*.

[5] Intuitively, regressing $P_2$ on $P_1$ gives us an idea of how well $P_1$ explains the true variable *H* because measurement error in the dependent variable ( $P_2$ in this case) will be absorbed by the disturbance without generating a bias. Hence, it is almost as if we were regressing the true variable on $P_1$ .

as may be expected in many cases, $\hat{r}_1$ will overestimate the reliability ratio and hence understate the extent of the attenuation bias induced by measurement error.

In de la Fuente and Doménech (2006), we develop an extension of this procedure that can be used to construct a minimum-variance estimator of the reliability ratio whenever more than two noisy proxies are available for the same underlying variable, working under the maintained assumption that measurement errors are uncorrelated across data sets. As in Krueger and Lindahl, the reliability ratio $r_k$ of a given series of average years of schooling (say $S_k$) is estimated by using $S_k$ to try to explain alternative estimates of the same variable ($S_i$ with $j \neq k$). The main difference is that, rather than running a set of independent pairwise regressions with different data sets, the efficient estimator of the reliability ratio for a given data set *j* can be obtained as the slope coefficient of a restricted SUR model of the form

$$P_k = c_k + r_j P_j + u_k \text{ for } k = 1..., K \tag{B.10}$$

where we constrain $r_i$ to be the same for all "reference" data sets, *k*, used in the left-hand side of the system (B.10), and *k* varies over the last available version of all data sets different from *j*. The reliability ratio of Barro and Lee's (2000) data set, for instance, is estimated using these authors' estimate of average years of schooling as the explanatory variable in a set of SUR regressions where the reference (dependent) variables are the average years of schooling estimated by Kyriacou (1991), Nehru et al. (1995), Cohen and Soto (2007) and de la Fuente and Doménech (2006) and the slope coefficient is forced to be the same in all cases. Other versions of the Barro and Lee data set, however, are not used as a reference because the correlation of measurement errors across the same family of schooling series is almost certainly very high and this will artificially inflate the estimated reliability ratio.

# 3. Some results for the OECD countries and the Spanish regions

The preceding discussion suggests two complementary ways of dealing with the problems caused by poor schooling data. One is to try to improve the quality of the data by drawing on new primary sources and introducing various corrections to neutralize the effects of changes in classification criteria, and the other is to use estimates of reliability ratios to correct for attenuation bias. In a series of related papers, Rafael Doménech and myself (D&D 2000, 2001a, 2001b, 2002, 2006, 2008) have followed both these strategies using data for 21 OECD countries and for the regions of Spain. In both cases, the first step has been to construct new schooling series which attempt to increase the signal to noise ratio. In the case of Spain (de la Fuente and Doménech, 2008), the task has been relatively simple since the required primary information is readily available in the decennial censuses and in municipal registers, both of which have been compiled using clear and relatively stable classification criteria.

## *3.1 A new data set*

The OECD series (de la Fuente and Doménech 2000, 2001b, 2006) required considerably more work. We first collected all information we could find on the distribution of the adult population by educational level in OECD countries. We used both international publications and national sources (census reports and surveys, statistical yearbooks and unpublished data supplied by national governments and by the OECD in response to a request for information that was accompanied by a preliminary version of our data set). Next, we tried to reconstruct a plausible time profile of attainment in each country using all available data and a bit of common sense. For those countries for which reasonably complete series were available, we primarily relied on national sources. For the rest, we started from the most plausible set of attainment estimates available around 1990 or 1995 (taken generally from OECD sources) and proceeded backwards, trying to avoid unreasonable jumps in the series that could only reflect changes in classification criteria. In some cases, the construction of the series involved subjective judgments to choose among alternative census or survey estimates when several were available. At times, we have also reinterpreted some of the data from international compilations as referring to somewhat broader or narrower schooling categories than the reported one.[6] Missing data points lying between available census observations were filled in by simple linear interpolation. Missing observations prior to the first census observation were estimated, whenever possible, by backward extrapolations that made use of census information on attainment levels disaggregated by age group.[7]

### *3.2 How good are different schooling series?*

In de la Fuente and Doménech (D&D, 2002, 2006), we use the procedure described in Box 3 to estimate the reliability ratios of the series of years of schooling most commonly used in the growth literature, restricting ourselves to the sample of 21 OECD countries covered by the data set described in the previous section. This indicator is constructed for several transformations of the series of average years of schooling after removing period means from all series so as to eliminate fixed time effects. In particular, we estimate reliability ratios for years of schooling measured in levels $(S_{it})$ and in logs $(s_{it})$, for average annual changes in both levels and logs measured across successive quinquennial observations ($\Delta S_{it}$ and $\Delta s_{it}$) and for log years of schooling measured in deviations from their country means $(s_{it} - s_i)$. Notice that $\Delta s_{it}$ corresponds to annual growth rates and $s_{it} - s_i$ is the "within" transformation often used to remove fixed effects.

The results are shown in Table 1 with the different data sets arranged by decreasing average reliability ratios. The last row of the table shows the average value of the reliability ratio for each type of data transformation

[6] Clearly, the construction of our series involves a fair amount of guesswork. Our methodology looks decidedly less scientific than the apparently more systematic estimation procedures used by other authors starting from supposedly homogeneous data. However, even a cursory examination of the data shows that there is no such homogeneity. Hence, we have found it preferable to rely on judgment to try to piece together the available information in a coherent manner than to take for granted the accuracy of the primary data. The results do look more plausible than most existing series, at least in terms of their time profile and, as I will show below, perform rather well in terms of a statistical indicator of data quality.

[7] A closely related paper, both in terms of its objectives and its methodology, is Cohen and Soto (2007). These authors construct a schooling data set for a much larger sample of countries using census and survey data from UNESCO, the OECD's in-house educational data base, and the websites of national statistical agencies, together with enrollment rates from UNESCO and other sources.

(taken across data sets) and the last column displays the average reliability ratio of each data set (taken across transformations). Our mean estimate of the reliability ratio for all series and transformations is 0.335. Since this variable must lie between zero and one (with zero indicating that the series only contains noise and one that it is measured without error),[8] this result suggests that the average estimate of the coefficient of schooling in a growth equation is likely to suffer from a substantial downward bias, even without taking into account the further loss of signal that arises when additional regressors are included in these equations (see de la Fuente and Doménech, 2006). The bias will be smaller when the data are used in levels or logs, but is likely to be very large in fixed effects or differenced specifications. The average reliability ratio is only 0.254 for the data in quinquennial log differences and 0.090 for level differences taken at the same frequency.

**Table 1. SUR estimates of reliability ratios, OECD sample**

| | $S_{it}$ | $s_{it}$ | $\Delta S_{it}$ | $\Delta s_{it}$ | $s_{it}$ - $s_i$ | $\Delta s_{it}$ - $\Delta s_i$ | Average |
|---|---|---|---|---|---|---|---|
| D&D (2002) | 0.754 | 0.775 | 0.337 | 0.769 | 0.917 | 0.246 | 0.633 |
| C&S (2001) | 0.806 | 0.912 | 0.330 | 0.467 | 0.547 | 0.185 | 0.541 |
| D&D (2000) | 0.720 | 0.761 | 0.100 | 0.550 | 0.818 | 0.074 | 0.504 |
| Kyr. (1991) | 0.723 | 0.600 | 0.024 | 0.065 | 0.111 | 0.026 | 0.258 |
| B&L (2000) | 0.707 | 0.603 | -0.018 | 0.045 | 0.178 | -0.016 | 0.250 |
| B&L (1996) | 0.559 | 0.516 | -0.017 | 0.039 | 0.146 | -0.007 | 0.206 |
| B&L (1993) | 0.526 | 0.436 | -0.019 | 0.029 | 0.121 | -0.017 | 0.179 |
| NSD (1995) | 0.278 | 0.330 | -0.021 | 0.066 | 0.095 | -0.115 | 0.106 |
| Average | 0.634 | 0.617 | 0.090 | 0.254 | 0.367 | 0.047 | 0.335 |

Source: de la Fuente and Doménech (2006).

Notes: All series are measured in deviations from their respective sample means in each period prior to estimation. Key: *D&D* = de la Fuente and Doménech (preliminary and final versions); *C&S* = Cohen and Soto (data taken from the working paper version published in 2001); *Kyr* = Kyriacou; *B&L* = Barro and Lee; *NSD* = Nehru et al.

Our results indicate that the importance of measurement error varies significantly across data sets, although their precise ranking depends on the

[8] This is true as long as the measurement error terms of the different series are uncorrelated with each other and with H. As can be seen in Table 1, some of our estimates of the reliability ratio lie outside this interval, which implies some violation of this assumption. In de la Fuente and Doménech (2002), we construct alternative estimates of reliability ratios under more general assumptions and find that the required corrections do not qualitatively change the results.

data transformation that is chosen. Two of the datasets that are most widely used in early cross-country empirical work, those by Kyriacou (1991) and Barro and Lee (various years), perform relatively well when the data are used in levels but, as noted by Krueger and Lindhal (2001), contain very little signal when the data are differenced. Efforts to increase the signal content of the schooling data seem to have been at least partially successful, although the attenuation bias continues to be potentially large even in these cases. Taking the average reliability ratio for the 1996 version of the Barro and Lee data set (0.206) as a reference, the 2000 revision of these series by the same authors has increased their information content by 21 percent, while the estimates reported in Cohen and Soto (2001) and in de la Fuente and Doménech (2002) raise the estimated reliability ratio by 162 percent and 207 percent, respectively.

### *3.3 Data quality and estimates of the growth effects of human capital in the OECD*

As we have seen in the previous section, the expected severity of the attenuation bias is a decreasing function of the reliability ratio of the series used in the estimation. This suggests that the estimated value of the coefficient of human capital in a growth regression should increase with the quality of the schooling data. In de la Fuente and Doménech (D&D, 2002, 2006) we show this to indeed be the case. We estimate various specifications of an aggregate production function using the different schooling series for the OECD countries analyzed in the previous section as alternative proxies for the stock of human capital. We find that both the size and the significance of the coefficient of schooling increase as expected with the reliability ratio. Finally, we exploit this correlation to construct a set of "meta-estimates" of the parameter of interest that corrects for measurement error bias.[9]

#### *Results with different schooling series*

The equations we estimate are derived from a Cobb-Douglas aggregate production function with constant returns to scale that includes as inputs

---

[9] A meta-estimate is an estimate that is not directly obtained from the data but that is constructed using other primary estimates.

the stock of physical capital, the level of employment and the average level of education of the adult population. This equation is estimated in levels (with the variables measured in logarithms), in levels with fixed country effects and in first differences. In de la Fuente and Doménech (2002) we also estimate a fourth specification in differences that includes fixed country effects and incorporates a process of technological diffusion or catch-up. In this specification, the rate of growth of TFP is directly proportional to the technological distance between each country and the US, and the fixed country effects capture permanent differences in TFP levels that will presumably reflect differences in R&D expenditure and other omitted variables.[10]

**Table 2. Alternative estimates of the human capital coefficient *($\alpha_s$)* using different specifications and schooling series**

| | NSD | KYR | B&L93 | B&L96 | B&L00 | C&S | D&D00 | D&D02 | Avge. |
|---|---|---|---|---|---|---|---|---|---|
| Levels | 0.078 | 0.186 | 0.141 | 0.165 | 0.238 | 0.397 | 0.407 | 0.378 | 0.249 |
| | (2.02) | (2.18) | (4.49) | (4.82) | (6.19) | (7.98) | (7.76) | (6.92) | (5.30) |
| Fixed eff. | 0.068 | 0.066 | 0.136 | 0.115 | 0.203 | 0.608 | 0.627 | 0.958 | 0.348 |
| | (0.76) | (1.86) | (3.30) | (1.80) | (3.74) | (4.49) | (3.99) | (6.51) | (3.31) |
| Differences | 0.079 | 0.009 | 0.089 | 0.083 | 0.079 | 0.525 | 0.520 | 0.744 | 0.266 |
| | (0.70) | (0.15) | (2.52) | (1.47) | (1.28) | (2.57) | (2.17) | (3.10) | (1.75) |
| Catch-up | - 0.206 | 0.014 | 0.056 | -0.007 | -0.019 | 0.573 | 0.587 | 0.540 | 0.192 |
| | (1.61) | (0.29) | (1.80) | (0.11) | (0.31) | (3.52) | (3.47) | (2.89) | (1.24) |
| Average | 0.005 | 0.069 | 0.106 | 0.089 | 0.125 | 0.526 | 0.535 | 0.655 | |
| | (0.47) | (1.12) | (3.03) | (2.00) | (2.73) | (4.64) | (4.35) | (4.86) | |

Source: de la Fuente and Doménech (2002).

Note: Key: See the notes to Table 1.

These specifications are estimated using quinquennial data for our OECD sample that cover the period 1960-1990. All equations include fixed period effects (dummy variables for the different sample subperiods).

[10] All specifications are derived from equation (B.2) in Box 1 using average years of schooling (*S*) as a proxy for the stock of human capital (*H*). To indicate this, I use $\alpha_s$ (rather than $\alpha_h$) for the coefficient of schooling in the production function. The last specification (omitted from the published version of the paper for space reasons) also incorporates a technical progress function similar to equation (B.4) in the same Box, except in that the stock of human capital is omitted. Hence, the estimated model does not allow for rate effects. We have tried to incorporate them but the results are not satisfactory. This problem frequently arises in the literature. See de la Fuente and Ciccone (2003) for a discussion of the reasons why it may be difficult to separate the rate and level effects of human capital.

The estimates of the coefficient that measures the elasticity of output with respect to the level of schooling $(\alpha_s)$ obtained with the different specifications and schooling series are shown in Table 2. The last two rows of the table show average coefficient values and *t* ratios for each data set computed across the different specifications, and the last column reports the average values of $\alpha_s$ and the corresponding *t* statistic computed across data sets for each specification.

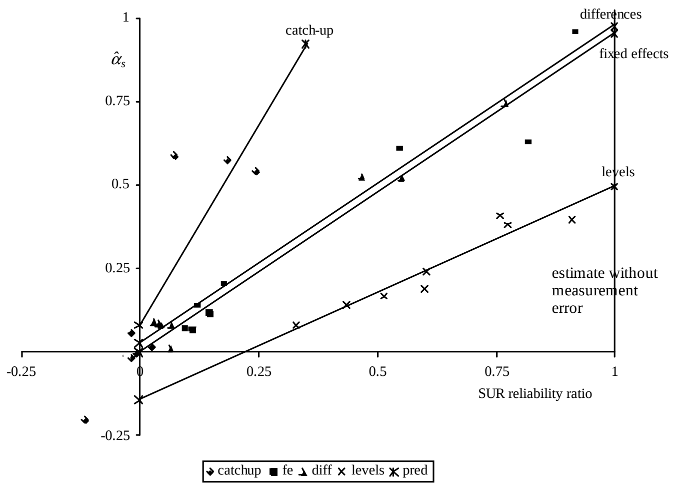
The pattern of results that emerges as we change the source of the human capital data is consistent with our hypothesis on the importance of educational data quality for growth estimates. For all data sets, the estimated value of $\alpha_s$ is positive and significant in the specification in levels without fixed country effects (first set of rows in the table), but the size and significance of the estimates increase considerably as we move to the data sets with higher reliability ratios (which correspond to the last columns of the table). The differences are even sharper when the estimation is repeated with fixed country effects (second set of rows) or with the data in growth rates with or without a catch-up effect (third and fourth blocks). The results obtained with the Kyriacou, Barro and Lee and NSD data in growth rates are consistent with those reported by Kyriacou (1991), Benhabib and Spiegel (1994) and Pritchett (2001), who find insignificant (and sometimes negative) coefficients for human capital in an aggregate production function estimated with differenced data. On the other hand, our series and those of Cohen and Soto produce rather large and precise estimates of the human capital coefficient in most equations and, in the case of our preferred catch-up specification, also yield plausible values of the remaining parameters of the model, with estimates of $\alpha_k$ close to the share of physical capital in national income and positive technological catch-up coefficients.

### *Correcting for measurement error bias*

The results summarized in Table 2 strongly suggest that measurement error induces a large downward bias in estimates of human capital coefficients. They also show that improvements in data quality reduce this bias and generate results that are generally more favorable to the view that investment in schooling contributes substantially to productivity growth. To make this point visually, Figure 1 plots the various estimates of $\alpha_s$ given in Table 2 against the corresponding SUR reliability ratios (taken from Table 1), along with the regression lines that summarize the relationship between these two variables for each of the specifications estimated in the previous section. The scatter shows a clear positive correlation between OLS estimates of schooling coefficients and reliability ratios within each specification and suggests that the true value of $\alpha_s$ is at least 0.50 (which is the prediction of the levels equation for $r=1$).

As suggested by Figure 1, it is possible to extrapolate the relationship between reliability ratios and estimated human capital coefficients that is observed across data sets in order to estimate the value of $\alpha_s$ that would be obtained in the absence of measurement error. In this manner, it is possible to construct meta-estimates of this parameter that will be free of attenuation bias, although this has to be done a bit more carefully than what is suggested by the figure when the growth equation includes additional regressors.

**Figure 1. Estimated $\alpha_s$ vs. SUR reliability ratio**




Source: de la Fuente and Doménech (2006).

In de la Fuente and Doménech (2002, 2006) we use a procedure of this type to obtain consistent meta-estimates of $\alpha_s$. Working with the three linear specifications estimated above (that is, with all of them except the catch-up model) and with different assumptions about the nature of measurement error (in particular about its correlation across data sets and with the remaining explanatory variables in the model), we obtain different estimates of $\alpha_s$ which are then adjusted to account for the possible bias generated by the fact that we are working with the average attainment of the entire population rather than that of employed workers. In this manner, we generate a rather broad range of possible values for $\alpha_s$. Under what we consider to be the most plausible assumptions, our results imply values of $\alpha_s$ between 0.70 and 0.80.

It is worth noting that our smallest lower bound for this parameter is 0.57. This is almost twice as large as Mankiw et al.'s (1992) estimate of 1/3, which could probably have been considered a consensus value for this coefficient in the early 1990s and then came to be seen as too optimistic in the light of the negative results in the literature reviewed in the last part of Section 1. Our estimates, in contrast, point to a considerably higher figure and suggest investment in human capital is an important growth factor whose effects have been underestimated in previous studies as a result of the poor quality of schooling data.

### *3.4 Regional results for Spain*

Our analysis of Spanish regional data yields qualitatively similar conclusions regarding the contribution of schooling to productivity. In de la Fuente and Doménech (2008), we estimate a catch-up specification using biennial data for the Spanish regions covering the period 1965-1995. The specification is identical to the one estimated above for the OECD sample except in that physical capital is now disaggregated into two components, one of which is the stock of productive infrastructures (transport and water supply networks and urban structures). As a proxy for the stock of human capital, we use our own census-based attainment series and an alternative estimate of average years of schooling constructed using Mas et al.'s (MPUSS, 2002) series on the breakdown of the working-age population by attainment level. This series is based on Labor Force Survey data and, presumably as a result of relatively small sample sizes in a number of regions, tends to be rather more volatile than our own series.

The estimates of the human capital parameter obtained with both schooling series are reported in Table 3. All equations contain period dummies. Equations (1) and (2) contain a full set of regional dummies, and equations (3) and (4) retain only those regional fixed effects that were significant in the first iteration. An inspection of the table reveals two interesting results regarding the coefficient of human capital $(\alpha_s)$. First, this parameter goes from being non-significant when the MPUSS (2002) data are used to having a large and significant value with our attainment series. This result is consistent with our estimates of the information content of the two series, as the relevant reliability ratio is 0.900 for our data and only 0.035 for MPUSS's attainment series when both are measured in logarithmic differences. Second, our estimate of $\alpha_s$ for the Spanish regions (0.835) is higher than those reported above for the OECD sample using a similar specification (0.540 with a full set of country dummies and 0.394 when only the significant fixed effects are retained). Once more, the explanation seems to lie at least partly in the information content of the different data sets (the relevant reliability ratio for the cross-country attainment series in de la Fuente and Doménech, 2006 was 0.246). In fact, our estimate of $\alpha_s$ using Spanish regional data lies well within the range of the meta-estimates obtained by de la Fuente and Doménech (2006) for OECD countries after correcting for measurement error.

**Table 3. Growth estimates with alternative schooling series and specifications**

| | **[1]** | **[2]** | **[3]** | **[4]** |
|---|---|---|---|---|
| **S data from:** | **MPUSS** | **D&D** | **MPUSS** | **D&D** |
| $\alpha_s$ | -0.013 | 0.835 | -0.013 | 0.835 |
| | (0.11) | (2.04) | (0.11) | (4.13) |
| adj. $R^2$ | 0.749 | 0.753 | 0.757 | 0.763 |
| regional effects | all | all | signif. | signif. |

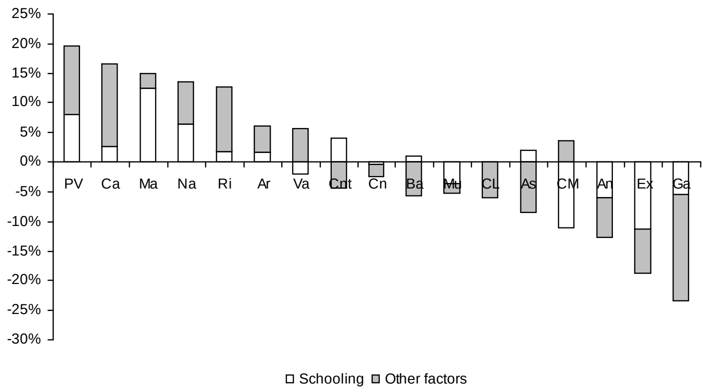
Source: de la Fuente and Doménech (2008).

Notes: All equations include period dummies. White's heteroscedasticity-consistent *t* ratios in parentheses below each coefficient. The employment ratio has been dropped from the equation due to its lack of significance.

### *3.5 Some implications*

The results summarized in the previous sections have some important implications. If average schooling enters the production function with a coefficient within the range of values we have estimated, differences in school attainment are one of the key sources of productivity differentials across both the OECD countries and the regions of Spain and investment in education yields a rather substantial return that, in most territories, compares quite favorably with that available from alternative investment opportunities.

**Figure 2. Contribution of schooling to relative productivity in 1995**




Source: de la Fuente and Doménech (2008).

Notes: Key: *An* = Andalucia; *Ar* = Aragón; *As* = Asturias; *Ba* = Baleares; *Cn* = Canarias; *Cnt* = Cantabria; *CL* = Castilla y León; *CM* = Castilla la Mancha; *Cat* = Cataluña; *Va* = Valencia; *Ex* = Extremadura; *Ga* = Galicia; *Ma* = Madrid; *Mu* = Murcia; *Na* = Navarra; *PV* = País Vasco; *Ri* = Rioja.

While I do not have the space that I would need to go into any detail, I do not want to close this section without at least a brief elaboration on these two statements. Using the estimates given in Table 3 and the underlying data, in de la Fuente and Doménech (2008) we have calculated the contribution of schooling to the relative productivity of the Spanish regions, defined as log real output per job measured in deviations from the (unweighted) sample average of the same variable. Figure 2 shows the decomposition of the relative productivity of each region into a schooling-induced component and a residual that captures the joint impact of all other factors. Using regression weights to average the different regions, we find that the share of schooling in relative productivity was 40 percent in 1995 – that is, for the typical Spanish region, schooling accounts for 4/10 of the productivity gap with the sample average.[11] A similar calculation for the

[11] We define the relative productivity of region i $(qrel_i)$ as the difference between the region's log output per employed worker and the average value of the same variable in the sample. The contribution of human capital to relative productivity $(cs_i)$ is obtained multiplying the coefficient of this factor, $\alpha_s$ by the relative level of schooling (measured in log differences with the

OECD sample implies a share of schooling in relative productivity of 30 percent.

Our estimates also imply that the social returns to education are quite respectable.[12] Combining our results on the productivity effects of human capital with rough estimates of its impact on employment and with data on educational expenditure, we estimate social rates of return ranging from 10.1 percent to 12.6 percent in Spain and from 8.3 percent to 11.5 percent in EU15.[13] In both samples, these returns compare quite favorably with those available from alternative investment opportunities in most territories and they tend to be higher than what may be expected on basis of microeconometric estimates of the private returns to schooling, suggesting that important externalities may be at work. Both findings suggest that in most of the territories we have studied, a marginal reallocation of investment resources in favor of education would be socially desirable.

## 4. Conclusion

Academic economists have traditionally been rather optimistic about the contribution of education to economic development and have often assigned a central role in formal models to the accumulation of human capital, particularly in the recent literature on endogenous growth. The results of early empirical studies on the determinants of economic growth have largely been consistent with this view. During the second half of the nineties, however, a new round of empirical papers produced rather disappointing results on the subject that sparked a lively controversy in the literature between "skeptics" and "believers" in the salutary effects of schooling on aggregate productivity growth.

---

geometric sample mean). After constructing these two variables for each region, we estimate a regression of the form $cs_i = a * qrel_i + e_i$, where $\varepsilon_i$ is a random disturbance. The coefficient obtained in this manner, $a \cong cs_i / qrel_i$, measures the fraction of the observed productivity differential that can be attributed to human capital in the sample as a whole.

[12] The social rate of return to schooling is defined as the discount rate that equates the present value of the increases in output induced by a marginal increase in average attainment to the present value of the explicit and opportunity costs of schooling. For further details on how this magnitude can be estimated, see de la Fuente (2003).

[13] For additional details, see de la Fuente and Doménech (2008) and de la Fuente (2003).

This paper contains a selective and rather opinionated review of some of the relevant literature. After setting the stage, it focuses on a problem (the poor quality of cross-country schooling data) that may help explain the discouraging results found in some influential studies, on possible ways of overcoming this problem, and on what happens when this is done. I have argued that due to various deficiencies of the primary data, the schooling series used in the early empirical literature on growth and human capital contain a considerable amount of noise that generates a very substantial downward bias in estimates of the parameter that measures the contribution of educational attainment to productivity. This conclusion is based on the estimation of a statistical indicator of the information content of the schooling series most commonly used in the literature. It is also reinforced by the finding of a clear tendency for human capital coefficients to rise and become more precise as the information content of the schooling data increases. When this relationship is extrapolated to construct estimates of the value of the schooling coefficient that would be obtained in the absence of measurement error, the exercise suggests that the true value of the elasticity of output with respect to the stock of human capital is almost certainly no lower than 0.60 – that is, around twice as high as the most optimistic estimate of reference in the earlier literature on growth and human capital. If this conclusion is correct, investment in the quantity and quality of schooling emerges as one of the most powerful policy levers available to governments in order to influence productivity growth and, ultimately, the standard of living of their citizens.

## References

Azariadis, C. and Drazen, A. (1990), Threshold externalities in economic development, Quarterly Journal of Economics 105, 501-526.

Barro, R. (1991), Economic growth in a cross section of countries, Quarterly Journal of Economics 106, 407-443.

Barro, R. (1997), Determinants of Economic Growth: A Cross-country Empirical Study, Lionel Robbins Lectures, MIT Press, Cambridge MA.

Barro, R. and Lee, J-W. (1993), International comparisons of educational attainment, Journal of Monetary Economics 32, 363-394.

Barro, R. and Lee, J-W. (1996), International measures of schooling years and schooling quality, American Economic Review 86, Papers and Proceedings, 218-223.

Barro, R. and Lee, J-W. (2000), International data on educational attainment, updates and implications, NBER Working Paper 7911.

Barro, R. and Lee, J-W. (2010), A new data set of educational attainment in the world, 1950-2010, NBER Working Paper 15902.

Baumol, W., Batey Blackman, S.A. and Wolf, E. (1989), Productivity and American Leadership: The Long View, MIT Press, Cambridge MA.

Behrman, J. and Rosenzweig, M. (1994), Caveat emptor: Cross-country data on education and the labor force, Journal of Development Economics 44, 147-171.

Benhabib, J. and Spiegel, M. (1994), The role of human capital in economic development: Evidence from aggregate cross-country data, Journal of Monetary Economics 34, 143-173.

Caselli, F., Esquivel, G. and Lefort, F. (1996), Reopening the convergence debate: A new look at cross-country growth empirics, Journal of Economic Growth 1, 363-389.

Cohen, D. and Soto, M. (2007), Growth and human capital: Good data, good results, Journal of Economic Growth 12, 51-76.

Coulombe, S., Tremblay, J.F. and Marchand, S. (2004), Literacy scores, human capital and growth across fourteen OECD countries, Statistics Canada, Ottawa.

de la Fuente, A. (2003), Human capital in a global and knowledge-based economy, part II: Assessment at the EU country level, European Commission, DG for Employment and Social Affairs, Brussels, http://digital.csic.es/bitstream/10261/1844/1/57603.pdf.

de la Fuente, A. and Ciccone, A. (2003), Human capital and growth in a global and knowledge-based economy, European Commission, DG for Employment and Social Affairs, Office for official publications of the European Communities, Luxembourg. http://ideas.repec.org/p/aub/autbar/576.03.html.

de la Fuente, A. and Doménech, R. (2000), Human capital in growth regressions: How much difference does data quality make?, OECD Economics Department Working Paper 262, Paris.

de la Fuente, A. and Doménech, R. (2001a), Schooling data, technological diffusion and the neoclassical model, American Economic Review 91, Papers and Proceedings, 323-327.

de la Fuente, A. and Doménech, R. (2001b), Educational attainment in the OECD, 1960-90, CEPR Discussion Paper 3390.

de la Fuente, A. and Doménech, R. (2002), Human capital in growth regressions: How much difference does data quality make? An update and further results, CEPR Discussion Paper 3587.

de la Fuente, A. and Doménech, R. (2006), Human capital in growth regressions: How much difference does data quality make?, Journal of the European Economic Association 4, 1-36.

de la Fuente, A. and Doménech, R. (2008), Human capital, growth and inequality in the Spanish regions, in U. Stierle et al. (red.), Regional Economic Policy in Europe. Challenges for Theory, Empirics and Normative Interventions, Edward Elgar, Cheltenham.

Hanushek, E. and Kimko, D. (2000), Schooling, labor-force quality and the growth of nations, American Economic Review 90, 1184-1208.

Hanushek, E. and Woessmann, L. (2008), The role of cognitive skills in economic development, Journal of Economic Literature 46, 607-668.

Hanushek, E. and Woessman, L. (2009), Do better schools lead to more growth? Cognitive skills, economic outcomes and causation, NBER Working Paper 14633.

Islam, N. (1995), Growth empirics: A panel data approach, Quarterly Journal of Economics 110, 1127-1170.

Krueger, A. and Lindahl, M. (2001), Education for growth: Why and for whom?, Journal of Economic Literature XXXIX, 1101-1136.

Kyriacou, G. (1991), Level and growth effects of human capital, A cross-country study of the convergence hypothesis, mimeo, New York University.

Landau, D. (1983), Government expenditure and economic growth: A cross-country study, Southern Economic Journal 49, 783-792.

Lau, L., Bhalla, S. and Louat, F. (1991), Human and physical capital stock in developing countries: Construction of data and trends, draft, World Development Report, World Bank.

Lau, L, Jamison, D. and Louat, F. (1991), Education and productivity in developing countries: An aggregate production function approach, Report WPS 612, World Bank, Washington.

Lucas, R. (1988), On the mechanics of economic development, Journal of Monetary Economics 22, 3-42.

Mankiw, G., Romer, D. and Weil, D. (1992), A contribution to the empirics of economic growth, Quartely Journal of Economics 107, 407-437.

Mas, M., Pérez, F. Uriel, E. Serrano, L. och Soler, A. (2002), Capital humano, series 1964-2001, Capital humano y actividad económica (CD Rom), Fundación Bancaja, Valencia.

Nehru, V., Swanson, E. and Dubey, A. (1995), A new database on human capital stocks in developing and industrial countries: Sources, methodology and results, Journal of Development Economics 46, 379-401.

Pritchett, L. (2001), Where has all the education gone?, World Bank Economic Review 15, 367-391.

# Comment on de la Fuente: Human capital and productivity

Pekka Ilmakunnas[*]

De la Fuente's paper provides a useful synthesis of the results on education and aggregate productivity from the empirical growth literature, emphasizing measurement issues and data quality. My comments deal with education and productivity at a more disaggregated level. I mention some issues that have been examined at the micro (individual, firm/plant) level that may have implications on the macro (country, region) level.

At the individual level, productivity is seldom directly observed but there are outcomes, like earnings, that are correlated with productivity. On the other hand, data on education (education years or degrees attained) are usually relatively reliable, coming from registers in many countries. Therefore, the measurement issues have not been central, as compared to the endogeneity of education. However, it is increasingly acknowledged that the data used do not really measure what happens in education. We do not really know what kinds of skills are generated and how, and which skills are relevant for productivity. Research has started to investigate the 'black box' of education in more detail. The use of student achievement test data is an example of this development, and it has also been applied at the macro level, as mentioned by de la Fuente.

At the firm or plant level, the connection between education and productivity can be studied using linked employer-employee data. In cases where information on the educational level of employees is based on the registered degrees of employees, the data are reliable, so that the problems

[*] Aalto University School of Economics, pekka.ilmakunnas@aalto.fi.

of measuring education have not been discussed much in the firm-level literature. There has been more discussion on measuring average unobservable skills by using individual effects from a wage equation estimated with both firm and individual effects (e.g. Abowd et al., 2005).

It is not uncommon to find that education (average education years of employees or the share of highly educated) is not significant or even negatively related to firm productivity (e.g. Ilmakunnas and Maliranta, 2005). Since educational data are more reliable than in the macro studies, this negative result is not likely to arise because of measurement problems. Rather, it seems to be related to lags in the effects of education. When firms hire new, highly educated workers, they may at the same time reorganize work and production processes, which leads to a disruption of productivity in the short run, but to positive productivity effects in the longer run.

Besides the level of education, the field of education may also be of importance. For example, Ilmakunnas and Maliranta (2005) found that highly educated workers with a business education have a higher productivity than those with a technical education (i.e. the coefficient of the share of workers with a business education is higher than that of the share of technically educated). A plausible explanation is that the latter are involved in developing products and processes and the former with their commercialization. The productivity effects do not show up until the new products are in the market.

In addition to the average educational level, also educational or skill diversity of the workforce is of importance. For example, there may be complementarities of workers of different skills. The evidence so far is mixed and partly depends on whether skills are measured by education or in some other way (see e.g. Ilmakunnas and Ilmakunnas, 2011).

Finally, besides the human capital stock, the flows of workers of different educational levels can also be used in measuring the productivity effects of education. When firms want to adjust the educational level of their workforce upwards, this happens through the inflow of highly educated and/or the outflow of less educated workers. This opens up the possibility of measuring the productivity of education separately for hired, staying and exiting workers (e.g. Maliranta et al., 2009). At the macro level, the flows would be cohorts entering into and exiting from the working age population, as well as immigration flows.

Whether these features of firm-level research that I have described can be applied at the macro level is an open issue and it is clearly dependent on the availability of data. At least at the regional level within some countries it is possible to calculate measures of the distribution of fields of education and the dispersion of education years. Even the construction of indices for average unobservable skills or flows of worker by educational level for regions could be possible in some countries. However, cross-country analysis is, of course, much more challenging.

At the firm/plant level, human capital (education and experience) only explains a relatively small part of the productivity dispersion across plants (Syverson, 2011). At the macro level, the role of education should be more important since it also picks up the external effects of education. Therefore, the gains from developing more accurate ways of measuring the aggregate productivity impacts of education are notable.

## References

Abowd, J.M., Haltiwanger, J., Jarmin, R., Lane, J., Lengermann, P., McCue, K., McKinney, K. and Sandusky, K. (2005), The relation among human capital, productivity, and market value – building up from micro evidence, in C. Corrado et al. (eds.), Measuring Capital in the New Economy, University of Chicago Press, Chicago.

Ilmakunnas, P. and Ilmakunnas, S. (2011), Diversity at the workplace: Whom does it benefit?, De Economist 159, Special Issue on Ageing Workforces, 223-255.

Ilmakunnas, P. and Maliranta, M. (2005), Technology, labour characteristics and wage-productivity gaps, Oxford Bulletin of Economics and Statistics 67, 623-645.

Maliranta, M., Mohnen, P. and Rouvinen, P. (2009), Is inter-firm labor mobility a channel of knowledge spillovers? Evidence from a linked employer-employee panel, Industrial and Corporate Change 18, 1161-1191.

Syverson, C. (2011), What determines productivity?, Journal of Economic Literature 49, 326-365.

# Productivity and international firm activities: What do we know?*

Joachim Wagner**

*Summary*

This paper summarizes in a non-technical way what we know from empirical studies based on firm-level data about the mutual links between international activities of firms and productivity. It is written with a view to inform policy makers in an evidence-based way. A special focus is on the empirical evidence we have from studies using firm-level data from the Nordic countries. It is argued that this empirical evidence does not provide a sound basis for searching for similarities and differences (and their causes) between the Nordic countries, and the empirical results reported cannot qualify as stylized facts that can inform policy makers in an evidence-based way.

*Keywords:* International firm activities, productivity, firm-level data, Nordic countries.
*JEL classification numbers:* F14, F23, D22, L25.

---

* Many thanks to Martin Andersson, Ragnhild Balsvik, Roger Bandik, Joakim Gullstrand, Pekka Ilmakunnas, Jakob Madsen and my discussant Markku Stenborg for helpful comments on an earlier draft version. Evidently, all errors and omissions are my fault.

** Leuphana University Lueneburg and IZA, Bonn, wagner@leuphana.de.

Productivity – the efficiency with which firms turn inputs (labour, physical capital, energy, materials, managerial know-how) into outputs (goods, services) – is important for the competitiveness of firms, regions and countries on local, national and international markets. Productivity is an important driver of growth and welfare. Therefore, the study of productivity has been a core topic in economics for a long time.

Empirical studies that use firm-level micro data to investigate the determinants and consequences of productivity differentials between firms, however, are of a more recent vintage. A case in point is the literature dealing with the links between productivity and international firm activities. This literature started with a Brookings paper by Bernard and Jensen (1995) that documents a positive exporter productivity premium in US manufacturing industries – exporting firms are more productive than non-exporting firms of the same size from the same narrowly defined industry. During the past 15 years, economists all over the world have used firm-level micro data to investigate productivity differences between exporting and non-exporting firms and the direction of causality between export activity and firm-level productivity (see Wagner, 2007, for a survey). More recently, researchers interested in the links between international firm activities and productivity have started to look beyond exports and investigate other forms of international firm activities (imports, foreign direct investment, offshoring), to look beyond manufacturing and investigate services and the role of countries of origin and the destination of imports and exports.[1]

This literature on the *micro-econometrics of international firm activities* inspired theorists to develop what is now labelled the *new new trade theory* where heterogeneous firms that differ in productivity are at the heart of the theoretical models.[2]

All this resulted in a mushrooming literature.[3] This paper summarizes in a non-technical way what we know from this literature about the mutual

---

[1] Furthermore, other dimensions of firm performance besides productivity – like growth, profitability, and wages paid – were also investigated. This literature, however, is beyond the scope of this paper that has a focus on productivity. For international firm activities and growth, see Wagner (2002), for profitability Fryges and Wagner (2010), and for wages Schank et al. (2007, 2010).

[2] The canonical model from this literature is Melitz (2003); Redding (2010) is a survey.

[3] For partial surveys, see López (2005), Helpman (2006), Greenaway and Kneller (2007), Wagner (2007) and Hayakawa et al. (2011).

links between international activities of firms and productivity with a view to inform policy makers in an evidence-based way. Given the focus on the Nordic countries of the conference for which this paper was prepared, Table 1 summarizes the empirical evidence we have from studies using firm-level data from the Nordic countries.[4] These studies will be discussed in turn in the appropriate sections of the paper.

The rest of the paper is organized as follows. Section 1 looks at exports and imports. Section 2 deals with outward and inward foreign direct investment. Section 3 looks at offshoring of activities to foreign countries. Section 4 summarizes what we know and what we do not know about the mutual links between international firm activities and productivity in Nordic countries and elsewhere and discusses the policy implications.

[4] I am grateful for hints to papers on international firm activities and productivity that are based on firm-level data from Nordic countries and that are not listed in Table 1. Given the topic of the conference, the focus is on productivity and international firm activities. Other important dimensions of firm performance that are closely linked to international firm activities include survival, employment, profitability, innovation, and wages. Empirical studies using firm-level data to investigate these links abound, and some of the papers covered in this survey at least touch upon other issues besides productivity. Studies for Nordic countries from this literature that do not discuss productivity but other dimensions of firm performance include the following: Survival (Bandick (2010) on foreign ownership and plant survival in Sweden; Bandick and Görg (2010) for foreign acquisition in Sweden; Greenaway et al. (2008a, 2009) for international trade and foreign ownership with relation to firm exit in Sweden); employment (Bandick and Görg (2010) for foreign acquisition in Sweden; Deschryvere and Kotiranta (2008) for offshoring of Finnish firms; Ekholm and Hakkala (2008) for offshoring of Swedish firms; Eliasson et al. (2010) on jobs and exposure to international trade within the service sector in Sweden; Hummels et al. (2010) for outsourcing in Denmark; Huttunen (2007) for foreign acquisitions in Finnish establishments; Ibsen et al. (2009) on the relation between export and import decision and employment growth in Denmark; Lehto and Böckermann (2008) on the effects of cross-border mergers and acquisitions on employment in Finland; Lööf (2009) on trade and employment growth in firms from manufacturing and services in Sweden; Munch (2010) on international outsourcing and job loss in Denmark; Pesola (2009) on labour market transitions following foreign acquisitions in Finland); innovation (Andersson and Lööf (2009b) for trade and innovative activities in small and medium-sized enterprises in Sweden; Dachs et al. (2008) for foreign-owned vs. domestic enterprises from Denmark, Finland, Norway and Sweden; Bandick et al. (2010) for foreign acquisition and R&D activities in Sweden); Laursen (2008) for innovation and export performance in Danish manufacturing and services firms; wages (Bandick (2009) on foreign acquisition in Sweden; Fosse and Maitra (2010) for offshoring to China from Danish firms; Geishecker et al. (2010) for offshoring by Danish firms; Heyman et al. (2007) on foreign ownership wage premium in Sweden; Hummels et al. (2010) for outsourcing in Denmark; Huttunen (2007) for foreign acquisitions in Finnish establishments; Lundin and Yun (2009) on international trade and wages in Sweden; Munch and Skaksen (2008) on exports and wages in Denmark; Pesola (2007) on foreign ownership in Finland).

**Table 1. Synopsis of empirical studies on productivity and international firm activities using data from Nordic countries**

| Author(s)<br>Year of publication | Data used | Type of international firm activity investigated | Important findings |
|---|---|---|---|
| ***Denmark*** | | | |
| Dachs et al. (2008) | Third European Community Innovation Survey (CIS 3), 1998-2000, 844 firms | Foreign ownership | Foreign-owned enterprises exhibit a significantly higher labour productivity than do domestically owned enterprises. |
| International Study Group on Exports and Productivity (2008) | All firms from manufacturing with at least 20 employees, 1999-2002. 29 161 observations | Exports | Exporters are substantially more productive than non- exporters. |
| Eriksson et al (2009). | All firms from manufacturing with at least 10 employees, 1993-2003. 95 365 firms in 1993,125 684 firms in 2003. | Exports, imports | Exports, imports of intermediate goods, and imports of final goods positively associated with higher productivity. |
| Smeets and Warzynski (2010) | Product-firm-level dataset for manufacturing firms, 1998-2005 | Exports, imports | Importing and exporting positively associated with productivity, firms that import and export are most productive. Evidence for self-selection into importing and exporting but no learning from trade effects. |
| ***Finland*** | | | |
| Ilmakunnas and Maliranta (2004) | Linked employer-employee data for 1994 to 2001; 2 200 to 2 500 plants annually. | Foreign ownership | Foreign ownership increases total factor productivity by between 9 and 11 per cent. |
| Ylä-Anttila et al. (2004) | 500 largest companies, 1986-1998 | Foreign ownership | Labour productivity has increased more rapidly companies with high foreign ownership. |
| Dachs et al. (2008) | Third European Community Innovation Survey (CIS 3), 1998-2000, 818 firms | Foreign ownership | Foreign-owned enterprises exhibit a significantly higher labour productivity than do domestically owned enterprises. |

**Table 1. Continued….**

| Author(s)<br>Year of publication | Data used | Type of international firm activity investigated | Important findings |
|---|---|---|---|
| Ilmakunnas and Nurmi (2010) | Linked employer-employee data for 1980-2005 | Exports | Highly productive plants are likely to enter the export market earlier and survive longer in the export market. The upper and lower tails of the productivity distribution are represented by plants that start exporting and those that are exiting. |
| Ilmakunnas et al. (undated) | Linked employer-employee panel data | Foreign ownership | Hiring workers from foreign multinationals has a positive effect on productivity in local domestic firms, indicating knowledge spillovers from multinationals to domestic firms. |
| ***Norway*** | | | |
| Dachs, et al (2008) | Third European Community Innovation Survey (CIS 3), 1998-2000; 2 286 firms | Foreign ownership | Foreign-owned enterprises exhibit a significantly higher labour productivity than do domestically owned enterprises. |
| Balsvik and Haller (2010) | Annual survey of all Norwegian manufacturing plants with at least 10 employees; 1992-2004. 65 740 observations | Acquisition of home firms by foreign firms | Prospective foreign owners pick high-productivity plants. Labour productivity increases after foreign acquisitions. |
| Irarrazabal et al. (2010) | Matched employer-employee data for all non-oil joint-stock firms 1996-2005; ca. 8 000 firms per year | Exports | Exporters are more productive; 15-40 per cent of the exporter premium reflects differences in workforce rather than true efficiency. |
| Castellacci (2010) | 814 firms in the service sector. Information for 2004-2006, but data are cross-section data only | Exports | Labour productivity positively related to export participation. |
| Balsvik and Haller (2011) | Panel data from the Norwegian Manufacturing Census | Foreign direct investment in Norway | Greenfield entry both in the same industry and in the same labour market region has a negative impact on the productivity of domestic plants; entry via acquisition affects productivity of domestic plants in the same industry positively. |

**Table 1. Continued….**

| Author(s)<br>Year of publication | Data used | Type of international firm activity investigated | Important findings |
|---|---|---|---|
| ***Sweden*** | | | |
| Hansson and Lundin (2004) | All manufacturing firms with at least 50 employees, 1990-1999. 3 275 unique firms, between 1 565 and 1 820 firms each year | Exports | Exporting firms have a significantly higher productivity than non-exporting firms. More productive firms self-select into the export market. Exporting further increases firm productivity. |
| Karpaty and Lundberg (2004) | All manufacturing firms with at least 50 employees, 1990-2000 | Inward FDI | The presence of foreign ownership in the same industry and region seems to enhance the total factor productivity of domestic firms. |
| Greenaway et al. (2005) | Manufacturing firms 1980-1997; at least 50 employees; 36 903 observations | Exports | Productivity growth of exporters upon entry does not appear to differ significantly from that of non-exporters either in the periods leading up to or after entry. |
| Dachs et al. Lööf (2008) | Third European Community Innovation Survey (CIS 3), 1998-2000; 1 197 firms | Foreign ownership | Foreign-owned enterprises exhibit a significantly higher labour productivity than do domestically owned enterprises. |
| Greenaway et al. (2008b) | Unbalanced panel of 1 570 firms in the food and beverage sector, 1997-2002 | Exports | Exporters are more productive than non-exporters; more productive firms tend to reach out to more markets. |
| Karpaty (2006) | Sample from FIEF longitudinal firm-level data base for Swedish manufacturing firms, 1993-2002. All firms with at least 50 employees. 23 848 observations | Foreign ownership | Foreign-owned firms have a higher total factor productivity than domestic firms; no productivity differential between Swedish and foreign multinational enterprises. Foreign greenfields in Sweden are more productive compared to both acquired firms and Swedish multinational enterprises. No evidence of reverse causality. |
| Karpaty (2007) | All manufacturing firms with at least 50 employees, 1986-2002 | Acquisition of home firms by foreign firms | Positive effect of foreign acquisition on productivity starting between 1-5 years post acquisition. |
| Andersson et al. (2008) | Detailed matched data from various sources for firms from manufacturing industries; unbalanced panel 1997-2004 | Exports, Imports | Export and import productivity premiums are significant and of similar magnitudes, and the premium is the highest for firms that both export and import. Premiums in both the number of markets and the number increase of products traded. |

**Table 1. Continued….**

| Author(s)<br>Year of publication | Data used | Type of international firm activity investigated | Important findings |
|---|---|---|---|
| International Study Group on Exports and Productivity (2008) | Manufacturing firms with at least 20 employees, 1997-2004 31 838 observations | Exports | Exporters are more productive than non-exporters (but not if firm-fixed effects are controlled for); no evidence for self-selection of more productive firms into exports or for learning-by-exporting. |
| Andersson and Lööf (2009a) | Manufacturing firms with at least 10 employees, 1999-2004, 38 929 observations | Exports | Learning-by-exporting effect on productivity found among persistent exporters with a high export intensity, but not among temporary exporters or persistent exporters with low export intensity. |
| Bandick (2009) | All manufacturing firms with 20 or more employees, 1993-2002; ca. 3 400 firms per year | Foreign acquisition | Targeted Swedish firms have better productivity growth after vertical foreign acquisition only; no such impact is found from horizontal foreign acquisition |
| Eliasson et al. (2009) | Manufacturing firms with less than 50 employees, 1997-2006. Some 200 000 observations | Exports | Evidence of self-selection of more productive firms into exports. Learning-to-export occurs but learning-by- exporting does not. |
| Gullstrand (2009) | Unbalanced panel of all firms in the food and beverage sector, 1997-2002. 1 570 firms, 9 858 firm-destination-year observ. | Exports | More productive firms are more likely to export. |
| Lööf (2009) | Unbalanced panel of firms from manufacturing and services, 1997-2006; ca. 260 000 observations | Exports | Self-selection of more productive firms into exporting in services and manufacturing; exporter premium larger in services. No evidence found that exporting increases productivity growth. |
| Lööf and Andersson (2010) | All manufacturing firms with at least 10 employees, 1997-2004. Ca. 57 000 observations | Imports | Imports lead to an increase in productivity; the effect is larger for imports from highly developed countries. |

Source: See References.

Note: For each country the studies are listed in chronological order by year of publication and in alphabetical order for each year.

# 2. International trade and productivity

## *2.1 Exports and productivity*

Discussions of the role of exports in promoting growth in general, and productivity in particular, have been going on for many years.[5] Earlier empirical studies in this field used data at the country or industry level to test whether exports promote productivity growth or vice versa. In 1995, Bernard and Jensen published the first of a series of papers that changed this research perspective (see Bernard and Jensen, 1995). They used large comprehensive longitudinal data from surveys performed regularly by official statistics in the US to look at differences between exporters and non-exporters in various dimensions of firm performance, including productivity. These papers started a literature. During the years following the publication of Bernard and Jensen's *Brookings* paper, researchers all over the world discovered the rich data sets collected by their statistical offices as a source for investigating the export activity of firms, and its causes and consequences. The extent and cause of productivity differentials between exporters and their counterparts which sell on the domestic market only is one of the core topics in this literature.

There are two alternative but not mutually exclusive hypotheses why exporters can be expected to be more productive than non-exporting firms (see Bernard and Jensen, 1999 and Bernard and Wagner, 1997). The first hypothesis points to self-selection of the more productive firms into export markets. The reason for this is that there exist additional costs of selling goods in foreign countries. The range of additional costs include transportation costs, distribution or marketing costs, personnel with the skill to manage foreign networks, or production costs in modifying current domestic products for foreign consumption. These costs provide an entry barrier that less successful firms cannot overcome. Furthermore, the behaviour of firms might be forward-looking in the sense that the desire to export tomorrow leads a firm to improve performance today to also be competitive on the foreign market. Therefore, cross-section differences between ex-

[5] This section is based on Wagner (2007) where technical details are discussed and detailed references to the literature can be found.

porters and non-exporters may, in part, be explained by ex ante differences between firms: The more productive firms become exporters.

The second hypothesis points to the role of learning-by-exporting. Knowledge flows from international buyers and competitors help improve the post-entry performance of export starters. Furthermore, firms participating in international markets are exposed to more intense competition and must improve faster than firms who sell their products domestically only. Exporting makes firms more productive.

A common approach to investigate differences in productivity between exporters and non-exporters is to follow (sometimes only in part, and sometimes with modifications and extensions) the methodology introduced by Bernard and Jensen (1995, 1999). Studies of this type use longitudinal data for plants (usually from the regular surveys conducted by official statistics) to document differences in levels and growth rates of productivity between exporters and non-exporters in a first step. Here, one starts by looking at differences in average labour productivity (total value of shipments per worker, or value added per worker) or average total factor productivity between exporters and non-exporters. The result is an unconditional productivity differential.

The next step is the computation of so-called exporter premia, defined as the ceteris paribus percentage difference of labour productivity between exporters and non-exporters. These premia are computed from a regression of log productivity on the current export status dummy and a set of control variables (usually including industry, region and firm size measured by the number of employees and year). The export premium shows the average percentage difference between exporters and non-exporters.

To shed some light on the empirical validity of the first hypothesis mentioned – namely, that the more productive firms go abroad – the pre-entry differences in productivity between export starters and non-exporters are investigated next. If good firms become exporters, we should expect to find significant differences in performance measures between future export starters and future non-starters several years before some of them begin to export. To test the second hypothesis mentioned – namely, that exporting fosters productivity – the post-entry differences in productivity growth between export starters and non-exporters are investigated.

Wagner (2007) gives a synopsis of findings from 54 empirical studies using firm-level data from 34 countries and investigating the relationship of exports and productivity. Among the countries covered are highly industrialised countries (e.g., US, UK, Canada, Germany); countries from Latin America (Chile, Colombia, Mexico); Asian countries (China, Korea, Indonesia, Taiwan); transition countries (Estonia, Slovenia); and least developed countries from sub-Saharan Africa. Given this wide range of countries, the big picture is amazingly clear-cut: With only a few exceptions, exporters are found to have higher productivity, and often higher productivity growth, and this also tends to hold after controlling for observed plant characteristics (like industry and size). Exporters are better.

The findings for pre-entry differences often present evidence in favour of the self-selection hypothesis: Future export starters tend to be more productive than future non-exporters years before they enter the export market, and they often have higher ex-ante growth rates of productivity. The good firms go abroad.

Evidence regarding the learning-by-exporting hypothesis is somewhat more mixed: The results for post-entry differences in performance between export starters and non-exporters point to faster productivity growth for the former group in some studies only. Exporting does not necessarily improve firms.[6]

Does the big picture sketched here – exporters are more productive than non-exporters, and the more productive firms self-select into export markets, while exporting does not necessarily improve productivity – also describe the situation in the Nordic countries? From the empirical studies summarized in Table 1, we see that exporters tend to be more productive than firms that serve the national market only in all four countries. Some evidence for self-selection of more productive firms into exporting has been found for Denmark, Finland and Sweden (but not in every study testing for it), while this hypothesis is not tested in the studies using data for Norway. The learning-by-exporting hypothesis has only been investigated with Danish and Swedish firm-level data – with mixed results. Apparently, in the case of the Nordic countries, the jury is still out on the direction of causality between exporting and productivity.

[6] Note, however, that De Loecker (2010) recently showed that current methods that are used to test for learning by exporting are biased towards rejecting the hypothesis of positive effects of exports on productivity. He provides evidence for this in the case of Slovenia. Comparable empirical results for other countries are, to the best of my knowledge, not available.

### *2.2 Imports and productivity*

While the causes and consequences of export and its mutual relationships with productivity are prominent topics in the recent literature on internationally active firms, imports are seldom dealt with. A case in point is the recently published Bruegel study on the internationalisation of European firms (Mayer and Ottaviano, 2007) where imports are not dealt with at all. As recently expressed by Bernard et al. (2007, p. 123), "(t)he empirical literature on firms in international trade has been concerned almost exclusively with exporting, largely due to limitations in datasets … . As a result, the new theories of heterogeneous firms and trade were developed to explain facts about firm export behavior and yield few predictions (if any) for firm import behavior."

This situation is rapidly changing, however.[7] With new datasets that include information on imports at the firm level becoming available for more and more countries, a new literature has been emerging since 2005 that has a focus on the links between productivity and imports. A number of recently published empirical studies based on data from a wide range of countries document the shares of firms that are exporters, importers, and two-way traders (that both export and import), or that sell or buy on the national market only, and they look at differences between these four types of firms. Differences in productivity and their relationship with different degrees of involvement in international trade are at the centre of these studies.

Details aside, the big picture that emerges from this literature can be sketched as follows: There is a positive link between importing and productivity at the firm level, documented by a significant productivity differential between firms that import and firms that do not trade internationally; the same holds for exporting. Two-way traders are more productive than firms that only import, or only export, or do not trade at all. Often, two-way traders are the most productive group of firms, followed by importers and then exporters, while firms selling or buying on the national market come last.

How can this empirical regularity of a positive relationship between importing and productivity at the firm level be explained theoretically? In

[7] For a comprehensive discussion of the literature on imports and productivity and technical details, see Vogel and Wagner (2009).

the literature, arguments for both a positive impact of productivity on importing (which is in accordance with the self-selection of more productive firms into import markets) and for a positive impact of importing on productivity ('learning-by-importing') are discussed.

To start with the arguments in favour of self-selection of more productive firms into importing, it is pointed out that the use of foreign intermediates increases a firm's productivity but, due to fixed costs of importing, only inherently highly productive firms import intermediates. Importing is associated with fixed costs that are sunk costs, because the import agreement is preceded by a search process for potential foreign suppliers, inspection of goods, negotiation, contract formulation etc. Furthermore, there are sunk costs of importing due to the learning and acquisition of customs procedures.

As regards learning-by-importing, it is stated that there are strong arguments in favour of a causal effect of imports on productivity, because by importing, a firm can exploit global specialization and use inputs from the forefront of knowledge and technology. Proponents of this view point to the literature on international technology diffusion that advances imports as an important vehicle for knowledge and technology transfer. Furthermore, importing intermediate products allows a firm to focus resources and specialize on activities where it has particular strengths. Importers may improve productivity by using higher quality foreign inputs or by extracting technology embodied in imported intermediates and capital goods. Furthermore, a variety effect is mentioned (in which the broader range of available intermediates contributes to production efficiency) and a quality effect caused by imported intermediates that might be of better quality than local ones. If importing increases productivity, this might lead firms to self-select into export markets and help improve their success in these markets, which might contribute to an explanation of the empirical regularity that two-way traders are the most productive firms on average.

Therefore, from a theoretical point of view, the direction of causality between productivity and importing can run from one of the two sides or from both sides simultaneously. Only some of the studies mentioned above tackle this issue (or at least part of it) empirically. The bottom line, then, is that we have convincing empirical evidence of a positive relationship between importing and productivity at the firm level for a large and growing number of developed and developing countries, while research on the direction of causality between productivity and import status is still in its infancy.

Vogel and Wagner (2009) use a newly available comprehensive panel data set for manufacturing enterprises from 2001 to 2005 to document the first empirical results on the relationship between imports and productivity for Germany, a leading actor on the world market for goods. Furthermore, for the first time, the direction of causality in this relationship is systematically investigated by testing for self-selection of more productive firms into importing, and for productivity-enhancing effects of imports ('learning-by-importing'). They find a positive link between importing and productivity. From an empirical model with fixed enterprise effects that controls for firm size, industry, and unobservable firm heterogeneity, they report that the premia for trading internationally are about the same in West and East Germany. Compared to firms that do not trade at all, two-way traders do have the highest premium, followed by firms that only export, while firms that only import have the smallest estimated premium. They find evidence for a positive impact of productivity on importing, pointing to self-selection of more productive enterprises into imports, but no evidence for positive effects of importing on productivity due to learning-by-importing.

Empirical evidence on the links between imports and productivity based on econometric studies using firm-level data from the Nordic countries is scarce. From Table 1, we see that imports and productivity are positively associated in Denmark and Sweden. For Denmark, we have evidence for self-selection into importing of more productive firms but not for learning from importing. Furthermore, for Sweden we have evidence that imports lead to an increase in productivity, and that this effect is larger for imports from highly developed countries. However, this evidence is based on only one study per country and topic. Studies with data from Finland and Norway are missing. The bottom line is therefore that we have no sound empirical evidence on the links between imports and productivity in the Nordic countries.

# 3. Foreign direct investment and productivity

## 3.1 Outward foreign direct investment

Besides international trade (exports and imports), other forms of international activities of firms and their relation to productivity are investigated both theoretically and empirically. A case in point is the multi-country, multi-sector general equilibrium model of Helpman et al. (2004) that explains the decision of heterogeneous firms to serve foreign markets either through exports or through foreign direct investment, i.e. by building new production facilities in a foreign county or by acquiring existing firms in that country. They show that, in equilibrium, only the more productive firms choose to serve the foreign markets, and the most productive among this group will further choose to serve these markets via foreign direct investment. The intuition behind this theoretical result is similar to the argument put forward in the case of exports and productivity. There exist additional costs of starting production activities in a foreign country, including costs for becoming familiar with all legal and economic aspects related to doing business abroad, and these costs can be expected to be even larger than the additional costs a firm that exports has to pay compared to a firm that sells its products on the national market only. Only the most productive firms can be expected to be able to pay these costs and to produce profitably in a foreign country.

Several empirical papers take the Helpman-Melitz-Yeaple model as a point of departure.[8] Wagner (2006) uses German establishment level data and a non-parametric test for first-order stochastic dominance to show that, in line with this hypothesis, foreign direct investors are indeed more productive (not only at the mean but over the whole range of the productivity distribution) than exporters which, in turn, are more productive (again, not only at the mean but over the whole range of the productivity distribution) than firms that sell their products on the national market only. Empirical evidence for other countries (including Japan and the UK) points in the same direction (see Wagner, 2006).

However, according to the summary of empirical studies on international firm activities and firm performance reported in Table 1, there is as

[8] See Wagner (2006) for a more complete discussion of this literature.

yet no evidence of the link of outward foreign direct investment and productivity from the Nordic countries.

### *3.2 Inward foreign direct investment*

Foreign direct investment cannot only take the form of domestic firms from country A buying a firm in a foreign country B or building a new production facility abroad in B (i.e. outward foreign direct investment from the point of view of country A). Foreign firms from country B can buy a domestic firm in country A or build a new production facility in A. This latter type of foreign direct investment is called inward foreign direct investment (from the point of view of country A) and it leads to foreign-owned firms in country A.

Productivity differences between domestic and foreign-owned firms have been investigated in a large number of empirical studies. From a theoretical point of view, it can be expected that foreign-owned firms are more productive than domestic firms because foreign-owned firms can use technological knowledge and management know-how owned by their parent company that made the parent company highly successful and productive (which can be viewed as a prerequisite for the parent company to become a multinational firm with production facilities in a foreign country). Furthermore, it might be the case that foreign investors engage in “cherry picking” – they buy the best and most productive domestic firms. Productivity differentials should then show up between domestic and foreign-owned firms of the same size and from the same industry. The big picture from empirical studies is in line with these hypotheses – foreign-owned firms tend to be more productive than domestic firms.[9]

The empirical evidence we have from studies based on firm-level data from the Nordic countries (summarized in Table 1) is in line with the findings from the international literature. Dachs et al. (2008) report that foreign-owned enterprises exhibit a significantly higher labour productivity than do domestically owned enterprises in all four Nordic countries. Ilmakunnas and Maliranta (2004) find that foreign ownership increases total factor productivity by between 9 and 11 per cent in Finland. Balsvik and Haller (2010) show that in Norway prospective foreign owners pick high-

[9] For a survey of empirical studies on productivity differentials between foreign-owned firms and domestic firms, see Barba Navaretti and Venables (2004, p. 155-162).

productivity plants and that labour productivity increases after foreign acquisitions. A similar positive effect of the acquisition of home firms by foreign firms on productivity is reported by Karpaty (2007) for Sweden; Bandick (2009), however, finds that targeted Swedish firms have a better productivity growth after vertical foreign acquisition only and not after horizontal foreign acquisition.

## 4. Offshoring

The third type of international firm activity besides international trade (exports and imports) and (outward and inward) foreign direct investment that is looked at in this paper with a view on its relation to productivity is offshoring. Offshoring means different things to different people. Here, offshoring describes the relocation of processes to any foreign country without distinguishing whether the provider is external or affiliated with the firm. It is a process whereby an activity which was previously undertaken in-house is contracted to another supplier in a foreign country. For example, a firm located in Germany that produces various types of clothes decides to stop the production of one of its products, simple shirts, in the factory in Germany and relocate this production to a plant in Poland that is owned by the German firm, or to an independent subcontractor in Romania.

There is evidence that offshoring firms differ systematically from non-offshoring firms.[10] In a comprehensive survey of the literature, Görget al. (2008, p. 34) ask "whether, among a random sample of firms we would expect all to engage in offshoring or whether it is only a certain group of firms that do so". According to the authors the "short answer to this is: only a certain group – and we would expect this to comprise the 'better' firms in any sample." Görg et al. (2008, p. 35) summarize empirical evidence from a number of studies showing that offshoring firms are more productive.

If firms that relocated parts of their activities abroad are more productive than non-offshoring firms at a certain point in time, this might be caused by self-selection of "better" firms into offshoring. Self-selection

[10] For a more complete summary of the literature on offshoring and productivity, see Wagner (2011).

would be in line with recent developments in economic theory of international firm activities. Offshoring involves substantial sunk costs related to searching for a foreign partner, doing market research, fixing contractual arrangements etc. Therefore, only the more productive firms will be able to overcome these sunk cost barriers and successfully start to offshore (see Antràs and Helpman (2004) and Görg et al. (2008, p. 34f.). Studies focusing on the consequences of offshoring for productivity are rare (see Olsen, 2006, p. 9). Görg et al. (2008, p. 8) summarize the findings by stating that for manufacturing firms, offshoring results in higher labour productivity.

For West German firms from manufacturing industries, Wagner (2011) finds that, compared to non-offshoring firms, offshoring firms are more productive. These differences existed in the year before some firms started to offshore, and this points to self-selection of more productive firms into offshoring. This finding is in line with results from recent theoretical models and with results for other countries.

To investigate the effects on productivity of relocation across borders, the performance of firms with and without offshoring was compared for 2000-2004 when some firms were relocating firms and others were not. Looking at first-time offshorers in 2001-2003, he finds no evidence for a causal effect of offshoring on productivity. When the group of offshoring firms includes firms with offshoring activities before 2001, however, he reports a positive and statistically significant causal effect, but this effect is rather small.

The bottom line, then, is that the empirical evidence points to higher productivity in offshoring firms compared to non-offshoring firms, and to self-selection of more productive firms into offshoring, while the jury is still out on the issue whether offshoring improves productivity in a firm or not. Note that due to the absence of empirical studies on offshoring and productivity that are based on firm-level data from the Nordic country, we have no evidence at all on this issue for these countries.

## 5. Discussion

The big picture shown by the literature on the micro-econometrics of international firm activities for the links between productivity and engagement on foreign markets can be summarized as follows. Details aside, internationally active firms are more productive than firms of the same size and from the same industry that are active on the domestic market only. The direction of causality between productivity and international activities is somewhat less clear. Higher productivity seems to be a prerequisite for international activities, and we have ample evidence for self-selection of more productive firms into exporting, foreign direct investment and offshoring. Whether international activities cause higher productivity or not, however, is still an unresolved question, and there are empirical studies reporting evidence for or against the learning-by-international-activities hypothesis.

Empirical evidence from firm-level data based studies for the Nordic countries on the links between international firm activities and productivity is scarce (see Table 1). While there are studies on exports and productivity and on productivity differences between foreign-owned firms and domestic firms for Denmark, Finland, Norway and Sweden, evidence for imports is only available for Denmark and Sweden, while there is no empirical evidence at all for the relationship between productivity and both outward foreign direct investment and offshoring. Furthermore, there is often only one study dealing with one form of international firm activity for a country. The empirical evidence available is therefore no sound basis that allows searching for similarities and differences (and their causes) between the Nordic countries, and the empirical results reported cannot qualify as stylized facts that can inform policy makers in an evidence-based way.[11] A suggested step on the way to a solid empirical basis for such an exercise is an ex-ante coordinated international comparative study that uses identically specified empirical models and comparable firm-level data for all forms of international firm activities covering firms from all Nordic countries and for other countries that are considered to be useful as a benchmark.[12]

An important topic that is not dealt with in this paper but that is closely related to the links between productivity and international activities of firms is the cause for productivity differences between firms. In the theoretical models à la Melitz (2003) that are at the core of the *new new trade*

---

[11] This is yet another case that illustrates what I call Bartelsman's Lament: "For policy makers, the state of affairs of productivity research is frustrating, at best" (Bartelsman 2010, p. 1891). For a comprehensive discussion of how to go from estimation results to stylized facts when empirically investigating international activities of heterogeneous firms, see Wagner (2010).

[12] See International Study Group on Exports and Productivity (ISGEP) (2008) for an example of such a study and Wagner (2010) for a broader discussion of this approach.

*theory*, the productivity of a firm is the result of a random draw from a productivity distribution. While this is for certain an appropriate approach to building a theoretical model for trade with heterogeneous firms, it is far from satisfactory from an empirical point of view. Obviously, there is a role for random shocks, or good or bad luck, in shaping the productivity level of a firm, but we have good reasons to believe that a high or low level of productivity is not a matter of luck alone. Productivity can be expected to be related to the quality of inputs used in the production process, and to the way these production factors are combined.

Using a knowledge production framework and a rich set of plant level data, Wagner (2008) demonstrates that in Germany, firms that are active on international markets as exporters or foreign direct investors do generate more new knowledge than firms which sell on the national market only. These differences are not only due to a larger firm size, or different industries, or the use of more researchers in these firms, but also to the fact that these globally engaged firms learn more from external sources. The importance of these knowledge sources varies with the type of innovation – for example, cooperation in R&D with universities and other research institutes matters in the case of new patents registered, while suppliers are important in the case of the share of new products in total sales and new production processes introduced. These results, which are broadly in line with the findings by Criscuolo et al. (2005) in their study using UK firm-level data, can help explain the strong positive correlation between productivity and international activities of firms. Firms that are active on markets beyond the national borders generate higher levels of new knowledge that feed into higher productivity.

Another important aspect that might help explain the positive productivity premium of internationally active firms is management quality. Although management quality has been considered as an important source of performance differences between firms for a very long time – Syverson (2010, p. 14) mentions a study published in 1887 that made this point – empirical evidence on this is scarce due to data limitations. As expressed by Syverson (2010, p. 14), "(t)he identity, much less the characteristics, practices, or time allocation of individual managers are rarely known. Furthermore, managerial inputs can be very abstract. It's not just time allocation that matters, but what the managers do with their time, like how they incentivize workers or deal with suppliers." A recent study by Bloom and

Van Reenen (2010) that relates management practices to productivity shows, among others, that firms that export (but do not produce) overseas are better-managed than domestic non-exporters, but are worse-managed than multinationals.

These findings demonstrate that there is more than sheer luck behind the observed productivity differences between firms. Before speculating about the implications of these results for the design of policy measures, however, one should remember that productivity differences at the firm level are notoriously difficult to explain empirically. "At the micro level, productivity remains very much a measure of our ignorance" (Bartelsman and Doms 2000, p. 586).

My suggested take-home message for policy makers and their advisers from my reading of the literature on productivity and international firm activities is therefore simple (if not trivial). Any policy measures that foster productivity are policy measures that help make more firms more fit for the world market, and any policy measures that make domestic and foreign markets more open to internationally active firms (like the reduction of barriers to international trade and investment, and measures that ease international financial transactions between firms) help improve productivity and foster economic growth. The positive effect of an increase in international firm activities on productivity does not only follow from learning effects, but also from reallocation effects. Reduced barriers and lower costs for international firm activities will increase these activities by internationally active firms (that are more productive than firms that are active on the domestic market only) and will induce the most productive firms that were previously not internationally active to become active on foreign markets. These highly productive firms will expand. The less productive domestic firms, on the other hand, will shrink or even exit because increased competition on factor markets due to an increase in factor demand by expanding high-productivity firms will drive factor prices up and lower the profitability of these low productive firms. This reallocation of factors of production from less productive to more productive, internationally active firms will lead to an increase in productivity.

Following the holy principle of comparative advantage, however, I leave any detailed suggestions for policy measures that foster productivity or make markets more open to international activities of firms to specialists in that field.

## References

Andersson, M. and Lööf, H. (2009a), Learning-by-exporting revisited: The role of intensity and persistence, Scandinavian Journal of Economics 111, 893-916.

Andersson, M. and Lööf, H. (2009b), Key characteristics of the small innovative firm, CESIS Electronic Working Paper Series 175, Royal Institute of Technology, Stockholm.

Andersson, M., Lööf, H. and Johansson, S. (2008), Productivity and international trade: Firm level evidence from a small open economy, Review of World Economics 144, 774-801.

Antràs, P. and Helpman, E. (2004), Global sourcing, Journal of Political Economy 112, 552-580.

Balsvik, R. and Haller, S.A. (2010), Picking "lemons" or picking "cherries"? Domestic and foreign acquisitions in Norwegian manufacturing, Scandinavian Journal of Economics 112, 361-387.

Balsvik, R. and Haller, S.A. (2011), Foreign firms and host-country productivity: Does the mode of entry matter?, Oxford Economic Papers 63, 158-186.

Bandick, R. (2009), Foreign acquisition, wages and productivity, mimeo, University of Aarhus.

Bandick, R. (2010), Multinationals and plant survival, Review of World Economics 146, 609-634.

Bandick, R. and Görg, H. (2010), Foreign acquisition, plant survival, and employment growth, Canadian Journal of Economics 43, 547-573.

Bandick, R., Görg, H. and Karpaty, P. (2010), Foreign acquisition, Domestic multinationals, and R&D, Institute for the Study of Labor IZA DP 5252, Bonn.

Barba Navaretti, G. and Venables, A.J. (2004), Multinational Firms in the World Economy, Princeton University Press, Princeton and Oxford.

Bartelsman, E.J. (2010), Searching for the sources of productivity from macro to micro and back, Industrial and Corporate Change 19, 1891-1917.

Bartelsman, E.J. and Doms, M. (2000), Understanding productivity: Lessons from longitudinal micro data, Journal of Economic Literature 38, 569-594.

Bernard, A.B. and Bradford Jensen, J. (1995), Exporters, jobs, and wages in U.S. manufacturing: 1976-1987, Brookings Papers on Economic Activity, Microeconomics 1, 67-119.

Bernard, A.B. and Bradford Jensen, J. (1999), Exceptional exporter performance: Cause, effect, or both?, Journal of International Economics 47, 1-25.

Bernard, A.B., Bradford Jensen, J., Redding, S.J. and Schott, P.K. (2007), Firms in international trade, Journal of Economic Perspectives 21, 105-130.

Bernard, A.B. and Wagner, J. (1997), Exports and success in German manufacturing, Weltwirtschaftliches Archiv/Review of World Economics 133, 134-157.

Bloom, N. and Van Reenen, J. (2010), Why do management practices differ across firms and countries?, Journal of Economic Perspectives 24, 203-224.

Castellacci, F. (2010), Firm heterogeneity, international cooperations and exportparticipation, NUPI Working Paper 781, Oslo.

Criscuolo, C., Haskel, J.E. and Slaughter, M.J. (2005), Global engagement and the innovation activities of firms, NBER Working Paper 11479.

Dachs, B., Ebersberger, B. and Lööf, H. (2008), The innovative performance of foreign-owned enterprises in small open economies, Journal of Technology Transfer 33, 393-406.

De Loecker, J. (2010), A note on detecting learning by exporting, NBER Working Paper 16548.

Deschryvere, M. and Kotiranta, A. (2008), Domestic employment effects of offshoring: Empirical evidence from Finland, ETLA – The Research Institute of the Finnish Economy Discussion Papers 1166, Helsinki.

Ekholm, K. and Nilsson Hakkala, K. (2008), The effect of offshoring on labor demand: Evidence from Sweden, mimeo, Stockholm University and Helsinki School of Economics.

Eliasson, K., Hansson, P. and Lindvert, M. (2009), Do firms learn by exporting or learn to export? Evidence from small and medium-sized enterprises (SMEs) in Swedish manufacturing, Örebro University Working Paper 15/2009, forthcoming in Small Business Economics.

Eliasson, K., Hansson, P. and Lindvert, M. (2010), Jobs and exposure to international trade within the service sector in Sweden, Uppsala University Center for Labor Studies Working Paper 2010:17.

Eriksson, T., Smeets, V. and Warzynski, F. (2009), Small open economy firms in international trade: Evidence from Danish transaction-level data, Aarhus School of Business, Aarhus University Working Paper 09-7.

Fosse, H.B. and Maitra, M. (2010), Impact of offshoring on firm average wage: Analyzing rent sharing and composition effect, mimeo, Copenhagen Business School and Columbia University.

Fryges, H. and Wagner, J. (2010), Exports and profitability – First evidence for German manufacturing firms, The World Economy 33, 399-423.

Geishecker, I., Görg, H. and Roland Munch, J. (2010), Do labour market institutions matter? Micro-level wage effects of international outsourcing in three European countries, Review of World Economics 146, 179-198.

Görg, H., Greenaway, D. and Kneller, R. (2008), The economic impact of offshoring, GEP Research Report, Leverhulme Centre for Research on Globalisation and Economic Policy, University of Nottingham.

Greenaway, D., Gullstrand, J. and Kneller, R. (2005), Exporting may not always boost firm productivity, Review of World Economics 141, 561-582.

Greenaway, D., Gullstrand, J. and Kneller, R. (2008a), Surviving globalization, Journal of International Economics 74, 264-277.

Greenaway, D., Gullstrand, J. and Kneller, R. (2008b), Firm heterogeneity and the geography of international trade, University of Nottingham Leverhulme Centre for Research on Globalisation and Economic Policy Research Paper Series 2008/41.

Greenaway, D., Gullstrand, J. and Kneller, R. (2009), Live or let die? Alternative routes to industry exit, Open Economies Review 20, 317-337.

Greenaway, D. and Kneller, R. (2007), Firm heterogeneity, exporting and foreign direct investment, Economic Journal 117, F134-F161.

Gullstrand, J. (2009), Firm and destination-specific export costs: The case of the Swedish food sector, mimeo, Lund University, forthcoming in Food Policy.

Hansson, P. and Lundin, N.N. (2004), Exports as an indicator on or promoter of successful Swedish manufacturing firms in the 1990s, Review of World Economics 140, 415-445.

Hayakawa, K., Kimura, F. and Machikita, T. (2011), Firm-level analysis of globalization: A survey of eight literatures, Journal of Economic Surveys 25, 19-68.

Helpman, E. (2006), Trade, FDI, and the organization of firms, Journal of Economic Literature 44, 589-630.

Helpman, E., Melitz, M.J. and Yeaple, S.R. (2004), Export versus FDI with heterogeneous firms, American Economic Review 94, 300-316.

Heyman, F., Sjöholm, F. and Gustavsson Tingvall, P. (2007), Is there really a foreign ownership wage premium? Evidence from matched employer-employee data, Journal of International Economics 73, 355-376.

Hummels, D., Jorgensen, R., Munch, J. and Xiang, C. (2010), The wage and employment effects of outsourcing: Evidence from Danish matches worker-firm data, mimeo, Purdue University and University of Copenhagen.

Huttunen, K. (2007), The effect of foreign acquisition on employment and wages: Evidence from Finnish establishments, Review of Economics and Statistics 89, 497-509.

Ibsen, R., Warzynski, F. and Westergard-Nielsen, N. (2009), Emplyoment growth and international trade: A small open economy perspective, Aarhus School of Business Working Paper 09-9.

Ilmakunnas, P. and Maliranta, M. (2004), Foreign medicine: A treatment effect analysis of the productivity effects of foreign ownership, Applied Economics Quarterly 50, 41-60.

Ilmakunnas, P., Maliranta, M. and Pesola, H. (undated), Spillovers from multinationals to domestic firms: An empirical analysis of the profitability effects of labour flows, mimeo, Aalto University and ETLA.

Ilmakunnas, P. and Nurmi, S. (2010), Dynamics of export market entry and exit, Scandinavian Journal of Economics 112, 101-126.

International Study Group on Exports and Productivity (ISGEP) (2008), Understanding cross-country differences in exporter premia: Comparable evidence for 14 countries, Review of World Economics 144, 596-635.

Irarrazabal, A., Moxnes, A. and Ulltveit-Moe, K.H. (2010), Heterogeneous firms or heterogeneous workers? Implications for exporter premia and the gains from trade, mimeo, University of Oslo and Dartmouth College.

Karpaty, P. (2006), Does multinationality matter? Evidence from Swedish firm data, Applied Economics Quarterly 52, 101-122.

Karpaty, P. (2007), Productivity effects of foreign acquisitions in Swedish manufacturing: The FDI productivity issue revisited, International Journal of the Economics of Business 14, 241-260.

Karpaty, P. and Lundberg, L. (2004), Foreign direct investment and productivity spillovers in Swedish manufacturing, FIEF Working Paper Series 2004:194, Stockholm.

Laursen, K. (2008), The effect of knowledge sources for export performance in manufacturing and services: Danish firm-level evidence, DRUID, mimeo, Copenhagen Business School.

Lehto, E. and Böckermann, P. (2008), Analysing the employment effects of mergers and acquisitions, Journal of Economic Behaviour & Organization 68, 112-124.

Lööf, H. (2009), Are services different exporters?, mimeo, Royal Institute of Technology, Stockholm.

Lööf, H. and Andersson, M. (2010), Imports, productivity and origin markets: The role of knowledge-intensive economies, The World Economy 33, 458-481.

López, R.A. (2005), Trade and growth: Reconciling the macroeconomic and microeconomic evidence, Journal of Economic Surveys 19, 623-648.

Lundin, N. and Yun, L. (2009), International trade and inter-industry wage structure in Swedish manufacturing: Evidence from matched employer-employee data, Review of International Economics 17, 87-102.

Mayer, T. and Ottaviano, G.I.P. (2007), The Happy Few: The Internationalisation of European Firms. New Facts Based on Firm-level Evidence, Bruegel Blueprint Series Volume III, Brussels.

Melitz, M.J. (2003), The impact of trade on intra-industry reallocations and aggregate industry productivity, Econometrica 71, 1695-1725.

Munch, J.R (2010), Whose job goes abroad? International outsourcing and individual job separations, Scandinavian Journal of Economics 112, 339-360.

Munch, J.R. and Skaksen, J.R. (2008), Human capital and wages in exporting firms, Journal of International Economics 75, 363-372.

Olsen, K.B. (2006), Productivity impacts of offshoring and outsourcing: A review, OECD Directorate for Science, Technology and Industry STI Working Paper 2006/1.

Pesola, H. (2007), Foreign ownership, labour mobility and wages, HEER – Helsinki Center of Economic Research Discussion Paper 175.

Pesola, H. (2009), Labour market transitions following foreign acquisitions, HEER – Helsinki Center of Economic Research Discussion Paper 251.

Redding, S.J. (2010), Theories of heterogeneous firms and trade, mimeo, Princeton University and CEPR, forthcoming in Annual Review of Economics.

Schank, T., Schnabel, C. and Wagner, J. (2007), Do exporters really pay higher wages? First evidence from German linked employer-employee data, Journal of International Economics 72, 52-72.

Schank, T., Schnabel, C. and Wagner, J. (2010), Higher wages in exporting firms: Self-selection, export effect, or both? First evidence from linked employer-employee data, Review of World Economics 146, 303-322.

Smeets, V. and Warzynski, F. (2010), Learning by exporting, importing or both? Estimating productivity with multi-product firms, pricing heterogeneity and the role of international trade, mimeo, Aarhus University.

Syverson, C. (2010), What determines productivity?, NBER Working Paper 15712.

Vogel, A. and Wagner, J. (2009), Higher productivity in importing German manufacturing firms: Self-selection, learning from importing, or both?, Review of World Economics 145, 641-665.

Wagner, J. (2002), The causal effects of exports on firm size and labor productivity: First evidence from a matching approach, Economics Letters 77, 287-292.

Wagner, J. (2006), Exports, foreign direct investment, and productivity: Evidence from German firm level data, Applied Economics Letters 13, 347-349.

Wagner, J. (2007), Exports and productivity: A survey of the evidence from firm-level data, The World Economy 30, 60-82.

Wagner, J. (2008), International firm activities and innovation: Evidence from knowledge production functions for German firms, Icfai Journal of Knowledge Management 6, 47-62.

Wagner, J. (2010), From estimation results to stylized facts. Twelve recommendations for empirical research in international activities of heterogeneous firms, University of Lüeneburg Working Paper Series in Economics 186, forthcoming in De Economist.

Wagner, J. (2011), Offshoring and firm performance: Self-selection, effects on performance, or both?, Review of World Economic 147, 217-247.

Ylä-Anttila, P., Ali-Yrkkö, J. and Nyberg, M. (2004), Foreign ownership in Finland – Boosting firm performance and changing corporate governance, ETLA – The Research Institute of the Finnish Economy Discussion Paper 904.

# Comment on Wagner: Productivity and international firm activities: What do we know?

Markku Stenborg*

Having spent a large part of my working life applying economics to various types of policy questions rather than generating more economics in ivory towers, my bias here is policy relevance. From that angle, this is an important and interesting survey of a well-defined research question. The story is tightly focused and useful for evidence-based policy advice.

The paper looks at the question why some firms choose to export, or be involved in other international activities, while some other firms do not. The main message here is that only more productive firms become exporters. This result is based on firm or plant level econometrics of heterogeneous firms. The approach allows due attention to details (which is not always granted in economics).

The literature has offered two suggestions as explanations for productivity premium separating exporters from other firms. First, only more productive firms choose to export and, second, exporting makes firms more productive. The main story behind the first suggestion is that international activities involve extra costs, which more efficient firms can cover. This is self-selection: only better firms go international. The narrative behind the second alternative is learning-by-international-activities: firms meet a wider set of circumstances, have to fend off tougher rivals and a more diverse set of strategies, etc., which makes them better. Note that the stories need not be pure alternatives, but both can be true simultaneously.

* Ministry of Finance, Finland, markku.stenborg@vm.fi.

The overarching empirical result widely supports the self-selection but there is much less evidence for the learning story. Furthermore, internationally active firms are better also in many other dimensions. More productive firms are also more likely to be involved in other international activities such as imports, cross-border M&As and foreign direct investments. The empirical story is broadly the same self-selection: only more productive firms find it profitable to expand the extra effort needed to surmount the hurdle at the border.

Next, I have three comments on the scope of the literature as seen in this survey and then I will briefly discuss some potential for policy conclusions.

First, I was somewhat surprised not to find comparative advantage, the main theoretical result behind trade, to even be mentioned in the paper. For instance, a layman might expect that some firms are in a position to better exploit the comparative advantage of their home nation, irrespective of their productivity difference *vis-à-vis* neighboring firms. Or could it be that exporting firms are more productive in part because of their ability to benefit from comparative advantage? Is this something we do, can or want to control for in empirical analysis? Or is the interplay between comparative advantage and the effect on productivity of international activities too trivial for the enlightened not to warrant a comment?

Second, productivity is not a total black box. It is shaped, in part, by resource allocation, R&D, innovation and "creative destruction", to name a few obvious candidates. It would be fruitful to have a better understanding of the interaction between success in international activities and innovation (or other drivers of productivity).

This leads to my third point, in the form of a question. Could it be that there are some common factors behind productivity and international activities? For instance, suppose that firms that are better at networking domestically become more productive and are also more able to network internationally. Then, the export driver is the common force behind both exports and productivity, not productivity *an sich*.

Turning to policy conclusions, note that this paper is quite humble. It concludes that there is not enough country-specific evidence to make sharp Nordic conclusions. How complete knowledge do we need here? Which institutions and other details matter for conclusions and prohibit us from using results from other industrialized countries? In some sense, this humility is admirable. It is not too uncommon to find papers trying to extend their policy conclusions well beyond the proper realm.

Are some firms more productive by accident? If so, there is not much new policy insight from this literature. Remove obstacles from productivity gains, and try to find means to lower the hurdle firms meet at the border, and wait and see more firms becoming active internationally, but that is about it. Are we equally confused but at a deeper level? Or can policy reforms affect the accident that shifts productivity to the level that permits exporting and other international business?

So, perhaps a wider scope would have been useful here. Policy conclusions from studies outside the Nordic countries and a better understanding of the drivers of productivity and their interplay with international activities would have been nice additions to this fine survey. In any event, to conclude, let me congratulate professor Wagner on his important contributions to this branch of literature and especially on a nice and tightly focused survey. It has been a pleasure to read and comment on this paper.

# Innovation and productivity*

Bronwyn H. Hall**

## *Summary*

What do we know about the relationship between innovation and productivity among firms? The workhorse model of this relationship is presented and the implications of an analysis using this model and the usually available data on product and process innovation are derived. Recent empirical evidence on the relationship between innovation and productivity in firms is then surveyed. The conclusion is that there are substantial positive impacts of product innovation on revenue productivity, but that the impact of process innovation is more ambiguous, suggesting that the firms being analyzed possess some market power.

*Keywords:* Innovation, TFP, revenue productivity, cross-country.
*JEL classification numbers*: O30, D20, L20.

---

* Paper prepared for the Nordic Economic Policy Conference on productivity and competitiveness, 29 April 2011, Helsinki, Finland. The revision of this paper has benefitted greatly from comments by Ari Hyytinen, Jakob Madsen, Jacques Mairesse and Anders Sørensen. I also thank John Jankowski for essential help with the US innovation data.

** University of Maastricht, University of California at Berkeley, NBER and IFS, bhhall@econ.berkeley.edu.

Early work on the sources of productivity growth revealed that growth in capital and labor explained less than half of such growth in the United States and many other countries. The remainder (the ‘residual’) was ascribed to technical change and a large literature emerged that attempted to find measures for technical change (improvements in capital and labor quality, R&D activities, and so forth) and use these measures to try to explain the residual growth in productivity (Griliches, 1996, 1998, among others). Considerable success has been achieved by this approach, to the extent that many countries are now moving to incorporate measures of R&D capital stock in their systems of national income accounts, and therefore to directly attribute some of economic growth to its contribution as well as adding the creation of knowledge capital to output itself.

Driven by an interest in the unexplained portion of productivity growth and partly in response to various economic slowdowns and productivity gaps among nations, a large body of research on innovative activity and productivity in firms has accumulated. For reasons of data availability, this work has mostly used two measures of innovative activity: R&D spending and patent counts.[1] As measures of innovation, each of these has both positive and negative attributes. Both pertain primarily to technological innovation and are more suited to measuring innovation in manufacturing firms than in other areas such as services. R&D spending has the advantage that it is denominated in comparable units (currency) and represents a (costly) decision variable on the part of the firm about its appropriate level of innovative activity. For the same reason, R&D is only an input to innovation and cannot tell us about innovation success. Patent counts are a measure of invention success, and can be considered to be at least a partial measure of innovation output, but they are inherently very noisy (a few are associated with very valuable inventions and most describe inventions of little value) and the extent of their innovation coverage varies by sector, with sectors like pharmaceuticals and instruments making heavy use of patents while other sectors use them to a very small extent.

As the industrial structure of advanced economies has shifted away from manufacturing and towards services, economists and others have gradually become aware that concepts like “technical change” and “R&D” only describe some of the sources of increased productivity in the econo-

[1] A recent survey of results for the R&D-productivity relationship is Hall et al. (2010).

my, and recent research has begun to look at innovation more broadly as a source of growth. This research has been greatly helped by the introduction of the Oslo Manual (Tanaka et al., 2005) with guidelines for the definition of various kinds of innovation and by the surveys of innovative activity in business firms that have been conducted in a large number of countries around the world, mostly using this manual as a guide (Mairesse and Mohnen, 2010). Several non-R&D kinds of innovative expenditure have been identified: the later phases of development and testing that are not included in R&D, capital expenditures related to the introduction of new processes, marketing expenditures related to new products, certain kinds of employee training, expenditures on design and technical specifications, etc.

Figure 1, which is based on data from these kinds of surveys, shows the distribution of the share of firms that report any kind of innovation during the three-year period 2006-2008 by country and size of firm.[2] The figure is instructive: it shows that in most countries, between 30 and 50 percent of the firms introduce a product or process innovation during a three-year period, and that the rate of introduction is much higher and also more even across countries among large firms, as might have been expected. In fact, the coefficient of variation for the innovation share across countries is 0.3 for SMEs and 0.12 for large firms, confirming the higher dispersion rate for SMEs.

Figure 2 shows a breakdown by product and process innovation, where innovation is defined as the development of a process or product that is "new to the firm" by the enterprise or its group.[3] In this case, we are able to compare the European countries to the United States, by restricting the population of firms to a common set of innovative sectors across the two regions.[4] The two types of innovation are roughly equal, with a slight pref-

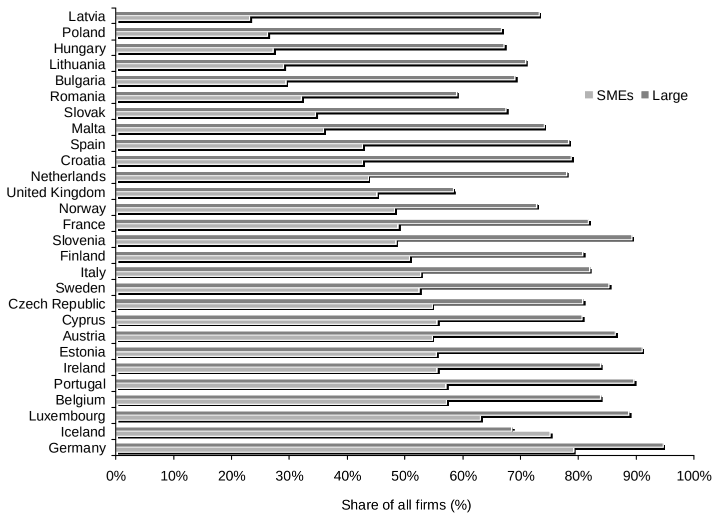
[2] The data for this and the subsequent figure mainly comes from the European Community Innovation Survey, data for the United States comes from the new 2008 Business R&D and Innovation Survey (BRDIS), conducted by the National Science Foundation and may not be exactly comparable to the European data.

[3] In the US case, the definition does not include the group to which the enterprise belongs. Because group structures are rare in the US, this distinction makes little difference. However, it does mean that the European numbers could be slightly higher given the broader definition of the firm doing the innovating.

[4] These sectors are manufacturing, telecommunications, computer services and software publishing, finance, and some technical professional services. The restriction is necessary because the US data does not contain enough detail outside manufacturing to match the innovative sector definition used by Eurostat, which is quite broad. The narrow definition used here is NACE activities C, J58, J61, J62, J63, K, M71. The broader definition used by Eurostat and elsewhere in this

erence for process innovation and some differences across countries. However, it is worth noting that the United States is by no means the most innovative among these countries by this measure, although this conclusion should be viewed with some caution given the slight noncomparability of the US data.[5]

**Figure 1. Innovating firms by size, as a share of all firms, 2006-2008**

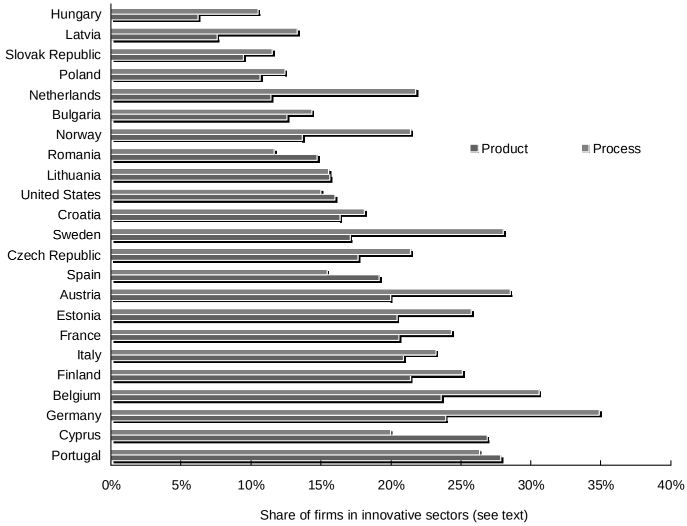



---

paper includes NACE activities B, D, E, G46, and H (mining, utilities, wholesale trade excluding motor vehicles, and transportation and storage).

[5] Although the sampling frame for the BRDIS was the population of US firms with five or more employees, this survey was the successor to the long-running RD-1 survey which was only filled out by RD-doing firms, and the innovation questions were at the end of a long survey, most of which concerned R&D. So there is some suspicion that they may not always have been accurately answered by non-R&D-doers. This suggestion has been informally confirmed by conversations with the NSF. The product innovation rate for 3 percent of the firms that report doing R&D in the BRDIS survey was 66 percent, whereas the rate for non-R&D-doers was 7 percent. The gap, which is much larger than that in Europe, does suggest some undercounting for the non-R&D firms.

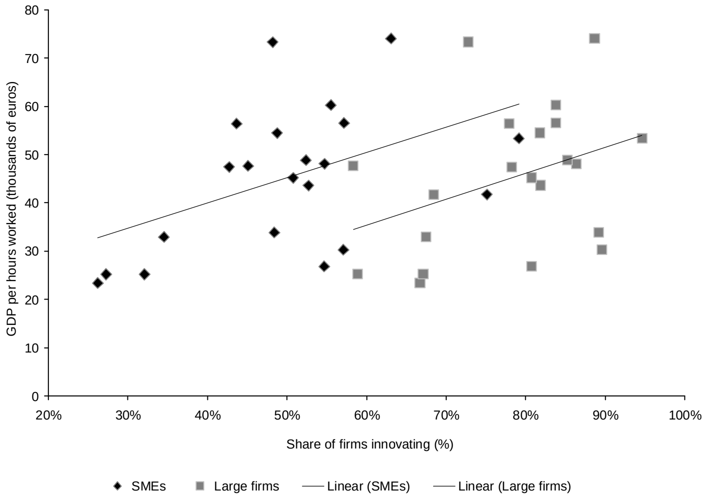
**Figure 2. Share of firms with innovation new to the firm or market, 2006-2008**




How does the aggregate innovation picture compare with aggregate productivity measures? To answer this question, I compared the innovation rates at the country level with overall labor productivity (GDP per hours worked, also from OECD data). The results are shown in Figure 3.[6] With the exception of an outlier (Norway), the share of both SMEs and large firms that innovate appears to be positively related to labor productivity at the country level. Simple univariate regressions for the relationship were moderately significant, and even more so when robust methods such as Least Absolute Deviations or Least Median of Squares were used.

[6] The innovation rate is defined as the share of all firms in innovating sectors that have introduced any new process or product in the past three years, including organizational and marketing innovations. The data are from Eurostat's database for the sixth Community Innovation Survey, variable INNO.

**Figure 3. Labor productivity levels 2009 and innovation 2006-2008**




Although the correlation displayed should not be taken too seriously, given the number of confounding influences and differences in industrial structure across countries, even at the aggregate level there does seem to be a relationship between innovative activity by firms and productivity, albeit one that leaves room for many other influences. It is natural to ask how this relationship comes about – what actions by individual firms lead to aggregate productivity improvements? One can think of two main channels through which the presence of more innovative firms can translate into productivity improvements: first, innovation in existing firms can both increase their efficiency and improve the goods and services they offer, thus increasing demand as well as reducing costs of production. Second, innovating firms are likely to grow more than others and new entrants with better products to offer are likely to displace existing inefficient firms with a concomitant increase in aggregate productivity levels. In both cases, the relationship between innovation and productivity is influenced by the institutional and macroeconomic environment where the firms operate, possibly leading to substantial differences across countries in the relationship between them.

The present paper will review the ways in which economists have analyzed the relationship between productivity and innovation, focusing on the use of such innovation survey data as well as other data on innovative output such as patents. The differing measures of innovation (dummy variables, innovative sales, and innovation expenditure) that the various surveys yield will be reviewed and their drawbacks and advantages discussed. The distinction between innovation input (expenditures and choices under the control of the firm) and innovation output (depending on inputs but also with a large element of chance) is important and there are rationales for using both concepts.

After discussing measures of innovation, the paper will review two approaches to measuring the relationship between productivity and innovation: the econometric or regression approach and the growth accounting approach. Both are in their relative infancy due to the fact that the appropriate data has been lacking until quite recently (and is still not widely available).

## 1. Innovation – the concept and its measurement

There were two early empirical efforts which generated datasets on innovation that have been used in some studies (regrettably few studies, in fact). They are the SPRU study of UK firms begun in 1970, and conducted over a period of 15 years through 1984 (Freeman and Soete, 1997) and the study by Acs and Audretsch during the 1980's that looked at US firm innovations. The SPRU study asked almost 400 experts in industry to identify significant technical innovations that were commercialized in the UK sometime between 1945 and 1983 and then surveyed the firms that had introduced the innovations. The database contains over 4 000 innovations, almost all of which are in the manufacturing sector. It has been used to show that the relationship between innovative activity and firm size is largely U-shaped, and that smaller firms show greater innovative activity than formal R&D activity (Pavitt et al., 1987). A couple of the papers surveyed below (Geroski, 1989 and Sterlacchini, 1989) make use of this database, but it has not been exploited extensively in the analysis of innovation and productivity.

The 1990 Acs and Audretsch study for the US Small Business Administration (SBA) was based on a survey of over 100 trade journals in 1982 that looked for an announcement of the market introduction of inventions. The definition used by the SBA was the following:

> "A process that begins with an invention, proceeds with the development of the invention and results in introduction of a new product, process or service to the marketplace."

This survey yielded over 8 000 US innovations, most of which probably dated 1978-1982, but all of which were introduced in 1982. Acs and Audretsch use these data to analyze the role of small firms in innovation, the growth of firms, and the evolution of market structure. Unfortunately, they do not provide any analysis of the relationship between these invention introductions and firm productivity.

Both the SPRU and the SBA surveys used the innovation as the unit of observation, and any firm-level analysis using these data is therefore only based on innovative firms. In contrast, the innovation surveys described below are conducted at the firm level and sometimes collect data on non-innovative firms as well. Thanks to work by the OECD and others, we now have a definition of innovation done by firms that is fairly standard across a wide range of countries and surveys:

> "An innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations."[7]

Most of the work on innovation described in this paper has been based on surveys that use a version of this definition. Thus, there has been consistency in the definition of the innovation variables across studies, although perhaps not consistency in the interviewees' understanding of the definition. However, note that there is at least one slightly ambiguous feature of the definition, in that it does not define "new" very precisely. Some of the surveys have made a distinction between "new to the firm" innova-

[7] Oslo Manual (OECD, 2005, p. 46).

tions and "new to the market" innovations, which can be a way of distinguishing more radical innovation from imitation. But in general, the interpretation of "new" is left to the survey respondent.

In spite of the apparent clarity of the definition of innovation in the Oslo Manual, measuring innovation in a form that is useful for statistical analysis has proved challenging. The central problem is that no two innovations are alike. Some innovations (e. g., the invention of the telephone or perhaps the telegraph) create a whole new market sector whereas others are useful but trivial, and there is a wide range in between. In general, we can say that smaller innovations are more numerous than game-changing ones. As shown in Table 1, this fact is very visible in the data collected by Acs and Audretsch. During the year 1982, over 85 percent of the innovations they identified were modest improvements of existing products, and none created entire new markets. Fewer than 2 percent were even considered to be the first of its type on the market in existing market categories.[8]

**Table 1. Manufacturing sector innovations by significance**

| | Number | | Share (%) | |
|---|---|---|---|---|
| | Large firms | Small firms | Large firms | Small firms |
| Establishes whole new categories | 0 | 0 | 0.00 | 0.00 |
| First of its type on the market in existing categories | 50 | 30 | 1.76 | 1.43 |
| A significant improvement on existing technology | 360 | 216 | 12.70 | 10.27 |
| Modest improvement designed to update existing products | 2 424 | 1 858 | 85.53 | 88.31 |
| Total | 2 834 | 2 104 | | |

Source: Acs and Audretsch (1990, Table 2.3).

The innovation surveys have typically measured innovation in two ways: first, by asking whether the firm introduced an innovation of a certain type (product, process, organisational, marketing, etc.) during a preceding period (usually the past three years) and second, by asking what share of the firm's sales is due to products introduced during the same preceding period. The first measure has a number of drawbacks, which

[8] Note that by using the 1982 date, Acs and Audretsch did miss two major innovations: the IBM personal computer and Microsoft DOS, both of which were introduced in 1981 and which arguably meet the definition of "created entire new market".

have become quite evident as it has been used in many empirical studies. When examined across a range of firm sizes, it produces the misleading results that larger firms are more likely to be innovative, whereas in truth larger firms are involved in a wider range of activities and are therefore more likely to have an innovation in at least one of them. So this variable cannot be used to make the kind of statements that one sometimes hears, such as “large firms are more innovative than small firms.”

Another problem is the previously mentioned unequal size of innovations and the failure in some surveys to distinguish between “new to the market” and “new to the firm.” Based on the Acs and Audretsch results, we know that many more of the innovative firms will have introduced improvements to existing products rather than entirely new goods and services, but the latter may be more important than the former. This view of the “skewness” of innovation values is supported by a large amount of research on the valuation of patented inventions (Harhoff et al., 1999; Scherer and Harhoff, 2000; Hall et al., 2005). Although patented inventions are not precisely the same as innovations, they are similar and share some of their distributional properties, with the majority being worth very little, and a few that are quite valuable to their owners.

Because of the imprecision and noisiness of the innovation dummies, many researchers prefer to use the second measure, the share of sales of innovative products, which does give a good indication of how important the innovation(s) were overall for the firm in question. Unfortunately, this measure is useful only for goods and services and cannot be used to capture process or organisational innovation. Nevertheless, it is the one relied on by more than half of the papers discussed in the following sections, often accompanied by a dummy for process innovation. Only one example exists where firms were asked to quantify the impact of process innovation on cost reduction (Peters 2006, for Germany).

## 2. Productivity – the concept and its measurement

What we mean by the term “productivity” is fairly easy to understand although difficult to measure: it is the quantity of output that can be produced using a given level of inputs. At this level of the definition, there is not even a presumption of optimality or efficiency in production. However, normally we assume that the entity whose productivity we wish to measure is "efficient" in the sense that it is using the minimum necessary level of inputs to produce a certain level of output, given its level of technological knowledge, its organization, its size, and other endowments, as well as the environment in which it operates.

Economists generally describe the relationship between output and the level of inputs using a production function, of which the most convenient for analysis is the following:[9]

$$Q = AC^{\alpha}L^{\beta}, \tag{1}$$

where $Q$ is output, $C$ is the level of capital stock, and $L$ is labor (and potentially other non-capital inputs).[10] $A$ is the overall level of productivity which may vary across entities. That is, because of organizational differences, frictions, or other constraints, entities with identical levels of $C$ and $L$ may not be able to achieve the same level of output $Q$.

For measurement purposes, the logarithm of equation (1) is taken:

$$q_{it} = a_{it} + \alpha c_{it} + \beta l_{it} \qquad i = entity, t = time, \tag{2}$$

where the added subscripts denote the fact that productivity levels are usually measured for a number of entities over several time periods. Equation (2) yields an expression for total factor productivity (usually denoted TFP):

$$TFP \equiv a_{it} = q_{it} - \alpha c_{it} - \beta l_{it}. \tag{3}$$

All well and good, but measuring TFP therefore requires measures of real output $Q$, real capital stock $C$, and labor input $L$ (as well as other pos-

[9] I ask the well-informed reader for patience with the elementary review provided here, which is primarily for the purpose of setting the notation for the subsequent discussion.

[10] The treatment here has been greatly simplified by omitting purchased inputs (such as materials, energy, etc.). In practice, these inputs are more important on a share basis than either capital or labor and need to be included in the estimation (typically accounting for about 0.7 of the inputs). Alternatively, one can measure output as value added, which is usually defined as output less purchased inputs. The precise choice of what to include or exclude depends to some extent on data availability, and several variations have been pursued in the literature discussed here. In particular, many of the available datasets do not include measures of the firm's capital stock and researchers are forced to resort to proxies such as current investment spending.

sible inputs, such as energy and materials), to say nothing of the coefficients $\alpha$ and $\beta$. I discuss the latter problem first.

There are two widely used approaches to estimating the weights $\alpha$ and $\beta$ to be applied to the inputs in the productivity measure: 1) assume that input markets are competitive, which implies that the coefficients are the shares of revenue received by each of the factors;[11] and 2) assume that the coefficients are (roughly) constant across entities and estimate them via regression. Solution (1) is favored by statistical agencies and others who simply need a measure of TFP for an individual entity and may not have a sample available for estimation, and solution (2) is the one typically used by econometricians and the main one employed in the literature discussed later in this paper, although there are some exceptions.[12]

The second problem, how to measure the inputs and outputs themselves, is subject to a multitude of solutions. Unfortunately, the choices can have a considerable impact not only on the measurement of TFP but also on the relation of that measure to innovation. The difficulty lies in the measurement of real inputs and outputs, holding constant the unit of measure over time. To take a concrete and well-known example, computers, which are a component of capital, have changed considerably over time. If we measure their contribution to the inputs simply as expenditure on computers, it is likely to be roughly constant over time, and TFP will grow as the computers become more productive. However, if we instead deflate the computer expenditure by an index of the effective price of computing power, which has fallen dramatically over the past thirty years, the real quantity of computers will grow substantially during the same period, and TFP growth will be correspondingly less. In essence, some technical change or innovation has been transferred from TFP to its inputs.[13] The same argument applies to labor input, where quality has probably generally increased over time so that a person-hour thirty years ago is not the same

---

[11] This approach can be modified to account for scale economies and market power as in R. Hall (1988), or indeed almost anything that implies homogeneity of some degree in the production function. See below for a modification that allows the firms to have some degree of market power.

[12] A large literature has developed on the methodologies for estimating the production function in the presence of simultaneity between input and output choice and errors of measurement. Some key papers are Blundell and Bond (2000), Griliches and Mairesse (1984) and Olley and Pakes (1996).

[13] Naturally, if the analysis is done at the aggregate level, the production of computers will be in the output measure, and their share of TFP will increase. See Denison (1966) and Jorgenson and Griliches (1967) for a discussion of this point.

as a person-hour today. All this means that TFP measures need to be used carefully, with an understanding of the approach used to deflation and quality adjustment.[14] That is, much of the effects of innovation may show up as higher quality inputs if they are quality adjusted, and will not appear in output.

For the output measure, the problem is even more striking when we look at the level of the firm or enterprise, because of the potential for variations in market power across firms, and for the role played by innovation in creating and/or increasing that market power. The easiest way of seeing this is to rewrite the TFP equation in terms of revenue rather than real output, under the assumption of an iso-elastic demand equation. The idea behind this approach is that each firm produces differentiated products and therefore faces its own downward sloping demand curve. Firms have idiosyncratic output prices, so that deflation of revenue by an overall deflator simply yields real revenue rather than an actual output measure. I denote the log of real revenue by $r_{it}$ and the log of the firm's output price by $p_{it}$, with $r_{it} = p_{it} + q_{it}$. Write the iso-elastic demand equation facing the firm in logarithmic form as follows:[15]

$$q_{it} = \eta p_{it}\,, \tag{4}$$

where $\eta$ is the (negative) demand elasticity. Combining equations (2) and (4) yields the following expression for the (observable) revenue as a function of the inputs and TFP:

$$r_{it} = \frac{\eta + 1}{\eta}(a_{it} + \alpha c_{it} + \beta l_{it})\,. \tag{5}$$

---

[14] On the output side, Hall (1996), Mairesse and Hall (1996) and Griliches (1994) present R&D-productivity regressions that illustrate the effect that a properly measured computing sector deflator can have on the measured returns to R&D via its impact on the measurement of TFP. Those authors show that using a hedonic price deflator for computing rather than an overall GDP deflator more than triples the elasticity of output with respect to R&D, from 0.03 to 0.11. That is, most of the returns to R&D during the period estimated (1980's) went to price reduction and real output increase, and very little was received by the firms in the form of increased revenues. See also OECD (2003, pp. 43-44), for a discussion of this issue.

[15] This treatment of the problem is drawn from Griliches and Mairesse (1984). See also Mairesse and Jaumandreu (2005) and Foster et al. (2008) for discussions of the differences between revenue productivity estimation and true productivity estimation.

The above equation implies that the estimated coefficients of capital and labor in the productivity equation will be negative if demand is inelastic $(0>\eta>1)$ and biased downward if demand is elastic $(\eta<-1)$. As $\eta$ approaches $-\infty$ (perfectly elastic, or price-taking), the bias disappears and the equation is identical to equation (2), but with revenue in place of output.

The conclusion is that if a regression based on equation (5) is used to estimate TFP $(a_{it})$, the estimate will typically be biased downward over a reasonable range of demand elasticities. Note also that for a profit-maximizing firm, the bias is equal to $1-m$, where $m$ is the markup. The further we are from perfect competition $(m=1)$ and the higher the markup, the greater is the downward bias. After presenting the basic model that shows the relation between innovation and productivity in the next section, I will derive the implications of equation (5) for the measurement of that relationship.

## 3. Modeling the relationship

When looking at the contribution of innovative activity to productivity, the usual starting point is to add a measure of the knowledge or intangible capital created by innovative activity to the production function:

$$Q = AC^{\alpha}L^{\beta}K^{\gamma}\ . \qquad (6)$$

Here, $K$ is some kind of proxy for the knowledge stock of the firm. $K$ can stand for a number of aspects of the entity's innovative capability: its technological knowledge obtained via R&D, its competency at transforming research results into useful products and processes, and so forth. It can even be based on innovative success rather than capability. Traditionally, $K$ has been measured as a stock of past R&D spending but as other kinds of data have become available, other measures involving patents or innovation indicators have been used.

As before, the logarithm of equation (1) is taken:

$$q_{it} = a_{it} + \alpha c_{it} + \beta l_{it} + \gamma k_{it} \qquad i = entity, t = time\ . \qquad (7)$$

Because much of innovative activity is directed towards new products and product improvement, it is useful to rewrite the demand equation to allow the knowledge stock to shift the demand curve facing the firm:

$$q_{it} = \eta p_{it} + \varphi k_{it} \qquad \varphi > 0 \tag{8}$$

Assuming that the knowledge stock has a positive coefficient implies that the effect of increased knowledge or innovative activity is to shift the demand curve outward by making the firm's products more attractive to its customers, at a given price.

Combining equations (7) and (8) as before, we obtain the following equation for revenue:

$$r_{it} = \left(\frac{\eta+1}{\eta}\right)\left(a_{it} + \alpha c_{it} + \beta l_{it}\right) + \left(\frac{\gamma(\eta+1)-\varphi}{\eta}\right)k_{it} \,. \tag{9}$$

This equation shows that the knowledge stock $K$ is likely to contribute to revenue and therefore to measured productivity growth via two channels: directly by increasing the efficiency of production and indirectly by shifting the demand curve for the firm's products outward (note that $\eta$ is negative so that $-\varphi/\eta$ is positive). It is usual to think of these two channels as process and product innovation.

For full identification of the system implied by equation (9), it would be desirable either to have data on individual firm output prices to allow a separate estimation of $\eta$ and $\varphi$ or to have some information on the components of $K$ that might be directed toward processes and/or products.[16] At the simplest level, one can gain some idea of the relative importance of the two types of innovation for productivity using the innovation dummy variables available from the various innovation surveys. One implication of the foregoing model is that process innovation will have ambiguous effects on

[16] Mairesse and Jaumandreu (2005) compare productivity estimates using revenue and output deflated at the firm level for France and Spain. They do not find any significant differences in the estimates, but they did not include R&D in the equation nor do they have true quality-adjusted price deflators. These two facts may account for the difference between their finding and that of Mairesse and Hall (1996) for the US.

revenue productivity, effects that depend on the firm's market power, whereas the effect of product innovation is likely to be positive.

In the studies reviewed here, the estimation of equation (9) is generally performed by regressing a measure of log revenue per employee $\left(r_{it} - l_{it}\right)$ on the logs of capital or investment, firm size measured in terms of employment, and various proxies for innovative activity. Industry dummies at the two-digit level are almost always included to control for things such as omitted inputs (in cases where value added is not available), differences in vertical integration, the omission of capital stocks (in cases where only current investment is available), and the overall level of technological knowledge. Although the model is in terms of the stock of knowledge or innovative capability, the usual proxies for this variable are the current level of innovative activity, measured as a dummy for some innovation during the past three years, or as the share of products sold that were introduced during the past three years. Because the estimation is almost always cross sectional, the fact that a flow of innovation rather than a stock is used will make little difference to the interpretation of the estimates, provided that innovation is persistent within firms. See Peters (2009) for evidence of this being the case.

## 4. The empirical evidence

Appendix Tables 1 and 2 summarize the studies which have attempted to explicitly estimate a quantitative relationship between firm-level productivity and innovation measures.[17] 25 papers are listed, of which all but two use data from the Community Innovation Survey (CIS) or its imitators in other countries. Of those using CIS-type data, 18 use some variant of the well-known CDM (Crepon, Duguet, and Mairesse) model for the analysis. One of these papers used both levels and growth rates to measure productivity (Loof and Heshmati, 2006), but most have chosen either levels (14 papers) or growth rates (10 papers) exclusively.

Use of the CDM model implies that most of the estimates are essentially cross-sectional ones that ignore issues of the timing of innovation and its contribution to productivity (exceptions are Masso and Vahter, 2008;

[17] The table ignores the large literature which studies R&D and productivity; see Hall et al. (2010) for a recent survey of this topic.

Belderbos et al., 2004; Peters, 2006). This is a reflection of the nature of the innovation surveys which ask about innovative behavior during the past three years and contain or are matched to other firm information that is contemporary with the innovation data. The data available is usually not sufficient to construct a time series (panel) for the firms involved since the samples are redrawn for each survey and there is little overlap.[18] Thus, the analysis usually relates productivity in one period to innovation in the same period or slightly before that period but it does not trace out any dynamic response. It is noteworthy that the results for the papers that do use lagged measures of innovation are not notably different from those using contemporary measures, reinforcing the cross-sectional and long-run nature of these results.

The CDM model has been described by many others in detail (see the references in appendix Tables 1 and 2) and I will only summarize it here. It generally consists of three sets of relationships, the first two of which can involve more than one equation. The first set of equations describes whether a firm undertakes R&D and, if so, how much, as a function of firm and industry characteristics. The second set describes the various types of innovation outcomes as a function of R&D intensity and other firm/industry characteristics. In many cases, the R&D variable in the innovation equations is computed as the expected R&D intensity, given the firm's characteristics. This procedure is grounded on the idea that many firms do informal R&D but do not report their spending separately to the statistical agency performing the survey. In a sense, the model fills in their R&D values with what might have been expected given their size, industry, nature of competition, etc. Looked at in another way, including the fitted value of R&D intensity for firms that actually report R&D is a form of instrumental variable estimation of the innovation equations, which helps correct for the simultaneity that might be present due to the fact that innovation is measured over the past three years, whereas R&D is frequently a current year measure.

The innovation equations in the CDM model can be probit equations for the probability of product, process, or organizational innovation or they

[18] For example, Criscuolo and Haskell (2003) report that there are 1 596 manufacturing firms in their CIS2 sample and 4 567 in their CIS3 sample, but only 509 appear in both surveys. Hall et al. (2008) have 9 462 firms in their sample drawn from three MCC surveys, but only 608 of these firms appear in all three surveys.

can also include an equation for the share of innovative sales (typically the sales share of products introduced during the past three years). In the latter case, the variable is sometimes transformed using logit transform which allows for infinite rather than finite support. That is, if $z$ is the share, ranging from 0 to 1, the logit transform $log(z/(1-z)) \epsilon (-\infty, +\infty)$ is used.[19] Following the logic used above, the predicted innovation probabilities or shares are then included in a productivity equation. The resulting estimates give the contribution of expected innovation conditional on R&D and other firm characteristics to productivity.

Tables 2a (levels, using innovative sales share), 2b (levels, using the product innovation dummy), and 3 (growth rates) summarize the results of estimating the productivity-innovation relationship from the papers listed in the appendix tables. I discuss each of these tables in turn. It should be noted that although I am treating the estimates as comparable, the precise regressions used in any particular paper will differ from those in other papers, as will the data construction itself. In addition, most researchers have included innovation variables that are predicted values from earlier regressions, as in the CDM model, while a few have included the actual innovation variables from the survey.

In spite of these variations, the results for the elasticity of output with respect to the innovative sales share (shown in Table 2a) are reasonably consistent across countries and time periods. The highest elasticities (0.23-0.29) are for knowledge-intensive or high-technology sectors. Most of the elasticities for Western Europe lie between 0.09 and 0.13, and less-developed countries, the service sector, and the low technology sectors have elasticities less than 0.09, with the exception of the weakly significant estimate for Chilean data. Thus, we can conclude that innovative sales are associated with revenue productivity, and that the association is stronger for higher technology sectors. For a typical Western European manufacturing firm, doubling the share of innovative sales will increase revenue productivity by about 11 percent.

[19] The alert reader will note that this expression is undefined for $z=0$ and $z=1$. Normally, this problem is solved by setting $z=0.01$ and $z=0.99$, respectively.

**Table 2a. Results for the productivity-innovation relationship in TFP levels (product innovation measured as innovative sales share)**

| Sample | Time period | Elasticity with respect to innov. sales share | Process innovation dummy |
|---|---|---|---|
| Chilean mfg sector | 1995-1998 | 0.18 (0.11)* | |
| Chinese R&D-doing mfg sector | 1995-1999 | 0.035 (0.002)*** | |
| Dutch mfg sector | 1994-1996 | 0.13 (0.03)*** | -1.3 (0.5)*** |
| Finnish mfg sector | 1994-1996 | 0.09 (0.06) | -0.03 (0.06) |
| French mfg sector | 1986-1990 | 0.07 (0.02)*** | |
| French Hi-tech mfg # | 1998-2000 | 0.23 (0.15)* | 0.06 (0.02)*** |
| French Low-tech mfg # | 1998-2000 | 0.05 (0.02)*** | 0.10 (0.04)*** |
| German K-intensive mfg sector | 1998-2000 | 0.27 (0.10)*** | -0.14 (0.07)** |
| Irish firms # | 2004-2008 | 0.11 (0.02)*** | 0.33 (0.08)*** |
| Norwegian mfg sector | 1995-1997 | 0.26 (0.06)*** | 0.01 (0.04) |
| Swedish K-intensive mfg sector | 1998-2000 | 0.29 (0.08)*** | -0.03 (0.12) |
| Swedish mfg sector | 1994-1996 | 0.15 (0.04)*** | -0.15 (0.04)*** |
| Swedish mfg sector | 1996-1998 | 0.12 (0.04)*** | -0.07 (0.03)*** |
| Swedish service sector | 1996-1998 | 0.09 (0.05)* | -0.07 (0.05) |

Source: Author's summary from Appendix Table 1.

Note: # Innovative sales share and process innovation included separately in the production function.

Table 2b presents the results of the productivity regression that uses a 0/1 measure of product innovation instead of the innovative sales share. For reasons mentioned earlier, this measure will vary by the size of the firm purely for measurement reasons and should be considered a much weaker proxy for innovative output. We do see that the results are more variable, although still positive for the most part. For manufacturing sectors in Western Europe, the typical values are around 0.05-0.10, implying that product innovating firms have an average productivity that is about 8 percent higher than non-innovators, but there is a wide dispersion.

The results for process innovation in both Tables 2a and 2b are even more variable, with some being negative, some zero, and some positive. Note that the few positive estimates in Table 2a are for the two cases where the authors included this variable alone in the productivity regression, without the innovative sales variable (Mairesse et al., 2005 for France and Siedschlag et al., 2010 for Ireland). The other positive estimates occur when product innovation is measured by a dummy rather than by the share of innovative sales, which suggests that they are partly due to the measurement error implicit in using a dummy to proxy for innovation. That is, we know from many of the surveys that process and product innovation go together. Therefore, if we have a weak measure of product innovation, we might expect the process innovation dummy to pick up more of the overall innovative activity. Recalling the discussion of equation (9), one could argue that the estimates in Table 2a, which are mostly negative for process innovation and positive for product innovation, suggest that firms are operating in the inelastic portion of their demand curves and that revenue productivity is enhanced mainly by the introduction of new and improved products, and not by efficiency improvements in the production process.[20]

**Table 2b. Results for the productivity-innovation relationship in TFP levels (product innovation measured as a dummy)**

| Sample | Time period | Product innovation dummy | Process innovation dummy |
|---|---|---|---|
| Argentinian mfg sector | 1998-2000 | -0.22 (0.15) | |
| Brazilian mfg sector | 1998-2000 | 0.22 (0.04*** | |
| Estonian mfg sector | 1998-2000 | 0.17 (0.08)** | -0.03 (0.09) |
| Estonian mfg sector | 2002-2004 | 0.03 (0.04) | 0.18 (0.05)*** |
| French mfg sector | 1998-2000 | 0.08 (0.03)** | |
| French mfg sector | 1998-2000 | 0.06 (0.02)*** | 0.07 (0.03)** |
| French mfg sector | 1998-2000 | 0.05 (0.09) | 0.41 (0.12)*** |
| French mfg sector | 2002-2004 | -0.08 (0.13) | 0.45 (0.16)*** |
| French service sector | 2002-2004 | 0.27 (0.52) | 0.27 (0.45) |
| German mfg sector | 1998-2000 | -0.05 (0.03) | 0.02 (0.05) |
| Irish firms # | 2004-2008 | 0.45 (0.08)*** | 0.33 (0.08)*** |
| Italian mfg sector | 1995-2003 | 0.69 (0.15)*** | -0.43 (0.13)*** |
| Italian mfg sector SMEs | 1995-2003 | 0.60 (0.09)*** | 0.19 (0.27) |
| Mexican mfg sector | 1998-2000 | 0.31 (0.09)** | |
| Spanish mfg sector | 2002-2004 | 0.16 (0.05)*** | |
| Spanish mfg sector | 1998-2000 | 0.18 (0.03)*** | -0.04 (0.04) |
| Swiss mfg sector | 1998-2000 | 0.06 (0.02)*** | |
| UK mfg sector | 1998-2000 | 0.06 (0.02)*** | 0.03 (0.04) |

Source: Author's summary from Appendix Table 1.

[20] The results surveyed here do not generally include the effects of organizational innovation, which has been shown to be associated with revenue productivity improvement, especially when accompanied by IT investment. However, in many cases, the data available on organizational innovation (a simple dummy variable) does not allow researchers to include this variable along with the other innovation variables in productivity regressions, due to the collinearity of the various innovation variables previously referred to.

**Table 3. Results for the productivity-innovation relationship in TFP growth rates**

| Sample | Time period | Elasticity wrt Innov sales share | Product innovation dummy | Process innovation dummy |
|---|---|---|---|---|
| Argentinian mfg sector | 1992-2001 | | 0.09 (0.08) | 0.18 (0.08)** |
| Dutch mfg sector | 1994-1998 | 0.009 0.001)*** | | -1.2 (0.7)* |
| Dutch mfg sector | 1996-1998 | 0.0002*** # | | |
| French mfg sector | 1986-1990 | | 0.022 (0.004)*** | |
| German mfg sector | 2000-2003 | 0.04 (0.02)** | | 0.14 (0.08)* @ |
| Italian mfg sector | 1992-1997 | | 0.12 (0.09) | 0.04 (0.12) |
| Spanish mfg sector | 1990-1998 | | 0.015 (0.004)*** | |
| Swedish mfg sector | 1996-1998 | 0.07 (0.03)** | | |
| Swedish service sector | 1996-1998 | 0.08 (0.03)*** | | |
| UK mfg sector | 1994-1996 | -0.02 (0.02) | | 0.02 (0.01)* |
| UK mfg sector | 1998-2000 | 0.07 (0.03)** | | -0.04 (0.02)** |

Source: Author's summary from Appendix Table 1.

Notes: # elasticity with respect to innovation expenditure per sales. @ elasticity with respect to cost reduction per employee.

Table 3 presents results for a productivity regression where the left-hand side is productivity growth, rather than its level. This relationship is not precisely the growth rate version of the regressions that lie behind Table 2, since it relates growth to the level of innovative activity, not to its growth rate. In general, the results are similar to but slightly lower than the level version of the equation, with an innovative sales elasticity focused on the range 0.04-0.08, and a product innovation dummy of about 0.02. As before, process innovation is negative when included with product innovation in the equation, although positive on its own. It is noteworthy that the only study with a true estimate of the cost savings due to process innovation rather than a dummy (Peters, 2006) yields a large and marginally significant elasticity of 0.14, implying that if we had better measures of process innovation, we might be able to considerably improve the measure of its impact.

From this summary of the empirical relationship between the various innovation measures and firm-level revenue productivity, we can conclude the following: first, there is a positive relationship, albeit somewhat noisy, between innovation in firms and their productivity, i.e. both the level and its growth. Second, the positive relationship is primarily due to product innovation. The impact of process innovation is more variable, and often negative. This can be interpreted in one of two ways: the typical firm enjoys some market power, but operates in the inelastic portion of its demand curve so that revenue productivity falls when it becomes more efficient. Alternatively, it is possible that there is so much measurement error in the innovation variables that only one of the two is positive and significant when entered in the productivity equation. Without instruments that are better targeted to predicting the two different kinds of innovation, this possibility cannot be ruled out.

## 5. Conclusions

The foregoing survey of empirical evidence on the relationship between innovation and productivity finds an economically significant impact of product innovation on revenue productivity and a somewhat more ambiguous impact of process innovation. As I have argued, the latter result is primarily due to the fact that we are not able to measure the real quantity effect of process innovation, which is the relevant quantity for social welfare. We can only measure the real revenue effect, which combines the impact of innovation on both quantity and price. So overall, we can conclude that in spite of the fact that innovative activity is not very well measured in many cases, it does generally increase an individual firm's ability to derive revenue from its inputs.

Naturally, this conclusion leads to new questions. What are the factors in the firm's environment that encourage such innovative activity? And how is aggregate productivity influenced by the innovative activities of individual firms? Although it is beyond the scope of this paper to answer these questions, some promising avenues to explore have recently been suggested in the literature. Taking the second question first, the approach of Foster et al. (2008), although intensive in its data requirements, has yielded interesting insights into the relative importance of productivity growth in existing firms and net entry in aggregate productivity growth. In addition, these authors perform a detailed analysis of the differences between revenue productivity growth and "physical" productivity growth, making the same distinction between efficiency and demand effects that I have made in this survey. They find that the use of revenue productivity will tend to understate the contribution of entrants to productivity growth, and that demand variation is a more important determinant of firm survival than efficiency in production.

A very interesting line of work would be to understand the extent to which innovative activity on the part of entrants and the existing firms is behind the results in Foster et al. (2008). That is, the paper provides evidence on the composition of aggregate productivity growth but not on its sources. Aghion et al. (2009) find that foreign firm entry in technologically advanced UK sectors spurs both innovation (measured as patents) and productivity growth, whereas entry by such firms in lagging sectors reduces innovation and productivity growth by domestic firms in those sectors, arguing that this is due to the fact that firms are discouraged by the cost of catching up. On the other hand, Gorodnichenko et al. (2010), using data from emerging market countries in Eastern Europe and the former Soviet Union, find a robust relationship between foreign competition (self-reported by the firms) and innovation in all sectors, including the service sector. Thus, we have evidence that at least some kinds of entry encourage innovative activity, although relatively little that traces the path from entry to innovation and then to productivity.

As to the regulatory and financial environment that encourages innovation on the part of firms, following important efforts led by the World Bank to collect data on entry regulation, the rule of law, and other country characteristics, a substantial cross-country growth literature has developed that relates these characteristics to entry (Djankov et al., 2002; Aidis et al., 2009; Ciccone and Papaioannou, 2006), investment (Alesina et al., 2003), productivity (Cole et al., 2005), and firm size and growth (Fisman and Sarria-Allende, 2004; Klapper et al., 2006). Briefly summarized, stronger entry regulation and/or higher entry costs are associated with fewer new firms, greater existing firm size and growth, lower TFP, less investment, and higher profits.[21] Most of the studies cited have made a serious attempt to find instruments or controls which allow them to argue that this relationship is causal. Thus far, none of these studies explicitly looks at the impact on innovative activity and its relationship with productivity, although one can argue that the entry of new firms is a form of innovation. Getting a full picture of the macro-economy that incorporates firm entry and exit, innovation, and the resulting productivity growth, a picture that would allow one to clearly understand the use of various policy levers, is a goal not yet achieved in the literature.

One avenue that looks promising is the work of Bartelsman et al. (2009) who extended Foster et al. (2008) to look at the allocative efficiency of entry and exit by firms to data on firms in the US and seven European countries. They develop a relative diagnostic measure of inefficient allocation of resources across firms based on the covariance of firm size and productivity within industry. The idea of this measure is that economies that are subject to inefficient regulation that prevents firms from growing or shrinking to their optimal size will display a lower correlation between firm size and productivity, since more productive firms will not be able to grow and displace less productive firms. They show that this measure changed in the way one would expect in three East European countries between the early 1990's and the 2000's. However, in spite of its promise for analyzing the sources of aggregate productivity growth, this kind of work has formidable data requirements. It also does not yet incorporate any measure of innovation as a causal measure, but it seems that extending this approach might be useful for exploring the simultaneous relationship between innovation, regulation, and productivity.

## 6. Policy implications

I close this survey with a few thoughts on what these results might mean for policy directed towards improving the productivity performance of European firms. The empirical results surveyed, which cover a large number of countries, mostly in Europe, do not suggest that firms are "underperforming". Innovation in European countries is as high or higher than it is in

[21] See Djankov (2009) for a recent survey of this literature.

those few non-European countries with which comparison is possible (notably Japan, and perhaps the United States), and it translates into productivity improvements in the way one might have expected. Without measures of innovation expenditures, it is not possible to compute rates of return, but innovative sales elasticities in the range of 0.09 to 0.13 are reasonable when compared to what we know about R&D elasticities in the production function.

A second implication of the results here is that because process innovation can increase real output while leaving revenue mostly unchanged at the individual firm level, evaluating the overall impact of process innovation requires consideration of its impact on price as well as quantity. In addition, one of the main consequences of innovation is likely to be the exit of some (inefficient) firms and the entry of new innovating firms, which implies that studying overall productivity impacts requires an examination of aggregate data as well as the micro evidence surveyed here. Taken together with the mostly good innovation-productivity performance of the individual firms studied here, this implication suggests that policymakers direct their attention to the extent to which entry and exit regulation impacts the rationalization of industry structure in response to innovative activity.

One final policy implication concerns the nature of the data derived from innovation surveys. Because of the inherent imprecision of a dummy variable for innovation and the problem associated with using this variable across a large size range of firms, researchers and policy makers should strongly resist using the innovation dummies to say anything analytical about innovation and firm size. For this purpose, the share of sales that are products new to the firm is a much better indicator, since it removes the scale problem from the data.

## References

Acs, Z.J., and Audretsch, D.B. (1990), Innovation and Small Firms, MIT Press, Cambridge MA.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. and Prantl, S. (2009), The effects of entry on incumbent innovation and productivity, Review of Economics and Statistics 91, 20-32.

Aidis, R., Estrin, S. and Mickiewicz, T. (2009), Entrepreneurial entry: Which institutions matter?, CEPR Discussion Paper 7278.

Alesina, A., Ardagna, S., Nicoletti, G. and Schiantarelli, F. (2003), Regulation and investment, NBER Working Paper 9560.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2009), Cross country differences in productivity: The role of allocation and selection, NBER Working Paper 15490.

Belderbos, R., Carree, M. and Lokshin, B. (2004), Cooperative R&D and firm performance, Research Policy 33, 1477-1492.

Benavente H.J.M. (2006), The role of research and innovation in promoting productivity in Chile, Economics of Innovation and New Technology 15, 301-315.

Blundell, R. and Bond, S.R. (2000), GMM estimation with persistent panel data: An application to production functions, Econometric Reviews 19, 321-340.

Chudnovsky, D., Lopez, A. and Pupato, G. (2006), Innovation and productivity in developing countries: A study of Argentine manufacturing firms' behavior (1992-2001), Research Policy 35, 266-288.

Ciccone, A. and Papaioannou, E. (2006), Red tape and delayed entry, CEPR Discussion Paper 5996.

Cole, H.L., Ohanian, L.E., Riascos, A. and Schmitz, Jr, J.A. (2005), Latin America in the rear view mirror, Journal of Monetary Economics 52, 69-107.

Crepon, B., Duguet, E. and Mairesse, J. (1998), Research, innovation, and productivity: An econometric analysis at the firm level, Economics of Innovation and New Technology 7, 115-156.

Criscuolo, C. and Haskel, J.E. (2003), Innovation and productivity growth in the UK: Evidence from CIS2 and CIS3, CeRiBA Working Paper, LSE, London.

Denison, E.F. (1966), Discussion of Jorgenson and Griliches, American Economic Review 56, 76-82.

Djankov, S. (2009). The regulation of entry: A survey, World Bank Research Observer 24, 183-203.

Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2002), The regulation of entry, Quarterly Journal of Economics 117, 1-37.

Duguet, E. (2006), Innovation height, spillovers and TFP growth at the firm level: Evidence from French manufacturing for company performance, Economics of Innovation and New Technology 15, 415-442.

Fisman, R. and Sarria-Allende, V. (2004), Regulation of entry and the distortion of industrial organization, NBER Working Paper 10929.

Foster, L., Haltiwanger, J. and Syverson, C. (2008), Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?, American Economic Review 98, 394-425.

Freeman, C. and Soete, L. (1997), The Economics of Industrial Innovation, Pinter, London.

Geroski, P.A. (1989), Entry, innovation and productivity growth, Review of Economics and Statistics 71, 572-578.

Gorodnichenko, Y., Svejnar, J. and Terrell, K. (2010), Globalization and innovation in emerging markets, American Economic Journal: Macroeconomics 2, 194-226.

Griffith, R., Huergo, E., Mairesse, J. and Peters, B. (2006 ), Innovation and productivity in four European countries, Oxford Review of Economic Policy 22, 483-498.

Griliches, Z. (1994), Productivity, R&D and the data constraint, American Economic Review 84, 1-23.

Griliches, Z. (1996), The discovery of the residual: A historical note, Journal of Economic Literature 34, 1324-1330.

Griliches, Z. (1998), R&D and Productivity: The Econometric Evidence, Chicago University Press, Chicago IL.

Griliches, Z. and Mairesse, J. (1984). Productivity and R&D at the firm level, in Z. Griliches (ed.), R&D, Patents, and Productivity, Chicago University Press, Chicago IL.

Hall, B.H. (1996), The private and social returns to research and development, in B.L.R. Smith and C.E. Barfield (eds.), Technology, R&D, and the Economy, Brookings Institution and American Enterprise Institute, Washington DC.

Hall, B.H., Jaffe, A.B. and Trajtenberg, M. (2005), Market value and patent citations, Rand Journal of Economics 36, 16-38.

Hall, B.H., Lotti, F. and Mairesse, J. (2008), R&D, innovation, and productivity: New evidence from Italian manufacturing microdata, Industrial and Corporate Change 17, 813-839.

Hall, B.H., Lotti, F. and Mairesse, J. (2011), Evidence on the impact of R&D and ICT investment on innovation and productivity in Italian firms, revised version of paper presented at the Concord 2010 Conference, Seville, Spain, March 2010.

Hall, B.H., Mairesse, J. and Mohnen, P. (2010), Measuring the Returns to R&D, in B.H. Hall and N. Rosenberg, Handbook of the Economics of Innovation, Elsevier, Amsterdam and New York.

Hall, R.E. (1988), The relationship between price and marginal cost in U.S. Industry, Journal of Political Economy 96, 921-947.

Harhoff, D., Narin, F., Scherer, F.M. and Vopel, K. (1999), Citation frequency and the value of patented inventions, Review of Economics and Statistics 81, 511-515.

Huergo, E. and Jaumandreu, J. (2004), Firms' age, process innovation and productivity growth, International Journal of Industrial Organization 22, 541-559.

Janz, N., Loof, H. and Peters, B. (2003), Firm level innovation and productivity – Is there a common story across countries?, ZEW Discussion Paper 03-26, Mannheim.

Jefferson, G.H., Bai, H., Guan, X. and Yu, X. (2006), R&D performance in Chinese industry, Economics of Innovation and New Technology 15, 345-366.

Jorgenson, D.W. and Griliches, Z. (1967), The explanation of productivity change, Review of Economic Studies 34, 249-283.

Klapper, L., Laeven, L. and Rajan, R. (2006), Entry regulation as a barrier to entrepreneurship, Journal of Financial Economics 82, 591-629.

van Leeuwen, G. (2002), Linking innovation to productivity growth using two waves of the community innovation survey, OECD Science, Technology and Industry Working Papers 2002/8, Paris.

van Leeuwen, G. and Klomp, L. (2006), On the contribution of innovation to multifactor productivity growth, Economics of Innovation and New Technology 15, 367-390.

Loof, H. and Heshmati, A. (2006), On the relationship between innovation and performance: A sensitivity analysis, Economics of Innovation and New Technology 15, 317-344.

Loof, H., Heshmati, A., Asplund, R. and Nåås, S.-O. (2001), Innovation and performance in manufacturing industries: A comparison of Nordic countries, SSE/EFI Working Paper Series in Economics and Finance 457, Stockholm.

Mairesse, J. and Hall, B.H. (1996), Estimating the productivity of research and development in French and US manufacturing firms: An exploration of simultaneity issues with GMM methods, in K. Wagner and B. van Ark (eds), International Productivity Differences and Their Explanations, Elsevier-North Holland, Amsterdam.

Mairesse, J. and Jaumandreu, J. (2005), Panel-data estimates of the production function and the revenue function: What difference does it make?, Scandinavian Journal of Economics 107, 651-672.

Mairesse, J. and Mohnen, P. (2010), Using innovation surveys for econometric analysis, in B.H. Hall and N. Rosenberg (eds.), Handbook of the Economics of Innovation, Vol II, Elsevier, Amsterdam and New York.

Mairesse, J., Mohnen, P. and Kremp, E. (2005), The importance of R&D and innovation for productivity: A reexamination in light of the 2000 French innovation survey, i J. Mairesse and M. Trajtenberg (eds.), Special issue of Annales d'Economie et de Statistique 79-80 in Memory of Zvi Griliches.

Mairesse, J. and Robin, S. (2010), Innovation and productivity: A firm-level analysis for French manufacturing and services using CIS3 and CIS4 data (1998-2000 and 2002-2004), CREST-ENSAE, Paris.

Masso, J. and Vahter, P. (2008), Technological innovation and productivity in late-transition Estonia: Econometric evidence from innovation surveys, European Journal of Development Research 20, 240-261.

OECD/OCDE (2003), The Sources of Economic Growth in OECD Countries, OECD, Paris.

OECD/OCDE (2005), Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, 3rd edition, OECD, Paris.

Olley, G.S. and Pakes, A. (1996), The dynamics of productivity in the telecommunications equipment industry, Econometrica 64, 1263-1297.

Parisi, M.L., Schiantarelli, F. and Sembenelli, A. (2006), Productivity, innovation and R&D: Micro evidence for Italy, European Economic Review 50, 2037-2061.

Pavitt, K., Robson, M. and Townsend, J. (1987), The size distribution of innovating firms in the UK: 1945-1983, Journal of Industrial Economics 35, 297-316.

Peters, B. (2006), Productivity effects of innovation activities, in B. Peters, Innovation and Firm Performance: An Empirical Investigation for German Firms, PhD thesis, University of Würzburg.

Peters, B. (2009), Persistence of innovation: Stylised facts and panel data evidence, Journal of Technology Transfer 34, 226-243.

Polder, M., van Leeuwen, G., Mohnen, P. and Raymond, W. (2009), Productivity effects of innovation modes, Statistics Netherlands Discussion Paper 09033, The Hague.

Raffo, J., l'Huillery, S. and Miotti, L. (2008), Northern and southern innovativity: A comparison across European and Latin American countries, European Journal of Development Research 20, 219-239.

Scherer, F.M. and Harhoff, D. (2000), Technology policy for a world of skew-distributed outcomes, Research Policy 29, 559-566.

Siedschlag, I., Zhang, Z. and Cahill, B. (2010), The effects of the internationalisation of firms on innovation and productivity, ESRI Working Paper 363, Economic and Social Research Institute, Dublin.

Sterlacchini, A. (1989), R&D, innovations, and total factor productivity growth in British manufacturing, Applied Economics 21, 1549-1562.
Tanaka, N., Glaude, M. and Gault, F. (2005), Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, OECD and Eurostat, Paris.
US National Science Foundation (2010), NSF releases new statistics on business innovation, SRS Info Brief NSF 11-300, Washington DC.

## Appendix Table 1. Empirical studies of the productivity-innovation relationship using productivity levels

| Authors (year) | Country | Observations | Method* | Output measure | Innov measure | Estimated impact of innovation | Comments |
|---|---|---|---|---|---|---|---|
| Benavente (2006) | Chile | 1995-1998<br>438 mfg plants | CDM model: ALS | Log VA per emp | Log innov sales share | 0.18 (0.11)* | SR prod not related to innovation or R&D, but related to engineers & admin (higher salaries); innovation due to capital, not in productivity. |
| Crepon et al. (1998) | France | SESSI 1986-1990<br>~5000 innov mfg firms | CDM model: ALS | Log VA per emp | Log innov sales share | 0.065 (0.015)*** | Positive impact of innovation sales share on productivity, as well as positive association of productivity with human capital in labor force. |
| Griffith et al. (2006) | France, Germany, Spain, UK | CIS3 1998-2000<br>FR 3625 mfg firms<br>DE 1123 mfg firms<br>ES 3588 mfgfirms<br>UK 1904 mfg firms | CDM model: sequential with IV | Log sales per emp | Product and process dummies | FR: 0.07 (0.03)** proc<br>0.06 (0.02)*** prod<br>DE: 0.02 (0.05) proc<br>-0.05 (0.03) prod<br>ES: -0.04 (0.04) proc<br>0.18 (0.03)*** prod<br>UK: 0.03 (0.04) proc<br>0.06 (0.02)*** prod | Estimation in 3 steps, no bivariate probit. Process innovation adds 0.07 in France, nothing in other countries; Product innovation positive except in Germany. |
| Hall et al. (2011) | Italy | MCC 1992-2003<br>14294 mfg firms | CDM with 4 types of innovation: FIML for selection; quadrivariate probit; IV | Log sales per emp | 4 innov dummies | prod: 0.69 (0.15)***<br>proc: -0.43 (0.13)*** | Innovation variables not separately well-identified in 196roductiveity equation; process appears to be negative and product positive for TFP. |

**Appendix Table 1. Continued….**

| Authors (year) | Country | Observations | Method* | Output measure | Innov measure | Estimated impact of innovation | Comments |
|---|---|---|---|---|---|---|---|
| Janz et al. (2003) | Germany Sweden | CIS3 1998-2000<br>1000 K-intensive mfg firms | CDM model: sequential with IV | Log sales per emp | Log innov sales per emp, process dummy | DE: 0.27 (0.10)*** prod<br>-0.14 (0.07)** proc<br>SE: 0.29 (0.08)*** prod<br>-0.03 (0.12) proc | Allowed for feedback from productivity to innovation output. Elasticity of 197roduc-tiveity wrt innov sales similar in both countries. |
| Jefferson et al. (2006) | China | 1995-1999<br>5500 R&D-doing large/medium sized firms | CDM model: sequential with IV | Log sales per emp | Log innov sales share | 0.035 (0.002)*** | No correction for innovation selection bias. |
| Loof and Heshmati (2006) | Sweden | CIS3 1996-1998<br>1071 mfg firms<br>718 service firms<br>92 utility firms | CDM variation: FIML on selection submodel; 3SLS; sensitiv-ity analysis | Log VA per emp | Log innov sales per emp, process dummy | prod: 0.12 (0.04)*** mfg<br>0.09 (0.05)** service<br>proc: -0.07 (0.03)*** mfg<br>-0.07 (0.05) service | Survey data less reliable than register data; sales not as good as VA in productivity eq |
| Loof et al. (2001) | Finland, Norway, Sweden | CIS2 1994-1996 (1995-1997 in Norway)<br>NO: 485 mfg firms<br>FI: 323 mfg firms<br>SE: 407 mfg firms | CDM variation: sequential with 3SLS | Log sales per emp | Log innov sales per emp, process dummy | FI: 0.090 (0.058) prod<br>-0.029 (0.060) proc<br>NO: 0.257 (0.062)*** prod<br>0.008 (0.044) proc<br>SE: 0.148 (0.044)*** prod<br>-0.148 (0.043)*** proc | Allows for simultaneity btwn innovation & output - feedback in NO but not FI and SE. Elasticity slightly higher for radical innovations. |

**Appendix Table 1. Continued….**

| Authors (year) | Country | Observations | Method* | Output measure | Innov measure | Estimated impact of innovation | Comments |
|---|---|---|---|---|---|---|---|
| Mairesse and Robin (2010) | France | CIS3 1998-2000<br>3500 mfg firms<br>CIS4 2002-2004<br>5000 mfg firms<br>3600 service firms | CDM model: FIML for selection eqs; bivariate probit; IV | Log VA per emp | Product and process dummies | mfg 98-00:<br>0.41 (0.12)*** proc<br>0.05 (0.09) prod<br>mfg 02-04:<br>0.45 (0.16)*** proc<br>-0.08 (0.13) prod<br>service: 0.27 (0.45) proc<br>0.27 (0.52) prod | Estimation is in 3 steps, but also in 2 steps, with innov & labor productivity equations combined. Process innovation enters productivity, but not product. Explores using a single innovation indicator, which works just as well. |
| Mairesse et al. (2005) | France | CIS3 1998-2000<br>2200 mfg firms | CDM & variations | Log VA per emp | Logit transform of innov sales share, process dummy, other dummies - all separately | HT: 0.23 (0.15)*<br>0.07 (0.03)*** radical<br>0.06 (0.02)*** process<br>LT: 0.05 (0.02)***<br>-0.08 (0.05)* radical<br>0.10 (0.04)*** process | TFP using output; going through innovation does not add much to estimates of return to R&D, after correcting for selectivity and endogenei-ty; endogeneity correction impt for innov variables. |
| Masso & Vahter (2008) | Estonia | CIS3 1998-2000<br>1467 mfg firms<br>CIS4 2002-2004<br>992 mfg firms | CDM variation: sequential with bivariate probit for innov | Log VA per emp | Product and process dummies (org dummies in 2nd period) | prod 98-00: 0.21 (0.08)***<br>02-04: 0.00 (0.05)<br>proc 98-00: -0.06 (0.10)<br>02-04: 0.15 (0.06)*** | Uses innov expenditure rather than R&D; proc & prod dummies; prod innovation increases productivity in recession; proc innovation in growth period. One and two year lag effects are roughly the same (cross sectional). |

**Appendix Table 1. Continued….**

| Authors (year) | Country | Observations | Method* | Output measure | Innov measure | Estimated impact of innovation | Comments |
|---|---|---|---|---|---|---|---|
| Masso and Vahter (2008) | Estonia | CIS3 1998-2000<br>1467 mfg firms<br>CIS4 2002-2004<br>992 mfg firms | CDM variation: sequential with bivariate probit for innov | Log sales per emp | Product and process dummies (org dummies in 2nd period) | prod 98-00: 0.17 (0.08)**<br>02-04: 0.03 (0.04)<br>proc 98-00: -0.03 (0.09)<br>02-04: 0.18 (0.05)*** | uses innov expenditure rather than R&D; proc & prod dummies; prod innovation increases productivity in recession; proc innovation in growth period. One and two year lag effects are roughly the same (cross sectional). |
| Roper | Netherlands | CIS 3.5-4.5 2002-2006<br>~1200 mfg & service firms | augmented CDM | Log VA per emp | 3 innov dummies (proc prod org) in combo | mfg:<br>1.7 (0.4)*** org alone<br>1.0 (0.5)** org & proc<br>0.9 (0.2)*** all<br>serv:<br>4.3 (0.5)*** org alone<br>17.1 (2.2)*** org & proc<br>-8.3 (1.3)*** proc & prod<br>3.9 (0.5)*** all | Org innovation has strongest TFP effects. Process and product, only when combined with org innovation. However, signs of coefficient instability due to correlation of 8 combi-nations when predicted. |
| Raffo et al. (2008) | France, Spain, Switzerland, Argentina, Brazil, Mexico | CIS3 1998-2001 mfg<br>AR 1308 firms<br>BR 9452 firms<br>MX 1515 firms<br>FR 4618 firms<br>CH 925 firms<br>ES 3559 firms<br>(2002-04) | CDM model: sequential with IV | Log sales per emp | product & organizational innov dummies | AR: -0.22 (0.15)<br>BR: 0.22 (0.04)***<br>MX: 0.31 (0.09)***<br>FR: 0.08 (0.03)**<br>ES: 0.16 (0.05)***<br>CH: 0.10 (0.06)* | Interaction of innovative activities with national systems weaker in developing coun-tries. Foreign and domestic subs are uniformly more productive, but do more R&D only in France and Brazil. |

**Appendix Table 1. Continued….**

| Authors (year) | Country | Observations | Method* | Output measure | Innov measure | Estimated impact of innovation | Comments |
|---|---|---|---|---|---|---|---|
| van Leeuwen and Klomp (2006) | Nether-lands | CIS2 1994-1996<br>1400 innov firms | CDM variation: 3SLS | Log sales per emp | Process dummy; innov sales share | prod: 0.13 (0.03)***<br>proc: -1.3 (0.5)*** | Includes market share eq; feedback from sales to innovation; revenue function approach better than VA prod function framework (innov sales do not enter VA function in the presence of R&D and markup coefficients). |
| Siedschlag et al. (2010) | Ireland | CIS3 2004-2006<br>CIS4 2006-2008<br>723 firms (balanced panel) | CDM variation: sequential with IV | Log sales per emp | Product, process, and organizational dummies, innov sales share - all separately | innov sales: 0.11 (0.02)***<br>prod: 0.45 (0.08)***<br>proc: 0.33 (0.08)*** | Uses innovation expenditure instead of R&D spending; includes FDI and foreign ownership characteristics. |

Source: Author's collection, supplemented by Table A.1 (Chudnovsky et al., 2006), Table 4.1 (Peters, 2006).

Notes: * CDM = Crepon, Duguet, Mairesse model described in text. ALS = asymptotic least squares on multi-equation model. 3SLS = three stage least squares. FIML = full information maximum likelihood on multivariate normal model. OLS = ordinary least squares. IV = instrumental variable estimation.

## Appendix Table 2. Empirical studies of the productivity-innovation relationship using productivity growth

| Authors (year) | Country | Observations | Method | Output measure | Innov meas-ure | Estimated effect | Result |
|---|---|---|---|---|---|---|---|
| Belderbos et al. (2004) | Nether-lands | CIS2, CIS3 1996-1998<br>2056 mfg firms | Productivity eq only | Log VA per emp | Innov exp per sales | elasticity ~ 0.0002 (0.00003)*** | Productivity and innov sales share on lagged innovative activity and various kinds of cooperation. |
| Chudnovsky et al. (2006) | Argentina | INDEC-SECYT 1992-1996<br>INDEC-SECYT 1998-2001<br>718 mfg firms in a panel | CDM variation: sequential estimation with FE | Log sales per emp | product and process dummies; interactions | prod only: 0.09 (0.08)<br>proc only: 0.18 (0.08)**<br>both: 0.14 (0.06)**<br>any: 0.13 (0.05)*** | Uses innov expend rather than R&D; fixed effect single eq estimation. Uses logit for prod/proc/both innovation dummies. R&D increases prob of prod innov; Tech acquisition increases prob of both. |
| Criscuolo and Haskel (2003) | UK | CIS2, 3 1994-2000<br>5000 mfg firms | single eq regression for TFP growth: OLS | TFP growth (not clear if sales or VA) | Process dummy; share of innov sales | proc 94-96: 0.016 (0.009)*<br>proc 98-00: -0.038 (0.019)**<br>prod 94-96: -0.022 (0.017)<br>prod 98-00: 0.065 (0.033)** | Process innovation lead to TFP growth but with substan-tial lag; novel process innova-tions negative at first. |
| Duguet (2006) | France | SESSI 1986-1990<br>~5000 mfg firms | TFP growth reg with latent innov or dummies (GMM) | Log VA per hour | dummies for radical & incremental innovation | 0.022 (0.004)*** radical<br>-0.01 (0.01) incremental | Only radical innovations affect TFP growth, with a coefficient of 0.02. Latent innovation does not enter. |
| Geroski (1989) | UK | 1976-1979<br>79 industries | panel reg (CRS) | Log output per capital | # ind innov (flow) during past 3 yrs | 0.025 (0.010)** | Distributed lag of innovation counts more important than entry for TFP. |

**Appendix Table 2. Continued….**

| Authors (year) | Country | Observations | Method | Output measure | Innov measure | Estimated effect | Result |
|---|---|---|---|---|---|---|---|
| Huergo and Jaumandreu (2004) | Spain | 1990-1998<br>2300 mfg firms | semiparametric estimate of TFP | TFP growth | process innov dummy | 0.015 (0.004)*** all<br>0 uncensored (innovators) | Process innovation leads to TFP growth immediately, then declines slowly over time. Primary interest is age distribution of investment returns. |
| Loof and Heshmati (2006) | Sweden | CIS3 1996-1998<br>~3000 mfg, service, + utility firms | sensitivity analysis using CDM model | Log VA per emp | Log innov sales per emp | 0.07 (0.03)** mfg<br>0.08 (0.03)*** service | Mfg, prod level - 0.12 elasticity with innov sales<br>higher for profits, lower for services<br>mfg, prod growth - elasticity 0.07 wrt innov sales<br>higher for profits; and for services<br>survey data less reliable than register data; sales not as good as VA. |
| Parisi et al. (2006) | Italy | MCC 1992-1997<br>465 mfg firms in both surveys | TFP growth regressions: IV | Log sales per emp | Product and process dummies | prod: 0.12 (0.09)<br>proc: 0.04 (0.12) | Process innovations add to prod growth; product innovations do not enter. R&D elasticity is 0.04. R&D enters product innovation but not process. |
| Peters (2006) | Germany | MIP 2000-2003<br>522 mfg innov firms | CDM variation: sequential estimation | Log sales per emp | Log innov sales per emp; Log cost reduction per emp | prod: 0.04 (0.02)**<br>proc: 0.14 (0.08)* | Uses survey estimates of cost savings due to process innovation as well as innovative sales share; lag between innovation and productivity growth. |

**Appendix Table 2. Continued….**

| Authors (year) | Country | Observations | Method | Output measure | Innov meas-ure | Estimated effect | Result |
|---|---|---|---|---|---|---|---|
| Sterlacchini (1989) | UK | 1954-1984<br>15 mfg inds | cross sections for 6-year periods | TFP growth averaged over 6 years | # ind innov produced; # ind innov used | 0.08 (0.04)** inn produced<br>0.07-0.30 innov used | Correlates R&D and SPRU innovations by industry of origin and use - ranking same. Prior to 73, ind of more impt for TFP. After, correlation btwn R&D growth and TFP, proba-bly due to simultaneity. In 1980's, relationship btwn R&D/innov & TFP breaks down. |
| van Leeuwen (2002) | Nether-lands | CIS2,3 1994-1998<br>1929 mfg innov firms<br>510 mfg innov firms pooled | CDM variation: FIML on submodels for selection | Log sales per emp | share of innov sales; process dummy | prod dyn: 0.006 (0.004)*<br>prod static: 0.009 (0.001)***<br>proc dyn: -1.2 (0.7)*<br>proc static: -0.20 (0.50) | Uses Griliches-Mairesse 1984 to connect revenue to knowledge stock via demand equation; also includes process innovation dummy. Estimation is both static (pooled across periods) and dynamic (second period only). |

Source*:* Author's collection, supplemented by Table A.1 (Chudnovsky et al., 2006), Table 4.1 (Peters, 2006).

Notes: * CDM = Crepon, Duguet, Mairesse model described in text. ALS = asymptotic least squares on multi-equation model. 3SLS = three stage least squares. FIML = full information maximum likelihood on multivariate normal model. OLS = ordinary least squares. IV = instrumental variable estimation.

# Comment on Hall: Innovation and productivity

Ari Hyytinen*

Bronwyn Hall's review essay focuses on the pros and cons of different (survey-based) measures of firms' innovative activities and the ways in which the relation between innovation and productivity has been analyzed in the recent economics literature.

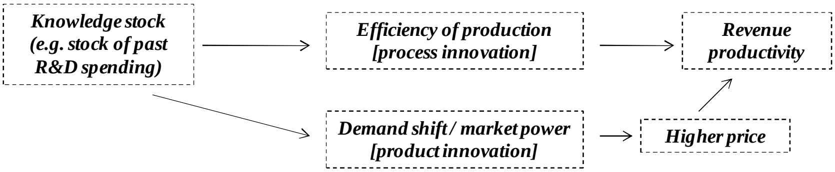
The essay is a careful review of a well-defined research question. It addresses all the important questions which a primer to the field or a policy-maker ought to ask: Why should we care? What are we trying to measure and what can we measure? How should we model the link between innovation and productivity? And, finally, what do we know? What are the policy implications of the findings in this literature?

Figure 1 provides a simplified summary of the economic set-up of the review. It illustrates the two main mechanisms via which the (formal and informal) research and development (R&D) efforts of firms may lead to process and product innovations and thereby to higher (revenue) productivity.

As is carefully documented in the review, most of the papers in the field use (cross-sectional) data from the Community Innovation Surveys (CIS) or its close cousins, with many analyses relying on the econometric model originally introduced by Crepon et al. (1998). The review concludes that despite the fact that innovative activity can often only be measured imprecisely, there is a clear positive relation between firms' innovative activity and their productivity. This positive relation can primarily be attributed to product innovations.

* University of Jyväskylä, ari.hyytinen@econ.jyu.fi.

**Figure 1. Innovations and (revenue) productivity**




The review is a very nice read, but both primers to the field and policy-makers ought to carefully note its limited scope: The review essay is not self-contained in the sense that it only covers a subset of the issues relevant for understanding productivity. Therefore, it should be read together with other surveys of the literature on the determinants of productivity. Relevant related surveys have recently been written both by Bronwyn Hall together with her co-authors (e.g., Hall et al., 2010) as well as by other scholars in the field (e.g., Syverson, 2011).[1]

Given the carefully defined scope of the review, I limit my comments to three viewpoints. The viewpoints I take are those of potential "consumers" of Hall's review essay (and this journal), i.e. researchers planning to enter the field, policy-makers considering appropriate interventions and statistical agencies collecting data on the innovative activities of firms.

Viewpoint 1: The review could perhaps be more explicit in its recommendations for a potential researcher in the field. Such a reader would clearly benefit if the review took a more systematic approach to the most important problems and some recent advances in the literature.

Are the problems in the current literature, such as a strong reliance on cross-sectional data and quite imperfect measures of innovations, serious enough to cast some doubt on the reliability of the empirical findings? Are the (econometric) identification strategies in the literature credible? In which direction might the results be biased? Given her wide experience of these issues, Hall's views and conjectures would no doubt benefit those who plan to begin research efforts in this field.

The framework presented in the review is perhaps a bit too simplified to put the recent literature into a proper context. For example, it does not allow for an assessment of the relative importance of firms' innovative

[1] See also Hall and Lerner (2010) for a survey of financing of innovation and Mairesse and Mohnen (2010) on further insights about the use of innovation surveys.

activities as a determinant of productivity. Syverson (2011) provides a wider perspective on this and is therefore worth taking a look at. Moreover, for readers from small open (Nordic) economies, some recent advances in the literature, such as Cassiman et al. (2011) emphasizing the triangle between innovations, exports and productivity, may be of interest.

Viewpoint 2: The review is quite brief in its policy discussion. It leaves open what the appropriate role for public policy is in this particular context. For example, a busy policy-maker should not rush to conclude from the findings of the review that product innovations should be supported more than process innovations, or vice versa. The busy policy-maker ought to be reminded that unfortunately, it is unclear whether the literature using innovation survey data makes it possible to draw any conclusions about the need for policy-interventions or the type of reforms that would be most effective.

Viewpoint 3: Given the potential readership of the *Nordic Economic Policy Review*, one might wonder if the findings in Hall's review essay have implications for statistical agencies (and their masters). What are the most important research questions that have not been addressed because of lack of appropriate data? Do we need better measures of product and process innovations, completely new innovation variables, or something else? The review could be more explicit about these questions. To me, Hall's review suggests that being able to study the dynamics between productivity and innovative activities and the intensive and extensive margins of firms' heterogeneous innovative activities would be a top priority for the researcher community. Such analyses call for panel data and better continuous measures of firms' various innovative activities.

## References

Cassiman, B., Golovko, E. and Martinez-Ros, E. (2011), Innovation, exports and productivity, International Journal of Industrial Organization 28, 372-376.

Crepon, B., Duguet, E. and Mairesse, J. (1998), Research, innovation and productivity: An econometric analysis at the firm level, Economics of Innovation and New Technology 7, 115-156.

Hall, B.H. and Lerner, J. (2010), Financing R&D and innovation, in B.H. Hall and N. Rosenberg, Handbook of the Economics and Innovation, Elsevier, Amsterdam.

Hall, B.H., Mairesse, J. and Mohnen, P. (2010), Measuring the returns to R&D, B.H. Hall and N. Rosenberg, Handbook of the Economics and Innovation, Elsevier, Amsterdam.

Mairesse, J. and Mohnen, P. (2010), Using innovation surveys in econometric analysis, in B.H. Hall and N. Rosenberg, Handbook of the Economics and Innovation, Elsevier, Amsterdam.

Syverson, C. (2011), What determines productivity?, Journal of Economic Literature 49, 326-365.

# A personnel economics approach to productivity enhancement

Edward P. Lazear[*] and Kathryn L. Shaw[**]

*Summary*

Personnel economics is a field that grew out of business education and the need to provide aspiring managers with methods for increasing firm productivity. Most of the literature focuses on compensation and management practices that affect productivity with the goal of explaining in a positive sense and guiding in a normative sense the approaches that are used by firms. Examples of the practices and methods include performance pay, based on relative or absolute performance, the use of teams and their formation among workers with complementary skills, and careful screening of workers. This essay discusses some of the empirically most significant productivity-increasing methods, the importance of which is growing over time. The use of these practices is closely aligned with technological change, especially that which has made measurement, implementation and the necessity to use these approaches more critical. As technology progresses, personnel economic approaches to productivity enhancement are likely to become even more important.

*Keywords:* Personnel economics, productivity, human resource practices.
*JEL classification numbers:* J08, J30, J38.

[*] Stanford GSB, Lazear_Edward@gsb.stanford.edu.
[**] Stanford GSB, Shaw_Kathryn@GSB.Stanford.Edu.

The essence of personnel economics is enhancement of productivity. Personnel economics is defined as the application of formal economic theory and sophisticated statistical techniques to traditional human resources issues. Like all economics, personnel economics has three features that distinguish it from standard personnel analyses. First, the approach assumes rational maximizing agents. This means that it is possible to predict individuals' behavior. Sometimes individuals can have preferences that are nonstandard in the sense that they care about others, care about their relative standing in an organization, or care about the way that supervisors view them, but all of these can be built into standard approaches that are used in economics. Personnel economics uses these approaches to derive predictions about the real world.

Second, personnel economics builds in the importance of equilibrium. No model in economics is complete without understanding what equilibrium behavior will be and the same is true in personnel economics. This is particularly true in understanding how firms interact with one another. For example, a firm might attempt to exploit its workers. Were it operating in an environment that was competition free, the firm might be able to get away with such behavior. However, in a competitive labor market, where firms have to attract workers and retain them, exploitation is driven out of business by firms that behave more appropriately. Sometimes this may take time and there is no presumption that equilibrium is achieved instantaneously, but fundamentally firms do not operate in a vacuum and neither do workers.

Finally, economics focuses on efficiency. That is, understanding whether there are actions that can be taken that can make both sides better off in a transaction. By doing so, the economic approach to productivity allows us to assess whether the gains exceed the cost and make welfare comparisons based on these assessments.

Those points are primarily methodological. They describe the field and its approaches in terms of tools used and assumptions made to perform the analytic work but what personnel economics is really about is enhancing productivity. The main focus of personnel economics is explaining what we see in the real world and also providing guidance to firms so that they can think about their human resource practices more systematically and achieve higher levels of productivity. The key requirement is that firms maximize profits or productivity consistent with attracting workers. An alternative way of thinking about this, which generally results in the same implications, is that the firm must make workers as happy as possible consistent with allowing the firm to remain viable. To a first approximation, both approaches yield the same implications although the distribution of profits and surplus may be different under the two different assumptions. But again, no matter in which way the structure is fundamentally set up, the analysis is about achieving high levels of worker productivity.

This paper will be divided into three parts. In the first part, the magnitudes of the effects of personnel practices on productivity are described. In the second part, theoretical models of the impact of personnel practices on productivity are examined. In the third part, the potential for future productivity growth for the macroeconomy is addressed, using the theoretical models and past evidence as a foundation for the discussion.

## 1. Assessing the impact on productivity

How much do innovations in personnel practices raise productivity? Is there room for future productivity gains due to these management practice innovations? These questions are addressed by relying on a variety of empirical evidence, often going "inside" firms, working with managers and using the firm's data to produce evidence of productivity gains from management practices.

### *1.1 Innovative people management practices cover all human resource domains*

The personnel practices that may have increased productivity fit into a list of broad categories: Incentive pay or rewards, teamwork, training, careful hiring, flexible job assignment, information sharing and delegation of authority. Detailed examples of these practices include the following:

- The use of incentive pay – in the form of piece rate pay, bonuses for performance, stock options, promotion tournaments, or status and recognition for performance

- The use of teamwork – in the form of problem solving by teams, often composed of workers who have complementary skills
- The use of the careful screening of workers – through multiple methods and group evaluations**.**

Due to data availability, there is better empirical evidence on some of these practices than others.

### *1.2 The "HR technology shocks" of the past have been sizable*

Over the last twenty-five years, firms have adopted innovative personnel management or human resource (HR) management best practices. These practices are "innovative" in that the techniques for managing people have changed over time: there are new "best practices" in people management. As shown in Figures 1 and 2, in the US there appear to have been increases in the use of incentive pay and in teamwork for Fortune 1000 firms. About 28 percent of Fortune 1000 firms used self-directed work teams in 1987; by 1999, about 70 percent did. About 38 percent of Fortune 1000 firms used individual incentive pay in 1987; by 1999, about 68 percent did.[1]

It is best to visualize a comparison between HR technology shocks and computer technology shocks. There have been clear computer technology shocks, arising from the falling prices of computing power which caused firms to invest heavily in new information technologies (IT). The use of new people management practices has occurred because there have been revolutions in personnel economics and HR practices that represent new management practices: managers have learned how to do things better.[2] Incentive pay and teamwork have become new management tools. In the past, engineers did the problem solving; today, groups of workers from many skill sets do the problem solving. Therefore, the new best

[1] This survey measures whether any unit within the large firm uses these HR practices. The survey also shows that the intensity of use of these practices has grown within these large firms. However, the sample size for this survey has fallen over time, so there is a potential bias in the results – this is not a longitudinal survey of the same firms.

[2] See Shaw (2003a, 2003b, 2002) and Lazear and Shaw (2007) for these data, and more description and details on the trends in the "technology shocks of HR practices". For articles emphasizing the productivity consequences of these trends across firms, see Bartel et al. (2007), Bloom and van Reenen (2007) and Bloom et al. (2011).

practices of people management are introduced into firms much like the "technology shocks" of IT.

When we delve more deeply into industries beyond the Fortune 1000 firms, we see that the HR innovations are everywhere and are correlated with IT innovations. Figures 3 and 4 show the adoption of new HR practices and IT innovations in the valve manufacturing industry, for the US and the UK.[3] The trends are an increasing use of new HR technologies, as well as new computer technologies. Moreover, the adoption of either new IT or new HR is correlated across firms.

These data largely pertain to the United States. Are similar trends found in Europe as well? The answer is largely yes. The European Working Conditions Survey (EWCS) was conducted in four waves (1990, 1995, 2000 and 2005) and is analyzed in the European Commission report, "Employment in Europe, 2007". The data show that the major HR innovations were primarily adopted in the 1990's, and perhaps earlier in the 1980's, as is true for the US. The specific innovations were the move to the "learning organization" and the "lean organization". Both emphasize problem-solving, with the lean organization emphasizing teamwork (often in manufacturing) and the learning organization emphasizing worker autonomy in decisions.

There are considerable differences across countries in the level and pace of the introduction of new management methods, with the Nordic countries standing out. The Nordic countries and to a lesser extent Germany were more likely to adopt the learning organizational form. [4]

At the firm level, the major people management innovations may have occurred in the 1980's and 1990's for large firms, but for smaller firms and within firms, innovations continue. Moreover, the easily measured innovations, such as greater teamwork, may have plateaued, but more subtle innovations may well be ongoing, as discussed below. The last survey measuring major changes was conducted in 2000; no surveys measuring subtle changes have been done since then.

Why has the use of these innovative HR practices gone up and will this trend continue in the future? The greater use of more innovative HR practices has arisen for four primary reasons: managers are adopting new

[3] See Bartel et al. (2007, 2009).

[4] In addition to the European Commission report, see also Eriksson (2004) on Denmark and Bauer and Bender (2002) on Germany.

ways of managing people as managers learn innovative "best practice" methods; HR practices are complements to new information technologies that make process or product innovations; HR innovations are themselves computer-based (like hiring through monster.com or social networking through LinkedIn); and firms adopt new methods as the markets for their products change. All these mechanisms are discussed below and evidence of the changes and the enabling mechanism is provided.[5]

**Table 1. Forms of performance pay**

<table>
<tr><th colspan="2">Incentive Pay</th></tr>
<tr><th>Individual</th><th>Team</th></tr>
<tr><td>• Piece rate pay (pay tied by a formula to production or sales) (7% of firms)<br>• Skill based pay (72% of firms)<br>• Individual incentives (bonuses or compensation tied to short-term or long-term performance) (93% of firms)<br>• Typical individual incentives (some unmeasured)<br>    • Merit based raises in base pay<br>    • Special targeted bonuses<br>    • Promotions (tournament)</td><td>• Work Comp Incentives (80% of firms)<br>• Gainsharing (53% of firms)<br>• Profit Sharing or Stock Plan (greater than 70% of firms)</td></tr>
<tr><td colspan="2">Intangibles incentives<br>• Worker or team assigned to better job tasks<br>• Achieve higher status relative to peers<br>Noon monetary recognition (96% of firms)</td></tr>
<tr><td colspan="2">Measurement of performance – examples<br>• Services<br>• Manufacturing<br>• Knowledge of workers' performance evaluations (tools such as Vitality Curve, Balanced Scorecard, and Software of Success Factors)</td></tr>
</table>

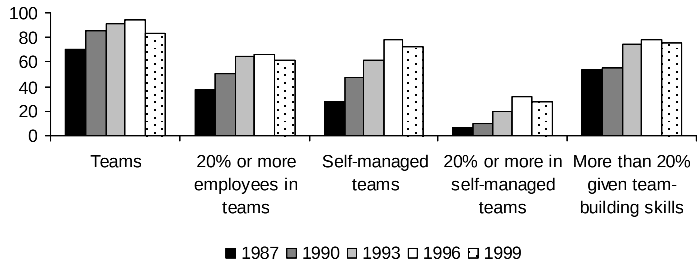
Sources of the numbers: Barkume and Moehrle (2001) and Lawler, Mohrman and Bensen (2001).

[5] Productivity has risen over time in industries that are computer-using industries, and computer use can be correlated with HR changes (Oliner and Sichel, 2002; Oliner et al., 2007; Jorgenson et al., 2002; Breshnahan et al., 2001, Bloom and van Reenen, 2007; Bartel et al. 2007).

## *1.3 Performance pay is rising and problem-solving skills are in demand*

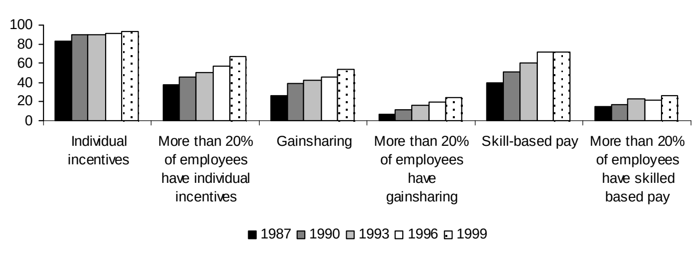
The use of performance pay may seem easier to measure than the other HR practices above, but it is not. Performance pay can go from very explicit pay for sales or output, to implicit rewards in the form of promotions, or even status rewards. Table 1 shows the alternative performance payment survey measures. Performance pay is any system that induces people to be more productive in return for increased rewards. It can even include more careful hiring: if firms hire more carefully and differentiate pay up-front across people, that is performance based pay. Such a broad definition makes it very difficult to measure past trends or forecast future trends in the use of performance pay.

**Figure 1. "HR technology shocks": Increased team use (US)**



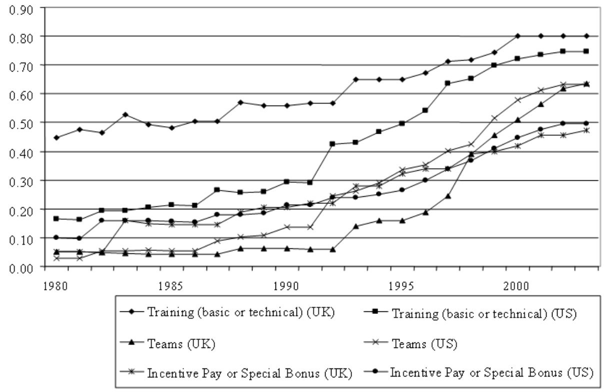

Source*:* Original data, see also Lawler et al. (2001), Shaw (2003a, 2003b) and Lazear and Shaw (2007).

**Figure 2. "HR technology shocks": Increased incentive pay (US)**



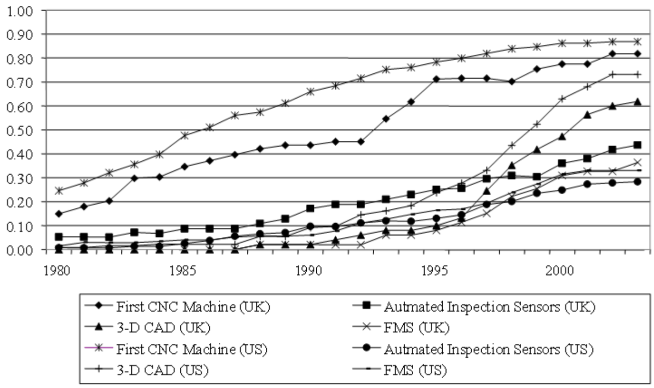

Source*:* See footnote to Figure 1.

**Figure 3. Proportion of plants with new HRM practices (UK v US) within the valve-making industry**




Source*:* Bartel et al. (2007) and Bartel et al. (2008).

**Figure 4. Proportion of plants with computer-aided production technologies (UK v US) within the valve-making industry**




Source*:* Bartel et al. (2007) and Bartel et al. (2008).

If we concentrate on compensation, the data suggest that pay for performance has risen over time, across a broad spectrum of jobs. Lemieux et al. (2009) show that from 1976 to 1998 (when their data end), the percent of workers who are classified as "working in performance pay jobs" grew from 33 percent to 40 percent. In Figure 2 and Figure 3 above, there is evidence of the increase in performance pay practices.

These measures of an increased use of performance pay do not capture the increases over time in careful hiring to match high-performing workers to demanding jobs. Time series data on employment by occupation show that firms have increased their demand for highly skilled workers in occupations that require problem-solving or decision-making skills. Autor et al. (2003) show that at all skill levels, from less than high school educated to college educated, the demand for non-routine cognitive skills has risen. Moreover, the technology shocks of an increased use of new information technologies have changed labor demand in the workplace. There has been an increased demand for workers who have the skills to work in computerized work places and the skills that are needed are the non-routine cognitive problem solving skills. People now make more sophisticated decisions using computers and more data. Additionally, there has been a growth in demand for workers who work with customers in ways that machines cannot substitute for labor.

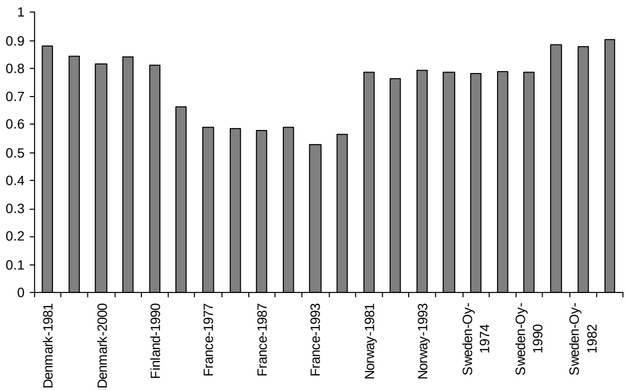
Wage rates have also responded to this increased demand for problem solving or skilled customer interaction. Among the lower tail of the wage distribution, wages have grown in jobs where personal service matters, such as childcare or specialty retail, e.g. Starbucks (Autor et al. 2008; Autor and Dorn, 2010). There is rising skill demand, and rising skills create greater value added per employee. Rising skill demand implies more careful screening and thus greater performance pay.

### *1.4 Star workers are far more productive than low-performers in virtually every industry. Star firms are far more productive than low-performing firms*

Data on pay and productivity from within companies show that the potential gains from improving the productivity of workers are substantial.

Assume, for the moment, that individuals' wages are a good proxy for personal productivity. The following is true. The variance of the distribution of wages across employees *within* the average firm is 60 to 70 percent as broad as the variance of wages across all employees in the economy. This is true across nearly all developed countries – the US and Europe – as shown in Figure 5. This means that there is a great deal of heterogeneity in what people do within firms. The average firm is not composed of identical lawyers, but is instead composed of people who work at all skill levels.[6]

**Figure 5. Compensation dispersion within firms**




Source: Lazear and Shaw (2009)

Note: This is the ratio of the average standard deviation of pay across workers within the firm to the standard deviation of pay across all workers in the economy.

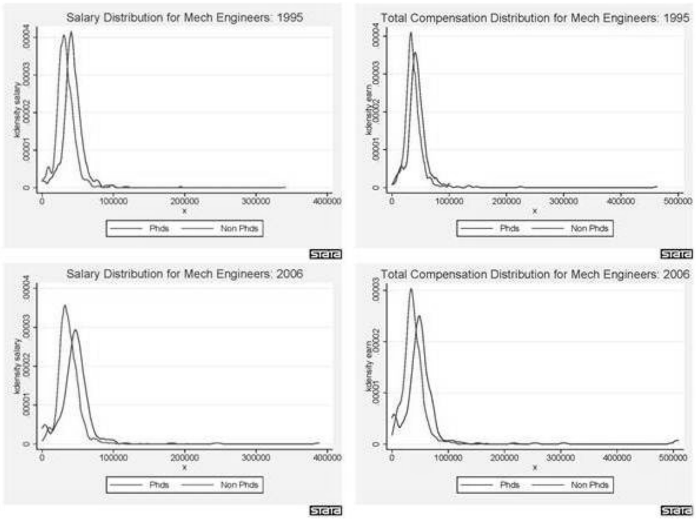
This fact has implications for human resource management practices. If there is so much wage variance within firms, it either means that the types of jobs that people do are very different (like clerk versus lawyer within a law firm), or that even when workers are doing identical jobs, these jobs are done with very different levels of individual productivity. Although there is no doubt that the former is true, data also show that the latter is equally true.

Even when people are doing jobs that are virtually identical, the variance in compensation and productivity across workers is large. Figure 6 shows the distribution of salary and total compensation for mechanical engineers in the US, in 1995 and in 2006 (in 2003 dollars). Figure 7 shows the distribution of wages for contract IT workers, comparing work for administrative IT tasks versus software development. Three points surface. There is a greater variance in productivity across workers in high skilled jobs than in lower-skilled jobs. As would be expected, skills and

[6] Shaw and Lazear (2008, Figure 1.4, p. 12).

effort and talent matter more in high-skilled than in low skilled jobs. Second, for mechanical engineers, the variance of compensation and the variance of salary have both gone up over time. This suggests that both the variance of personal productivity (skills and effort) and rewards for productivity, in the form of variable compensation, have gone up over time.

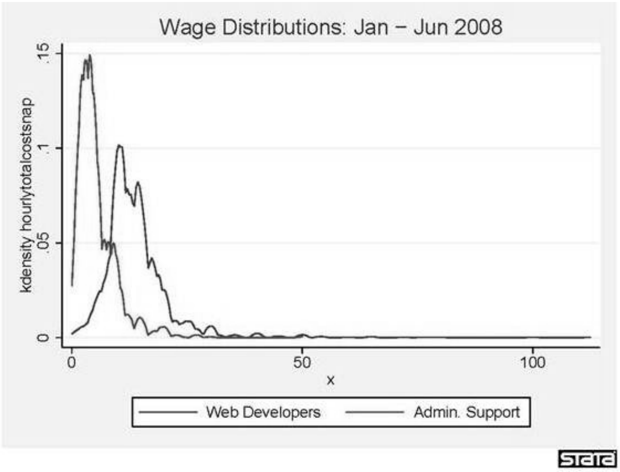
**Figure 6. The distribution of salaries and total compensation for mechanical engineers**




Source*:* Calculations by Chris Stanton (GSB, Stanford University) using confidential NSF SESTAT survey of scientists and engineers.

A third point is striking. Productivity data are much harder to obtain than compensation data, but the data show that even within very narrowly defined jobs, productivity varies tremendously across workers within narrowly defined occupations. Productivity is measured differently for each occupation – for teachers, it is measured as student test scores, for call center workers, it is seconds per call. Economists have asked the question, how much does individual talent/skill/effort determine productivity? Table 2 displays the answer in the last column. The percent of the variance of productivity that is explained by personal differences goes from 71 percent for windshield installers, to 32 percent for call center workers, to 13 percent for fruit pickers, to 7 percent for teachers.[7]

**Figure 7. Hourly wage distributions for contract workers in IT jobs**




Source*:* Calculations by Chris Stanton using confidential data from oDesk.com

Consider the teaching occupation. The evidence that 7 percent of the variance in student performance is due to teacher quality is a big impact for teachers, given little evidence of other factors that can improve students' performance. Thus, in the US and elsewhere, new management practices, like performance pay and training, are being put in place as a means of increasing the skill and effort of teachers.

The data above look at the variance of productivity across individuals, but more research explores patterns in the variance of productivity (or value added) across firms. Researchers have shown that firms that are categorized within very narrowly defined industries have very different productivity levels and productivity growth rates.[8] Thus, some firms are hugely productive, and some are not as productive. It may be that these

[7] The aim of measuring the variance of person fixed effects is that it takes out the effects of learning and time series demand shocks that affect productivity. Because the measure of productivity differs across all jobs, levels of productivity cannot be compared in the same way as levels of wages across jobs in Figures 6 and 7.

[8] These data refer to the average value added across all workers for manufacturing plants. In Haltiwanger (2008), the interquartile range in labor productivity across establishments within narrowly defined industries averaged 66 log points; in narrowly defined retail service industry categories, it averaged 57 log points.

firms are all making very different products. Or, it may be that some firms are better managed: cross sectional surveys of firms show that even firms within narrowly defined industries do choose very different HR practices (Bloom and van Reenen, 2007).

**Table 2. The variance of productivity across individuals, by occupation**

| | **No incentive pay** | | **Incentive or team pay** | | |
|---|---|---|---|---|---|
| | **Pay** | **Output** | **Pay** | **Output** | **Percent of variance explained by person effects** |
| *a)* Lazear (2000), Performance pay Auto glass installation | | 2.7<br>(-0.526) | | 3.24<br>(-0.491) | 0.710 |
| *b)* Bandiera et al. (2005), Social incentives Agricultural workers | | 5.01<br>-0.049 | | 7.98<br>-0.026 | 0.132 |
| *c)* Nagin et al. (2002), Monitoring call center employees | | 0.021 | | 0.019 | 0.320 |
| *d)* Hamilton et al. (2003), Teams and heterogeneity (USD%) Garment Production | 221.77<br>[131.87] | 86.8<br>[36.95] | 301.03<br>[142.4] | 105.45<br>[32.8] | |
| *e)* Rivkin et al. (2005). Teachers | | | | | 0.070 |

Notes: Means are reported with the coefficient of variation reported in parentheses below unless otherwise noted. The change in *r*-squared is (*r*-squared including fixed effects) - (*r*-squared with no heterogeneity). *a)* Output is units-per-worker-per-day. b) Output is kilograms picked per hour. *c)* Output is "suspicious or bad calls" / "good calls" per week. The incentive column is for sites after an audit experiment. The no incentive column is for all sites. Change in *r*-squared comes from monthly income regressions. The equivalent statistic for "bad call" regressions is .08. *d)* Reported statistics are the median. The 75 - 25 percent range is in brackets below. Earnings are weekly. Output is garments per week. *e)* The reported r-squared difference is from regressions of student math scores with teacher fixed effects versus school effects. The equivalent difference in r-squared for reading is .04.

### *1.5 The potential upside gains from human resource management practices are big*

Given the large difference between star performers and low-performers, there is plenty of room for improvement in individual productivity within firms. The use of more innovative HR practices can narrow the gap between stars and laggards within firms and across firms. What are these HR practices? Will firms have the incentive to invest in these new HR practices in the future? Which firms will invest the most?

The high variance in personal productivity within firms, described in the above section, implies that when a high quality worker is employed by a firm that demands his skills, the productivity gains will be large.

Consider the gains and methods of hiring and sorting to match workers to firms.

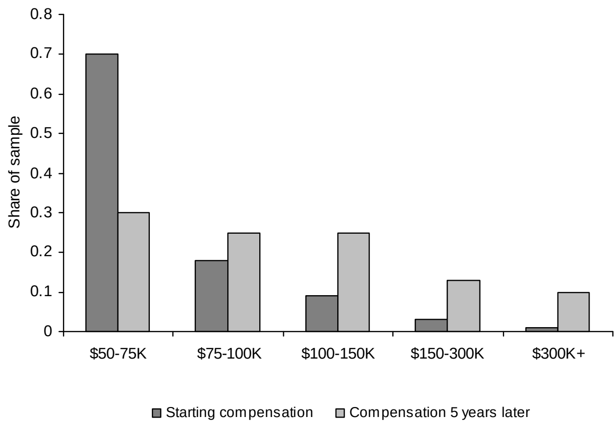
- There are very high gains to sorting workers well *across* firms – either by hiring carefully, or firing early. That is, there are gains from matching the right workers to the right firms. Firms can accomplish this sorting through better hiring practices, such as improved screening, or by using the management and compensation practices that attract the right workers to the firm.
- There are high gains from sorting workers *within* firms – sorting workers to the right jobs or teams within the firm.

Two management practices that are used to accomplish sorting across firms are:

- New hiring methods and IT that assists in hiring (like Monster.com). There is a much greater use of testing and informed interviews in hiring today.
- Personnel practices also induce people to self sort to firms. All personnel practices send signals to potential employees that induce them to choose the firm that best matches their skills or pay demands. Consider forms of incentive pay. Piece rate pay is preferred when the production function is simple, output is observable, and the value of teamwork is small. The firm that offers piece rate pay induces workers who will be productive to join the firm. Thus, as firms introduce a range of new personnel practices, workers are matched to the firms at which the worker will be most productive. Stock options are another example. They induce creative workers to join small firms. Promotion tournaments induce good managers to join large firms.

There is a great deal of evidence that workers are sorted within and across firms. Figure 8 shows the dispersion of compensation for all employees in the software industry, for employees when they are hired (starting compensation), and for the same employers when they have experience with the firm (about five years later). Notice that the variance of pay rises with tenure. The rising variance is a function of some sorting (low performers leave the firm) and internal forces that sort workers and train and reward. Once workers are hired, the number of high performers appears to grow and the number of low performers to fall. What personnel practices within the firm can be used to raise productivity within firms?

**Figure 8. Compensation in the software industry income dispersion for 51 000 software employees in 619 firms**




Source: Andersson et al. (2009).

### *1.6. Insider econometric evidence using data from within companies shows that the productivity gains from innovative HR practices can be very significant*

Personnel economists have gone “inside” companies and obtained the companies’ knowledge and datasets to estimate the effect of personnel practices, like performance pay and teamwork, on productivity.[9] The gains can be large. Across a variety of industries, the effect of personnel practice innovations can be as high as 50 percent of output.[10] The eco-

[9] See Shaw (2009) and Ichniowski and Shaw (2009) for detailed descriptions of the studies, with explanations for how much and why HR practices raise productivity.

[10] These papers estimate the productivity effects in very narrowly defined firms or industries – in firms that install auto glass, that pick fruit, in grocery stores, in steel mills. Why turn to evidence from such narrow firms and industries? There are two reasons. There are no data on both management practices and productivity across industries from which to estimate productivity effects, and if such data existed, the time lags between adoption and effect would be

nomic value attached to these productivity gains varies tremendously across firms: in capital intense steel mills, the value added from small worker productivity gains is huge; in labor intense windshield installation, the productivity gains are bigger, but the value added is smaller per employee.

These insider econometric studies show that the firm's choice of its optimal personnel practices can raise productivity independent of the firm's choice of its information technology. In addition, new investments in IT often motivate the firm's choice of its optimal personnel practices.

Why do HR practices raise productivity? Section 2 below, on Foundations on Personnel Economics, addresses this very question. Therefore, a few points are summarized here. Specific HR practices succeed in giving workers the incentive, the opportunity, and the ability to raise their productivity. Incentive pay gives workers the incentive, team problem-solving gives them the opportunity, and training gives workers the ability to raise performance. In addition, there are social influences with non-conventional tastes as their premise that try to explain personnel behavior. For example, for fruit pickers, when paid a fixed wage, managers picked as their team members the workers who were their friends, but after piece rate pay, managers picked the best workers (Bandiera et al., 2007). Social comparisons matter, but so do incentives.

There is an overall conclusion. New personnel practices operate on the intensive and the extensive margins of workers' decisions. Consider pay for performance as an example. When the firm moves to performance pay (from hourly pay), it operates on the worker's intrinsic margin – performance pay raises workers' performance conditional on holding the job. It also operates on the workers' extrinsic margin of behavior – it causes workers to sort across firms to form better matches.

Lastly, it is very important that HR management practices have "internal fit" – that they complement each other – and "external fit" – that they match the external conditions of the firm. These are described next.

---

unknown. In addition, researchers seek estimates of causal relationships – rather than correlations between management practices and productivity. By going deep inside firms – and obtaining their data on productivity and human resource practices – the causal impact of HR on productivity can be identified.

### *1.7 The value of adopting complementary human resource management practices*

Empirical studies show that the greatest gains from innovative personnel practices are achieved when firms introduce a set of complementary human resource practices. For example, when steel mills moved from hourly pay to performance pay, the mills also began to train workers more, use more problem solving teams that tap into workers' talent, and do more information sharing (Ichniowski et al., 1997; Gant et al., 2002). Existing theoretical and empirical studies emphasize the greater gains that arise when management practices are combined.

### *1.8 Which firms will adopt innovative HR practices?*

If HR practices can significantly improve productivity – through greater workers' effort, sorting, or bundling complementary HR practices – why do some firms adopt incentive pay while others do not? Best practices may evolve over time – as all firms tend to use more team work or incentive pay – but within narrow industries there can be a marked difference across firms.

The firm's choice of its optimal HR practices depends on the firm's choice of its product market strategy. Therefore, optimal HR practices vary across firms due to differences in firms' product markets, production processes and labor markets. Consider the airline industry. Southwest Airlines offers a low-cost service on short flights, and uses team-based HR practices with high levels of incentive pay. United offers premium services and uses HR practices to complement those high-level services. There are trends in the 'best practices' in the industry – both of these airlines now pay for performance (like flight time). But the set of HR practices that support Southwest would be different from those that support United, even if there were not any union influences.

Because HR practices follow individual firm's strategy, a within industry analysis is required to understand the optimal use of HR practices and the productivity gains. This is done for the software industry in Andersson et al. (2009). In most software companies, the employees work on new product innovations. But software companies differ markedly from each other. In firms that produce products like video games, there are huge potential upside gains to producing a new big-selling game. The potential gains to innovation are very skewed in this product market – there is a set of big winners. In firms that produce software for big firms, like mainframe software, the potential upside gains are small. Their goal is to produce innovations that keep their current customers happy. Anderson et al. (2009) show that the firms' HR practices reflect their product market strategy. The video game firms with the high potential upside gains pay higher levels of pay and a higher incentive pay to all employees, whether the firm actually succeeds or not.

Another example is provided by the valve-making industry, as described in Figures 3 and 4 above (Bartel et al., 2007). Valves are manufactured products that control the flow of water or air through pipes. Using data on 212 firms within the industry, the finding is that new information technologies have clearly raised productivity. However, new IT is most often adopted by the firms that produce customized products, rather than commodities. Furthermore, when new IT is adopted, new HR practices are also more likely to be adopted (better training especially).

In the steel minimill industry, the story is similar. The mills that gain the most from problem solving by line employees are those that produce complex products.

In these examples, product market strategy drives the HR choices that firms make. Within very narrowly defined industries, there is a wide range of HR practices in use. Firms that have different product market strategies, different production functions, or face different labor markets choose different HR practices, even within very narrowly defined industries.

Finally, firms do not immediately reach the new 'optimal' HR practices because there are adjustment costs that lower the profitability of adopting new practices. Workers, and firms, are earning rents from their investments in old practices, and until those rents are clearly gone, it is costly to invest in new practices. However, when competition lowers prices (or raises product quality), firms innovate or exit from the market.

## 2. Foundations of Personnel Economics

At this point, it is hoped that the reader has been convinced that various personnel practices have substantial effects on productivity and that personnel economists have gained some understanding of the nature of those effects and the primarily causal factors. In this section, we examine in more depth the specific personnel practices that can be used to affect productivity. Much of personnel economics was developed in business schools with the purpose of training future business leaders to enhance productivity and the welfare of their workforces. To be useful, it should be possible to delineate those strategies that are most effective in affecting productivity and giving guidance as to when each of the various tactics is likely to be most effective.

This section has three parts: In the first part, compensation practices that affect productivity will be examined; second, skill structure and productivity with an emphasis on entrepreneurship and leadership will be discussed briefly; and finally, specific management practices that can be used to enhance productivity will be analyzed in some depth.

### *2.1 Compensation*

Compensation is important because it is the primary method by which individuals' productivity can be affected. Although workers sometimes say that they care about other things than compensation, this cannot be true on the margin. Almost no one would be willing to work at their current jobs if they did not receive positive compensation for doing so. Compensation is a necessity of life and it is essential that individuals are compensated sufficiently to attract them to do the job in the first place. But the structure of compensation also matters a great deal in terms of affecting worker productivity and one of the findings of the empirical personnel literature is that compensation can have significant effects on worker productivity both in terms of choice of method of compensation and level.

Perhaps the most important point that comes from the compensation literature is that pay compression is the enemy of productivity. The moral position on equality often becomes confused with that on pay compression. The strongest case in favor of equality can be made for equality of opportunity, not for equality of outcome. It is relatively uncontroversial that accidents of birth, upbringing, or random misfortune should not determine a person's state in life. But pay compression that is justified on the grounds of fairness does not necessarily further equality of opportunity and may actually work in the opposite direction, disadvantaging those who offset misfortune by increased effort. Most of the theoretical and empirical literature argues that pay that is related to performance, either in an absolute or a relative sense, enhances productivity.

Another source of pay compression is a progressive tax structure. Although the compression induced by the tax structure can be offset by compensation policies at the level of the firm, the offset does not come without the standard distortion and deadweight loss that are associated with any tax system. There is now a significant literature that documents the adverse consequences of high taxes on growth at the aggregate level.[11] The personnel economics literature provides a microeconomic rationale for the adverse effects of taxation and growth. When pay is compressed through the tax structure, effort is reduced and workers are less well allocated to their most productive uses.

There are essentially three ways to structure compensation. The first is a direct payment on measured output sometimes referred to as a piece rate or commission. Second is payment on a measure of input almost always with a minimum requirement of effort that is necessary in order to retain the job. Third is compensation based on relative performance that usually takes the form of some kind of implicit contest or tournament-like structure. Let us consider each of these in turn.

The most basic kind of incentive pay is a piece rate where individuals are paid on basis of some measured level of output. For this purpose, the measurement must be relatively straightforward and inexpensive. In some jobs, it is simply impossible to do this in an effective way and compensating on the basis of output creates more problems than it solves. We return to these in a moment but right now, consider a simple technological process where individuals can be monitored and paid on the basis of their output. Such would be the case in agriculture and basic manufacturing where individuals produce an entire good or separable pieces of that good on their own. Historically, garments were fabricated and compensated in

[11] For example, Barro (1991) finds that growth is inversely related to the share of government consumption in GDP. Hansson and Henrekson (1994) find that government total outlays have negative effects on productivity. Widmalm (1999) finds that taxing personal income is negatively related to growth and the more progressive tax structures are associated with lower economic growth. Prescott (2002) argues that the difference in taxation between the US and France explains much of the difference between the two countries' growth rates. Bergh and Karlsson (2010) find that government size robustly correlates negatively with growth.

this way where each individual would be paid for the number of pockets sewed onto a pair of pants or the number of sleeves sewed onto a shirt. Agricultural workers are often paid on basis of the amount of produce that they pick or the number of seeds that they are able to plant. Service workers can also be paid on the basis of output when the service is easily observed, as is the case with sales people where revenues are a good metric for output. These are the production processes best suited to piece rate pay, but even here, distortions can be created when workers do not take all aspects of their behavior that are relevant to firm profitability into account.[12]

When individuals are paid on basis of output, two things happen. First, effort is directly affected by the compensation formula. Payment schemes that place a higher weight on output relative to input are likely to induce more output. As a result, a shift from a simple compensation scheme that pays on effort to one that pays on output can often work in the direction of enhancing productivity.[13]

Second, and sometimes equally important, is that different workers are attracted to piece rate environments. Consider a situation where worker productivity and ability are heterogeneous so that some workers are more able than others. If one firm pays a straight salary to all workers irrespective of output or ability and another firm pays on the basis of observed output, the latter is likely to attract higher ability individuals than the former. Not only is this conceptually the case, but evidence from the economic literature documents this quite clearly.[14] In this example, an auto glass company that installs windshields into cars changed their method of pay from straight hourly wages to a performance structure that guaranteed a minimum amount, but also allowed workers to earn more by paying them piece rates if their output were above a certain level. The theory, as

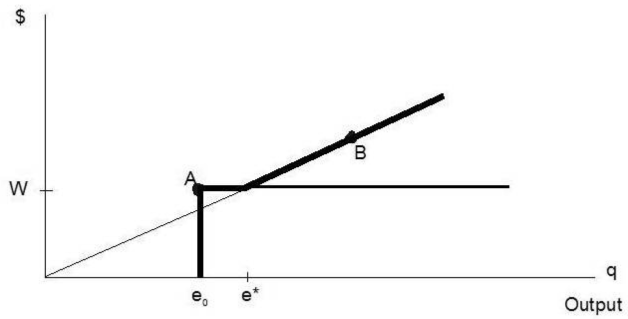
[12] The theory of these distortions is outlined in Lazear (1986) and more carefully examined in Baker (1992). See also Holmstrom and Milgrom (1991).

[13] Most of the literature that studies piece rates focuses on the incentive effect. Lazear (2000) notes a twenty percent increase in productivity from the incentive effect of piece rates. Shearer (2004) finds almost exactly the same magnitude. Bandiera et al. (2007) also find about a 20 percent increase in productivity. Some authors point out some of the adverse consequences of the piece rate schemes, e.g., Freeman and Kleiner (2005) who find a decrease in profits and Asch (1990) who shows that the timing of effort depends on the compensation scheme. Oyer (1998) examines the effects of quotas on distorting the timing of sales.

[14] Perhaps the most obvious example of this is found in Lazear (2000).

illustrated by Figure 9, has two predictions. First, average output will increase. Second, the firm will tend to attract the most ambitious workers. Before the switch to performance pay, individuals were paid a straight wage, *w*, for working at the firm. Their pay was independent of output as long as their output exceeded some minimal level shown in the diagram as $e_0$. Taken literally, there was no incentive to put forth output levels higher than $e_0$ except to the extent that worker pride and other non monetary factors were involved. Indeed, output before the switch to performance pay tended to be relatively low and concentrated around levels near $e_0$. After the switch to performance pay, workers were paid the maximum of *w* and the amount that they would earn if pay were calculated on a piece rate basis. In Figure 9, if their output were to exceed $e^*$, they would reap the benefit of that by being paid on a per unit basis. Given the compensation scheme, the break even level implied an increase in output from around 2.5 units per day to a level near 4 units per day. Thus, the worker was faced with a choice. He could continue to produce at level $e_0$ and still receive compensation level *w* or he could produce at some level higher than $e^*$ and receive a compensation level commensurate with this higher level of output. A point like *b* illustrates the choice made by a worker who opts to put forth more effort, produce higher output, and receive a compensation that exceeds *w*.

**Figure 9. Safelite: Before and after switch to performance pay**




Source*:* Based on Lazear (2000).

In addition to the effort effect that results because some workers choose to move from *a* to *b*, there is also a selection effect. The workers who find the new compensation scheme most attractive are those who will actually take advantage of the higher pay associated with higher levels of output. The workers most enamored of the scheme tend to be the most ambitious ones. As a result, over time there should be an increase in productivity for the typical worker not only because of incentives associated with the performance pay structure, but also because the nature of the workforce changes. In fact, that is exactly what was seen at the firm.

Table 3 shows that in the movement from hourly wages to piece rate, average output went up from 2.7 units per day to 3.24 units per day and the cost per unit went down from $44 per unit to $35 per unit. At the same time, actual pay went up and the standard deviation of pay went up because some workers were able to take advantage of the new output based program. What is not shown in Table 3, but comes from a slightly more sophisticated analysis of the data, is that about half of the increase in productivity was a result of an increase in effort by the existing workforce. The other half resulted from a change in the workforce from less productive to more productive workers. Indeed, workers hired after the shift to piece rates were about 25 percent more productive than those who were at the firm before that period.

**Table 3. Safelite output in two regimes**

| | Hourly Wages | | Piece rates | |
|---|---|---|---|---|
| | Mean | Standard deviation | Mean | Standard deviation |
| Number of observations | 13 106 | | 15 246 | |
| Units-per-worker-per day | 2,7 | 1,42 | 3,24 | 1,59 |
| Actual Pay | $2 228 | $794 | $2 283 | $950 |
| PPP pay | $1 587 | $823 | $1 852 | $997 |
| Cost-per-unit | $44.43 | $75.55 | $35.24 | $49.00 |

The fact that sorting is such an important part of increases in productivity is a key point not only in the context of the current study, but in understanding how changes in an industry or, even more generally, at a country level can affect productivity. At least in the short run, it is difficult for the workforce of an entire country to change in any kind of significant way, which means that the sorting effects that one sees between firms are not as likely to be as pronounced at the country level. To be sure, in an open environment, say like that permitted in the EU or between the states in the United States, individuals can move and will move over time to those countries or states that have firms and industries with the kind of compensation that is most suitable for them. But this is likely to be a slow process and one cannot expect to see the kinds of dramatic productivity effects from sorting that one sees between firms.

Another kind of sorting is relevant, even within country. When pay is compressed so that individuals do not directly experience the rewards of their output, there is less motivation to leave a job at which one is not good to move to another. This is a major problem in teaching (and perhaps other state sponsored jobs), where workers' salaries and job security are not very closely tied to their productivity. Even very poor teachers may have little incentive to seek other employment if they still receive relatively high wages and their teaching jobs are secure. Moving them from teaching into something to which they are better suited would raise aggregate productivity. Were teachers paid on the basis of their output, poor teachers would have more incentive to move into other occupations.[15]

The incentive effects that are observed can be expected both at the individual and the country level as long as the compensation structure is appropriate for those levels of aggregation. Still, the effect on output of switching from one scheme to another is pronounced. In this case, it is anywhere between 20 and 40 percent and these gains accrued to the firm within a six-month period. Additionally, because some of that increase in productivity was passed on to workers, the level of compensation particularly went up for those who were able to take advantage of the new system.

Although piece rates are effective at increasing productivity in certain kinds of industries, they are not appropriate for all workers and all occupations. In particular, some skills and aspects of output are difficult to measure. For example, consider government workers at the Ministry of Finance. Although it might be possible to describe the typical worker's duties and functions, it would be a difficult task to provide clear metrics of output that would be reliable and that would induce finance ministry employees to do exactly the right thing.

[15] See Champion and Shaw (2011) for an analysis of pay compression and its effects on teacher mobility.

One of the major problems associated with compensating on the basis of output is that all attributes of output cannot necessarily be spelled out. Under these circumstances, individuals may do exactly the wrong thing. This is sometimes put in terms of a distortion between quality and quantity. We all know the apocryphal stories of Soviet workers who were paid on the weight of nails that they produced and opted to make one five-hundred pound nail. While this is an extreme example of how output can be distorted by a particular compensation method, it illustrates the point. Some aspects of output are easier to measure than others and those are most likely to be the ones that receive compensation and are produced. Under these circumstances, it is often better to move to an input rather than an output based system.

An input based system has two advantages. First, it does not focus on one aspect of output while ignoring another. Second, the risk tends to be borne by the party best able to bear the risk, namely the firm. One disadvantage of compensating workers on the basis of their actual output is that it forces workers to bear the risk. For example, imagine a real estate sales person who is paid on the basis of sales. Real estate is a bulky item and there are some time periods when houses move quickly and others during which house sales are sluggish. If a sales person is paid solely on the basis of his or her sale revenue, then the sales person's income is highly volatile and the worker is in less of a position to bear that risk than is a firm which can diversify across workers, across regions, and over time periods. Consequently, in a risky environment and particularly one in which measurement is difficult and complex, firms pay on the basis of input.

It might appear that payment on input does not provide motivational incentives to workers. In some sense that is true, but it is an over-simplification because input-based pay is generally coupled with a minimum effort or output requirement. The incentives can be quite strong especially if that minimum requirement is set sufficiently high so that it is a stretch for the typical worker. For example, a worker who is paid solely on the basis of a time unit, like a month, (e.g. a monthly salary) could be required to work 14 hours per day six days a week during that month. A requirement of this sort would be onerous and would need to be coupled with very high compensation in order to induce people to take the job. But the job would certainly not suffer from lack of input. Naturally, the question remains of how much is produced per hour on the job. But if the job were one where input was a good proxy for output, then maintaining very high levels of input could provide the appropriate incentives.

Another major problem with this person-invariant input-based schemes is that a simple scheme with a minimum standard does not cater well to unobserved worker heterogeneity.[16] Piece rates have the advantage of being more flexible by allowing high productivity for a variety of workers in the same scheme. Take, for example, the piece rate discussed above in the context of Safelite Autoglass. The firm could accommodate a large variety of types of worker ambition with the one piece rate scheme. Those who want to work very hard find that they are permitted to do so and are rewarded for having high levels of output by having a take-home-pay that is significantly greater than that for those who work at lower levels of effort. A salary structure with a minimum windshield installation requirement would be too blunt an instrument in this context, causing productivity to suffer. For example, suppose that the average output requirement were three windshields per day. Some workers would find that level of effort too high and would be likely to leave the firm. Others would find that level of effort below that which they would like to offer if only they could receive higher compensation. The hybrid scheme results in too much effort for some and too little for others, whereas the piece rate scheme, which is variable, caters to heterogeneity in a much more continuous fashion.

Firms are somewhat reluctant to adopt piece rate pay because of the inability to measure all aspects of output and compensate them appropriately. Even if it were possible to identify the various aspects of output that are important on the job, it is difficult to set the prices correctly so that one factor is not over-weighted relative to another factor. Given these constraints, it is important to think about other compensation schemes that motivate and enhance productivity, but that do not steer workers in the wrong direction. The real world has largely solved this problem by using relative comparisons. The primary model that economists use to analyze this has come to be called “tournament theory”.

[16] When schemes are tailored to provide a menu of compensation and effort requirements they become tantamount to a piece rate.

### *2.2 Tournaments*

Most comparisons in life are relative. It is very difficult to do anything in the absolute without some basis for comparison and labor markets, particularly managerial ones, are no exception. Indeed, it is the relative nature of productivity that allows incentives to be created and productivity to be enhanced even when output is difficult to define and quantify. The standard way that this is done in managerial settings is through relative comparisons; this is the primary way in which economists think about promotion and the hierarchical structure of the firm.

The analogy is that of a sports tournament where contestants are paid on the basis of their relative performance. Specifically, the winner receives one prize and the loser receives another, lower prize. There are a few points to be made in this context that have parallels in the business environment.

First, in a tournament (like a tennis match), prizes are fixed in advance and depend on relative rather than absolute performance. The contestant who wins the championship at Wimbledon wins it not because he is good, but because he is the best player in the finals match. All contestants at Wimbledon are excellent, but what it means to win is that at least on that particular day, the winner was better than his opponent. Similarly, in the business world, the person who is promoted to CEO from the ranks of say executive vice presidents is promoted not only because he or she is able, but because the individual is viewed to be better than the competition. The CEO's compensation, while not necessarily fixed in advance, is generally within a range that is known and that is significantly different than that applicable to the executive vice presidents. Note also that the comparison and selection of the winner in the business tournament can be made on the basis of subjective evaluation rather than objective ones. It is certainly possible that in determining the choice of CEO, the board may use quantitative metrics, but they can also use "intangibles" and any other criteria that they deem reasonable to select the CEO.

Second, returning to the tennis match, the spread between the winner's and loser's prize affects motivation. Although contestants may want to win even if prizes equal one another (one enjoys winning at tennis even when one is playing merely for fun on a weekend outing), individuals become much more focused when, say, a million pounds are at stake. Similarly, in the context of the business environment, the difference between the compensation that goes to the CEO and that which is paid to the executive vice presidents affects the desire of individuals to become CEO. Compensation need not be merely monetary. It can take the form of increased power, flexibility, notoriety, or any other attribute. Productivity is affected by the spread between the winner's and loser's prize, monetary or other spreads. Furthermore, the effort that is affected is not the effort of the individual who has won the contest, but rather the individuals who are competing for the prize. This is pervasive in all aspects of economic activity. In business, associates in consulting, accounting and law firms work long and hard, hoping to become partners in the firm because partners are rewarded richly. In corporations, young managers largely strive to obtain results because they are cognizant of the positive career implications of achieving success for the business in which they are employed. Assistant professors put in long hours hoping to receive tenure at the university because tenured positions carry with them high salaries, job security, and other benefits that are viewed as valuable to the young faculty members. In government, assistant and deputy level individuals put forth a great deal of effort in part because they know that this will enhance their probability of getting higher positions later. Indeed, politics is an explicit tournament and substantial effort goes into winning elections because of the power and other aspects of reward that go with the jobs.

Third, there is such a thing as too much incentive. It is possible even in a tennis match to set the difference too high between the winner's and the loser's prize. If that difference is too large, individuals will not want to compete in the tournament because the amount of work required will exceed the expected return of the activity. Similarly, in the business environment, while it might appear that productivity can always be enhanced by increasing the difference between the wages of winners and losers, that is not necessarily the case in a competitive labor market where firms must attract workers to come to the firm. For example, individuals who go to consulting firms, accounting firms, law firms, or other firms where promotion to partnership is a goal, find that the expected number of work hours is extremely high. The upside may be very high, but if the probability of obtaining a partnership in the firm is sufficiently low, firms will find it difficult to attract individuals who are willing to put forth the additional amount of effort required for a low probability of a very high prize.

Thus, productivity can be adversely affected by setting the spread too high.

There are other reasons why increasing the spread is problematic. They relate to collusion and competition.[17] Firms need to be aware of these factors in determining the optimal wage structure within a firm so that the compensation scheme enhances rather than detracts from productivity.

Before elaborating on that, it is important to note that sometimes firms explicitly think about their tournament-like structure, while sometimes the tournament is merely implicit. However, we would argue that tournaments are a feature of almost every industrial business or political environment simply because it is impossible to avoid relative comparisons. Once relative comparisons are being made, then a tournament is at work and individuals will behave as if they are contestants in that tournament. Indeed, there is a great deal of academic evidence that documents the importance of tournaments in a variety of different contexts.[18]

One problem that arises when rewards are based at least in part on relative performance is that individuals may attempt to collude to circumvent the system. For example, in sports tournaments, on rare occasions players bet against themselves and attempt to bias the tournament by reducing their effort and allowing their opponent to win. A similar thing could happen in a business environment where, for example, individuals decide to alternate who gets the bonus in any particular year by simply taking it easy in one year and letting the other person win. While possible,

---

[17] See Dye (1984) and Lazear (1989).

[18] Eriksson (1999) tests the predictions of tournament theory by examining wage structures within Danish firms. Devaro (2006) uses a sample of recent hires and their initial promotions to fit a structural model of tournaments. He shows the importance of relative, rather than absolute performance in determining promotions. Knoeber and Thurman (1994) provide empirical evidence consistent with several specific predictions of tournament theory in the context of rewards for producers of broiler chickens. They show, for example, that spreads between prize values affect output. Using a survey of Australian firms, Drago and Garvey (1998) show that individuals are less helpful and work harder when the promotion incentives are strong. Ehrenberg and Bognanno (1990) show that golfers are affected by prize spreads in tournaments. Finally, Bull et al. (1987) do laboratory experiments (so, “effort” is chosen and stated rather than experienced) on how people respond to contests with piece rates and tournaments. The results of their experiments are generally supportive of tournament theory, though they find that less able people expend more effort than what would be predicted by the theory. There are numerous other papers that are authored by a large number of scholars, listed on a tournament website http://tournamenttheory.org/. Finally, an annual conference on tournaments and relative pay is held in North Carolina.

this kind of collusion is unlikely to be a major factor in affecting productivity. One way in which firms counteract this is by reserving the right to hire and promote from the outside when they believe internal candidates not to be sufficiently good. The ability to bring in outsiders who are not part of the potentially collusive agreement is a force that tends to break any kind of collusive arrangement.[19]

An effect that goes in the opposite direction is more problematic. Relative comparisons reduce cooperation. When individuals are paid on the basis of their relative performance, they do well not only by making themselves look good, but also by making their rivals look bad, which may discourage them from acting cooperatively. There are a couple of methods that firms can use to mitigate this problem.

First, the firm can move away from relative compensation and towards absolute compensation. Workers on straight piece rates do not have any incentives to act uncooperatively, but do not have much incentive to act cooperatively either. Second, firms can reduce the spread between wages in the firm's hierarchical structure. Thus, the firm reduces the incentive to engage in destructive or uncooperative behavior. At the same time, a contraction in salary difference reduces the incentives to put forth effort so that there is a trade off and one that may be unavoidable. This is particularly true when workers are especially aggressive toward one another and also when cooperation between workers is particularly important.

This points to a third way of reducing the adverse consequences of relative compensation on cooperation, which is to make sure that those who should cooperate with one another are not competing for the same job. For example, if cooperation between workers A and B is important and cooperation between workers C and D is important, it is better to have A compete against C and B compete against D for promotions. This can be done geographically or across different product lines.

A final method by which cooperation can be enhanced is to attempt to change workers' tastes so that they find cooperation to be a positive aspect of the job rather than a negative one. The military, for example, is well known for its effective methods of creating comradery among soldiers who might otherwise want to compete with one another to outshine their peers on the battlefield and enhance their probabilities of promotion.

---

[19] See Chan (1996).

Instead, during training programs and other activities, the military helps create strong bonds between soldiers that make them feel extreme loyalty towards one another, which offsets any competitive desires and enhances the cooperative ones. To the extent that individuals can be made to find competition and anti-productive behavior odious, productivity is enhanced.

### *2.3 Skill structure*

Because this volume has other chapters that are directly devoted to human capital, we do not spend much time on that here despite the incentives being quite important in determining the skill structure of an economy. But it is worth noting that economies move forward because leaders and entrepreneurs have multiplicative rather than additive effects on their firms and industries. A great breakthrough can be used not only by the individual who invents it, but by many others. Consequently, its effect on productivity is multiplied by the number of potential users of the technology. Creating an environment that fosters great leadership and significant entrepreneurship is an important component of economic growth and enhances short-run productivity, but more importantly affects technological change over time. In a series of papers, it has been argued that leadership and entrepreneurship are enhanced by the existence of general rather than specific skills (Lazear 2005, 2011). The basic idea is consistent with the “weakest link” concept, which states that a chain is no stronger than its weakest link. A leader or entrepreneur must have a variety of skills to either perform the many tasks necessary to make a business successful or to choose others wisely who can act as team mates in carrying out that effort.

Consider, for example, an individual who wants to start a restaurant. If he is not a good chef, the restaurant will go nowhere because the cuisine will be sub-par. But even the best chef cannot create a successful restaurant without business skills. He must be able to raise capital, acquire supplies, manage people, and do this all with a sufficient amount of expertise so that the business survives. It is certainly possible for the chef to delegate this to others but in putting together the team, he must understand enough about the other required skills to be able to choose the right people.

Similarly, leaders encounter problems in a variety of different contexts and in order to make good decisions, they must have at least some knowledge in each of a large number of areas. True, leaders can delegate some of the information gathering and analysis to others but fundamentally, the decision making comes down to the leaders and without the ability to evaluate the work produced by their team, they will not make good decisions. Because the problems are varied, it is necessary for leaders to have sufficiently general skills.

The above mentioned papers test the theory by looking at the skill set of successful managers; those who are most likely to succeed and rise to leadership positions or to found their own businesses are those who have general skills either acquired in school or at the firm.[20]

The notion that general skills are necessary for leadership has important policy implications. Most important, the need for general skills suggests that an educational system that tracks people into specialties too early in life may be detrimental to entrepreneurship. Although such a system may be excellent for producing high quality, specialized individuals who are good at doing what they are trained to do, those workers may not have sufficiently general skills to allow them to become leaders or entrepreneurs. If an economy is too specialized, it may be well-suited to the present tasks, but less well-suited to the uncertainties and technological changes that will confront the economy in the future.

## 3. The future

Will future innovations in personnel practices within firms result in increased productivity growth for the entire economy? There are no data on management innovations that can be used to predict future productivity gains from these innovations. However, the above models, the past data, and the future competitive pressures suggest that management innova-

[20] Wagner (2003) finds supporting evidence for the model using German data, as do Backes-Gellner and Moog (2007) who examine the actual skills used on the job among those who are leaders and entrepreneurs. In Wagner (2006), using regional stratification and rare events of entrepreneurial events, it is found that the generalist notion is supported. See also Backes-Gellner et al. (2010) and Tuor and Backes-Gellner (2010).

tions have significantly increased productivity in the past and could also do so in the future.

### *3.1 Human resource management practices could play a significant role in future trend productivity improvements*

The evidence presented above suggests that HR practices contributed to trend productivity gains in the last fifteen years in the US. The data on time trends in the adoption of HR practices show that firms have increased their use of innovative management practices – like performance pay and teamwork – over time, especially in manufacturing but also for the entire economy. Empirical studies show that firms' use of these new management practices can have large performance gains. Combined, there is a general belief that improved management practices have raised productivity. There is further empirical evidence that innovative HR practices were especially important in making the information technology investments of the 1990's more productive.

Will management practices raise the trend productivity for developed economies in the future? Here, the evidence is mixed. Some aggregate time trends suggest that performance gains have been realized and will not be repeated. There are two reasons for this:

- Firms are currently investing less in information technologies and software than in the 1990's. If innovative HR practices are complements to IT, then HR innovations will also decline (Gordon, 2008).
- The data suggest that firms' investments in easily measured HR innovations have tapered: Firms have leveled the degree to which they use performance pay and teams – it is not spreading more (see Figures 1 to 4 above).

However, case study evidence suggests that future gains from innovative personnel practices could be substantial, as described next.

### *3.2 Case study examples strongly suggest that HR innovations are ongoing and that software firms are producing important new IT-based HR innovations*

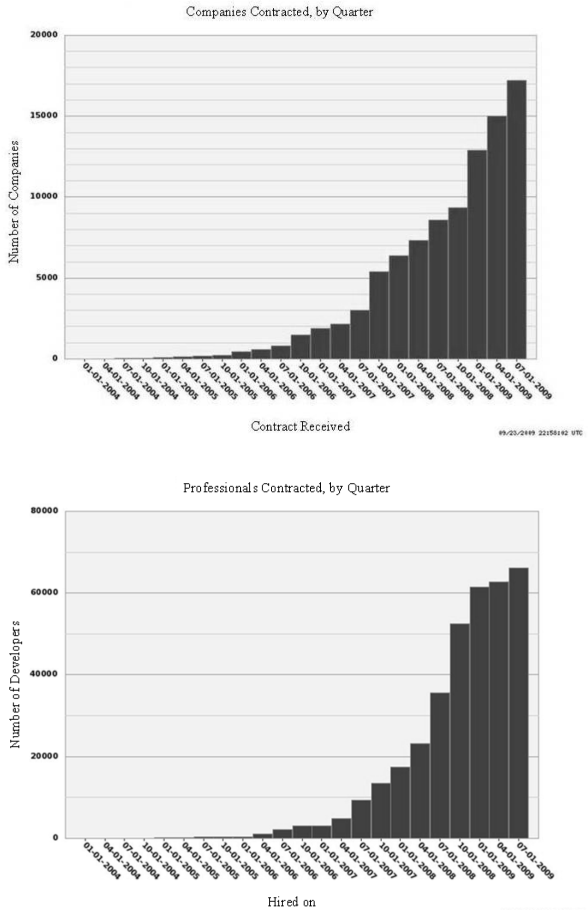
Productivity gains in the future may come from new HR practices and, often, from new HR practices that are built upon new information technologies. There are no databases measuring the current trends in HR adoption or IT-based HR investments by firms. Therefore, case study evidence is required. In a series of quick examples, numerous HR innovations – in hiring, job matching, training, teamwork, and performance evaluation – are highlighted.

The recent innovations in hiring are clear: all firms are using some internet search to look for potential employees. And, firms and workers continue to improve the use of powerful Internet search tools. For example, the online e-recruiting website of Monster.com does not just post ads and help people build their résumés, it also obtains a large part of its revenue from creating custom online application forms and processes for corporate websites.[21] By 2003, Monster Global had built up a huge network of interlocking sites with local content and language in North America, Europe and the Asia Pacific Region. US companies are sourcing workers from around the world through Monster Global, for direct employment and for outsourcing.

By 2007, prior to the recession and the drop in hiring, companies stated that 44 percent of their new hires in the previous year came through e-recruiting on-line. The time trend is clear as well. In 1998, 29 percent of Global 500 companies had a corporate website and, by 2003, 94 percent had one. North American companies were the early adopters of corporate employment websites, but now all large companies around the world have these websites (usually run by US companies). Moreover, when firms can package jobs into tasks that can be outsourced, they do so by hiring workers from around the world through many online search and matching websites that are experiencing a huge growth in usage. Figure 10 displays the rapid growth, as represented by the contracting company oDesk.

[21] See Nakamura et al. (2009) for details.

**Figure 10. The growth of the oDesk market matching on the internet**




Source*:* Calculations by Chris Stanton using confidential data from oDesk.com.

Note: ”Companies contracted” are the number of companies that use oDesk matching software to hire IT employees. ”Professionals contracted” are the number of workers hired for a contract IT job through the oDesk software.

These online job search tools are improving the quality of job matches between employer and employee, by enabling currently employed people to hunt for alternative jobs. In the past, job hunting in person or through friends was time consuming and awkward. Firms that hire from the pool of currently employed people reduce the “lemons” problem – of hiring from the pool of people who other employers did not want.[22] In 2007, Nakamura and Freeman ran an international survey of people who were using on-line employment websites. The vast majority of the respondents to the survey reported that they were employed at the time that they were logging on to the e-recruiting job sites, and a third of all respondents were uploading their résumés to the jobsites (Nakamura et al., 2009).

Many firms are increasing their investments in workers' training, as firms aim to reach broader world-wide markets and have more data on workers' performance following training. Some big new investments are in service industries. Workers are being trained to follow the rules of standard operating procedures and to make better decisions when solving problems. These training expenditures are not well measured or reported, but are well known. Many have heard of "corporate universities", starting with Hamburger University of McDonald's, and General Electric and Microsoft and Cisco Systems have all built "universities". Examples of other universities are Starbuck's University, Men's Wearhouse (men's clothing), and Cheesecake Factory (restaurants). Cheesecake Factory, a restaurant chain, estimates that they spend $2,000 a year on training each individual server (http://www.workforce.com/section/11/feature/24/35/18/). Investments are also growing for high-skilled service jobs. Tata Consulting Services (TCS) is a well-known example. Tata reports that it spends six percent of its operating revenue on training: it built its most massive training center in 1997, which trains 3,500 employees per year. Most major Tata locations have training centers around the world.[23] A consulting firm, the Corporate University Xchange, provides estimates of what it thinks is the number of these "universities", rising from 400 in 1993 to 2,000 in 2001 and to perhaps as much as 3,700 by 2010.[24]

There are expansions in the IT-based HR innovations aimed at managing team performance within sites and across international sites. Consider the software innovations. Teamwork software focuses on collaboration between people. There are hundreds of web-based shared workspace applications available.[25] These are created both within companies and by

---

[22] However, it increases the 'winner's curse' – of paying too much for the person in an active market.

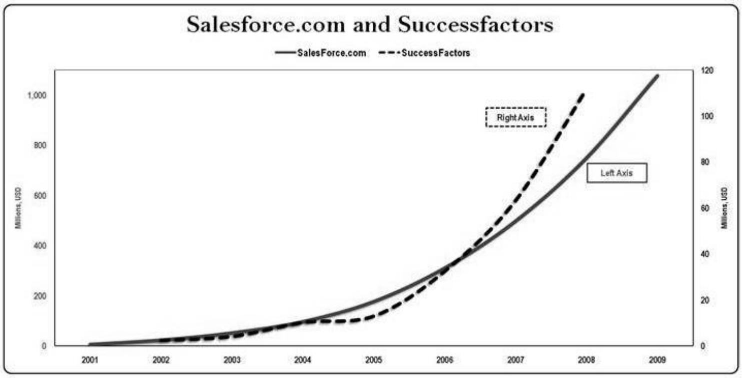
[23] Williams (2009).

[24] See Hearn (2002).

[25] Consider Microsoft's Groove, a web-based common workspace, and many other products, such as BaseCamp, intranets.com and drupal.

external vendors. Social network software also focuses on making the links between people that are useful for hiring and for the sharing of intellectual capital. Social networking tools, like Facebook, are increasingly aimed at building networks for work across firms and also within firms. These social networking sites will change the hiring and internal management of workers in the future.

**Figure 11. Operating revenues for two firms that offer web-based HR management tools**




Source*:* Calculations by Nick Obradovich from annual reports for operating revenue.

The quality of performance evaluation of employees is greatly enhanced with the use of internet-based tools for posting and sharing evaluations. Two well-known firms are highlighted. The firm SuccessFactors creates software for online performance evaluation of people who do all kinds of jobs, and for real-time analysis of the performance data. Most US Fortune 500 firms, and many firms around the world, have begun to purchase SuccessFactors software and the company's growth accelerated dramatically since it went public in 2001: SuccessFactors states that its services are translated into 32 languages and with more than 3 000 firms purchasing their software for use by 6 million employees in 60 industries. Similarly, salesforce.com has seen a massive growth. Figure 11 displays the rapid growth that now occurs.

All these IT-based HR innovations are very recent. The diffusion and optimal use of them is certainly in their infancy. When we look at technological innovations – going as far back as the introduction of hybrid corn into agriculture in the 1940's – we see that the diffusion and application of the innovations takes many years. Thus, the potential productivity gains from the IT-based managerial innovations are as yet unknown, and could be rising, in developed countries and developing countries.

### *3.3 What next? The future is data and HR innovations that use data*

All firms have more data. For example, in the services industries, like retail trade, firms gather data on customer satisfaction and on the productivity of every worker who sells a product at a checkout stand. In knowledge-driven businesses, firms have data residing in project management tools and in new performance evaluation data sets. All large firms invested in Enterprise Resource Planning software in the last 20 years, linking different types of data within the firms. These data will prompt a greater HR optimization. Firms have focused on product and process innovations in the past, but not on optimizing the management of employees. That has yet to come.

The new datasets within firms may also improve productivity by using "rules" based management. The "rules versus discretion" debate that is prominent in macroeconomics also applies to personnel economics. When a patient with chest pain enters an emergency room, should the team of nurses and doctors have rules to follow to treat the patient, or should the lead doctor make decisions that are specific to the care of that individual? Data on the distribution of productivity across individual employees suggest that there could be a huge variance in doctors' productivity. Thus, leaving decisions up to the lead doctor could invite a lower quality of care, if that doctor is not highly skilled, or a higher quality of care, if that doctor is highly skilled. Firms are aware of the tradeoff between rules and discretion; you need rules for basic performance but discretion for abnormal performance issues. Fast food chain stores train workers in the exact rules to follow to make the perfect cup of coffee – there is no discretion – employees follow standard operating procedures.[26] These same employees, who are trained to follow rules, are also trained to use discretion when warranted. When unusual problems arise, discretion is desirable. Such discretion could push decisions to lower levels of the firm, to increase the speed of operations (Wulf, 2007), or could push the decision

[26] These rules have benefits for time consistency: employees know how their behavior will now affect future rules and outcomes.

up the hierarchy, so that each tier evaluates the quality of the decision.[27] In all cases, new data are resulting in firms rethinking their use of rules versus discretion.

Are there market pressures that are likely to increase HR innovation? For firms producing goods in tradable sectors, trade pressure increases innovation. Evidence from microeconomic papers shows that firms with more exports are more productive.[28] For workers in tradable goods sectors, productivity in the US should rise to offset the higher costs of US labor. This has been the case in manufacturing in the past. As prices abroad fall, or wages abroad fall, new investments in personnel practices become profitable. However, increased trade protection is a danger that could lower productivity.

Domestic market pressure may also increase competition. Forecasts of higher savings rates, or reduced consumption rates, can contribute to that pressure.

Note, finally, that this past recession is likely to be raising productivity by introducing structural changes that favor firms that have the most innovative HR practices, just as past recessions favored the most productive manufacturing firms. When manufacturing output recovered after the recessions of 1982 and 2001, manufacturing employment was permanently down. Firms also used the 1982 recession as an opportunity to close plants and lay off workers who had become redundant as a result of technical change in production methods. The same thing is happening in service sector jobs. Firms that are less productive, and often have poorer HR practices, are permanently closing.

In sum, HR innovations and IT innovations that support HR are ongoing. There remains a marked difference in the productivity of workers doing similar jobs within firms, and firms are seeking ways of lowering this variance and increasing the mean levels of workers' productivity. Firms are also seeking ways of producing higher value added products. These influences combine to lead to greater HR innovations.

---

[27] Pushing decisions down in the hierarchy will increase the number of false positive errors, where employees begin a project believing it is valuable when it really is not, and pushing decisions up the hierarchy will increase the number of false negative errors, or rejecting projects even when they would have been good.

[28] See Bloom et al. (2009) and Guadalupe and Wulf (2010) and references therein.

## References

Andersson, F., Haltiwanger, J., Freedman, M., Lane, J. and Shaw, K.L. (2009), Reaching for the stars: Who pays for talent in innovative industries?, Economic Journal 119, F308-F332.

Asch, B.J. (1990), Do incentives matter? The case of navy recruiters, Industrial and Labor Relations Review 43, 89S-106S.

Autor, D. and Dorn, D. (2010), Inequality and specialization: The growth of low-skill service jobs in the United States, NBER Working Paper 15150.

Autor, D., Katz, L.F. and Kearney, M.S. (2008), Trends in U.S. wage inequality: Revising the revisionists, Review of Economics and Statistics 90, 300-323.

Autor, D., Levy, F. and Murnane, R. (2003), The skill content of recent technological change: An empirical investigation, Quarterly Journal of Economics 118, 1279-1333.

Backes-Gellner, U. and Moog, P. (2007), Who chooses to become an entrepreneur? The jacks-of-all-trades in social and human capital, Institute for Strategy and Business Economics Working Paper 76, Zurich.

Backes-Gellner, U., Tuor, S.N. and Wettstein, D. (2010), Differences in the educational paths of entrepreneurs and employees, Empirical Research in Vocational Education and Training 2, 83-105.

Bandiera, O., Barankay, I. and Rasul, I. (2005), Social preferences and the response to incentives: Evidence from personnel data, Quarterly Journal of Economics 120, 917 - 962.

Bandiera, O., Barankay, I. and Rasul, I. (2007), Incentives for managers and inequality among workers: Evidence from a firm-level experiment, Quarterly Journal of Economics 122, 729-773.

Barkume, A.J. and Moehrle, T.G. (2001), The role of incentive pay in the volatility of the employment cost index, Compensation and Working Conditions, Summer 2001, 13-18.

Barro, R.J. (1991), Economic growth in a cross section of countries, Quarterly Journal of Economics 106, 407-443.

Bartel, A., Ichniowski, C. and Shaw, K.L. (2007), How does information technology affect productivity? Plant-level comparisons of product innovation, process improvement, and worker skills, Quarterly Journal of Economics 122, 1721-1758.

Bartel, A., Ichniowski, C., Shaw, K.L. and Correa, R. (2009), International differences in the adoption and impact of new information technologies and new HR practices: The valve-making industry in the U.S. and U.K., in R. Freeman and K. Shaw (ed.), International Differences in the Business Practices and Productivity of Firms, University of Chicago and National Bureau of Economic Research, Chicago and Boston.

Bauer, T. and Bender, S. (2002), Flexible work systems and the structure of wages: Evidence from matched employer-employee data, IZA Discussion Paper 353, Bonn.

Bergh, A. and Karlsson, M. (2010), Government size and growth: Accounting for economic freedom and globalization, Public Choice 142, 195-213.

Bloom, N. and van Reenen, J. (2007), Measuring and explaining management practices across firms and countries, Quarterly Journal of Economics 122, 1351-1408.

Bloom, N., Genakos, C., Sadun, R. and van Reenen, J. (2009), Management practices across Asia, Europe and the US, manuscript, Centre for Economic Performance, London.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D. and Roberts, J. (2011), Does management matter? Evidence from India, NBER Working Paper 16658.

Breshnahan, T., Brynjolfsson, E. and Hitt, L. (2001), Information technology, work organization and the demand for skilled labor: Firm-level evidence, Quarterly Journal of Economics 117, 339-376.

Champion, S. and Shaw, K.L. (2011), Teachers' pay compression: Leaving for opportunity, Working Paper, Stanford University Graduate School of Business.

Baker, G. (1992), Incentive measures and performance measurement, Journal of Political Economy 100, 598-614.

Chan, W. (1996), External recruitment versus internal promotion, Journal of Labor Economics 14, 555-570.

Devaro, J. (2006), Internal promotion competitions in firms, RAND Journal of Economics 37, 521-542.

Drago, R. and Garvey, G.T. (1998), Incentives for helping on the job: Theory and evidence, Journal of Labor Economics 16, 1-25.

Dye, R.A. (1984), The trouble with tournaments, Economic Inquiry 22, 147-149.

Ehrenberg, R.G. and Bognanno, M.L. (1990), The incentive effects of tournaments revisited: Evidence from the European PGA tour, Industrial and Labor Relations Review 43, 74-88.

Bull, C., Schotter, A. and Weigelt, K. (1987), Tournaments and piece rates: An experimental study, Journal of Political Economy 95, 1-33.

Eriksson, T. (1999), Executive compensation and tournament theory: Empirical tests on Danish data, Journal of Labor Economics 17, 262-280.

Eriksson, T. (2004), How common are the new compensation and work organization practices and who adopts them?, Department of Economics Working Paper 01-8, Aarhus School of Business.

European Commission (2007), Employment in Europe 2007, October, European Commission, Brussels.

Freeman, R.B. and Kleiner, M.M. (2005), The last American shoe manufacturers: Changing the method of pay to survive foreign competition, Industrial Relations 44, 307-330.

Gant, J., Ichniowski, C. and Shaw, K.L. (2002), Social capital and organizational change in high-involvement and traditional work organizations, Journal of Economics and Management Strategy 11, 289-328.

Gordon, R. (2008), The slowest potential output growth in U.S. history: Measurement and interpretation, for Presentation at CSIP Symposium on The Outlook for Future Productivity Growth, Federal Reserve Bank of San Francisco, November 14 2008.

Guadalupe, M. and Wulf, J. (2010), The flattening firm and product market competition: The effect of trade liberalization on corporate hierarchies, American Economic Journal: Applied Economics 2, 105-127.

Hamilton, B.H., Nickerson, J. and Owan, H. (2003), Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation, Journal of Political Economy 111, 465-497.

Haltiwanger, J. (2008), Top ten questions for understanding firm dynamics and productivity growth, Keynote address at CAED 2008.

Hansson, P. and Henrekson, M. (1994), A new framework for testing the effect of government spending on growth and productivity, Public Choice 81, 381-401.

Hearn, D.R. (2002), Education in the workplace: An examination of corporate university models, http://www.newfoundations.com/OrgTheory/Hearn721.html.

Holmstrom, B. and Milgrom, P. (1991), Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, Journal of Law, Economics, and Organization 7, 24-52.

Ichniowski, C. and Shaw, K.L. (2009), Insider econometrics: Empirical studies of how management matters, forthcoming in R. Gibbons and J. Roberts (eds.), Handbook of Organizational Economics and NBER Working Paper 15618.

Ichniowski, C., Shaw, K.L and Prennushi, G. (1997), The effects of human resource management practices on productivity, American Economic Review 86, 291-313.

Jorgenson, D.W., Ho, M.S. and Stiroh, K.J. (2002), Projecting productivity growth: Lessons from the U.S. growth resurgence, Economic Review, Federal Reserve Bank of Atlanta, Third Quarter, 1-13.

Knoeber, C.R. and Thurman, W.N. (1994), Testing the theory of tournaments: An empirical analysis of broiler production, Journal of Labor Economics 12, 155-179.

Lawler, E.E., Mohrman, S.A. and Benson, G. (2001), Organizing for High Performance: The CEO Report on Employee Involvement, TQM, Reengineering, and Knowledge Management in Fortune 1000 Companies, Jossey-Bass, San Francisco.

Lazear, E.P. (1986), Salaries and piece rates, Journal of Business 59, 405-431.

Lazear, E.P. (1989), Pay equality and industrial politics, Journal of Political Economy 97, 561-580.

Lazear, E.P. (2000), Performance pay and productivity, American Economic Review 90, 1346-1361.

Lazear, E.P. (2005), Entrepreneurship, Journal of Labor Economics 23, 649-680.

Lazear, E.P. (2011), Leadership: A personnel economics approach, NBER Working Paper 15918.

Lazear, E.P. and Shaw, K.L. (2007), Personnel economics: The economist's view of human resources, Journal of Economic Perspectives 21, 91-114.

Lazear, E.P. and Shaw, K. (2009), Wage structure, wages, and mobility, in E. Lazear and K. Shaw (eds.), The Structure of Wages: An International Perspective, University of Chicago, National Bureau of Economic Research, Chicago.

Lemieux, T., MacLeod, B. and Parent, D. (2009), Performance pay and wage inequality, Quarterly Journal of Economics 124, 1-49.

Nagin, D.S., Rebitzer, J., Sanders, S. and Taylor, L. (2002), Monitoring, motivation and management: The determinants of opportunistic behavior in a field experiment, American Economic Review 92, 850-872.

Nakamura, A.O., Shaw, K.L., Freeman, R.B, Nakamura, E. and Pyman, A. (2009), Jobs online, in D.H. Autor (ed.), Studies of Labor Market Intermediation, University of Chicago Press, Chicago.

Oliner, S.D. and Sichel, D.E. (2002), Information technology and productivity: Where are we now and where are we going?, Federal Reserve Bank of Atlanta Economic Review 87, 15-44.

Oliner, S.D., Sichel, D.E. and Stiroh, K.J. (2007), Explaining a productive decade, Brookings Papers on Economic Activity 1, 81-137.

Oyer, P. (1998), Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality, Quarterly Journal of Economics 113, 149-185.

Prescott, E.C. (2002), Prosperity and depression: 2002, American Economic Review 92, 1-15.

Rivkin, S.G., Hanushek, E.A. and Kain, J.F (2005), Teachers, schools, and academic achievement, Econometrica 73, 417-458.

Shaw, K.L. (2002), By what means does information technology affect employment and wages?, in N. Greenan, Y. L'Horty and J. Mairesse (eds.), Productivity, Inequality, and the Digital Economy: A Transatlantic Perspective, MIT Press, Cambridge MA.

Shaw, K.L. (2003a), Technology shocks and problem-solving capacity, in G. Donna and M. Zavodny (eds.), Technology, Growth, and the Labor Market, Kluwer Academic Publishers, Boston.

Shaw, K.L. (2003b), The human resources revolution: Is it a productivity driver?, in J. Adam, J. Lerner and S. Scott (eds.), Innovation, Policy and the Economy, University of Chicago, National Bureau of Economic Research, Chicago.

Shaw, K.L. (2009), Insider econometrics: A roadmap with stops along the way, Labour Economics, 16, 607-617.

Shaw, K.L. and Lazear, E.P. (2008), Tenure and output, Labour Economics 15, 710-724.

Shearer, B. (2004), Piece rates, fixed wages and incentives: Evidence from a field experiment, Review of Economic Studies 71, 513-534.

Tuor, S.N. and Backes-Gellner, U. (2010), Risk-return trade-offs to different educational paths: Vocational, academic and mixed, International Journal of Manpower 31, 495-519.

Wagner, J. (2003), Testing Lazear's jack-of-all-trades view of entrepreneurship with German micro data, Applied Economics Letters 10, 687-689.

Wagner, J. (2006), Are nascent entrepreneurs 'jacks-of-all-trades'? A test of Lazear's theory of entrepreneurship with German data, Applied Economics Letters 38, 2415-2419.

Williams, J.R. (2009), Commitment to corporate education, Working Paper, Tepper School of Business, Carnegie Mellon University.

Widmalm, F. (1999), Tax structure and growth: Are some taxes better than others?, Public Choice 107, 199-219.

Wulf, J. (2007), Authority, risk, and performance incentives: Evidence from division manager positions inside firms, Journal of Industrial Economics 55, 169-196.

# Comment on Lazear and Shaw: A personnel economics approach to productivity enhancement

Tuomas Pekkarinen[*]

In their article, Edward P. Lazear and Kathryn Shaw, both of whom have been at the forefront in creating the field of personnel economics, concentrate on the normative aspects of this line of research. Whereas much of the personnel economics literature concentrates on describing why the internal labor markets of firms are structured the way they are, the literature also provides important results on the optimal ways of designing payment schemes, how to assign workers to tasks, how to evaluate workers' performance etc. These normative results imply that firms can enhance their productivity by adopting improved human resource management (HRM) practices.

Lazear and Shaw outline these normative results in their article and show that they correspond to several large trends in the behavior of firms. Indeed, it is true that firms in the US and also in other countries seem to be adopting more and more of the kind of HRM practices that are outlined in the personnel economics literature. It also seems obvious that these HRM practices have had an impact on aggregate productivity growth although quantifying this effect is probably difficult.

An interesting question for the readers of *NEPR* is obviously whether HRM practices prescribed by the personnel economics literature are a "policy" that can be used to enhance productivity in the future. Here the answer is uncertain. We simply do not know whether the benefits of

[*] Aalto University School of Economics, Tuomas.Pekkarinen@hse.fi.

HRM innovations have already been reaped. However, it is still true that we observe a great deal of heterogeneity in the use of HRM practices across firms. That is, even within narrowly defined sectors, firms apply different sets of HRM practices. My main comment on the article by Lazear and Shaw is that the interpretation of this heterogeneity is key for assessing the future scope of productivity enhancement of HRM practices. Indeed, if HRM practices are so effective in enhancing productivity, they should be a relatively "cheap" form of production technology that should rapidly spread from one firm to another and we should not really observe that much heterogeneity in the data.

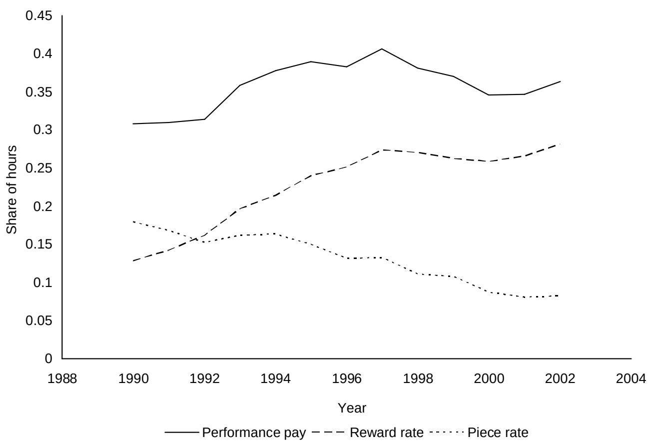
There are, broadly speaking, two ways of interpreting the variation in the use of HRM practices across firms and these have been elegantly outlined by Bloom and van Reenen (2011). The first view on the heterogeneity of HRM practices is an equilibrium approach. According to the first view, firms choose optimal policies but face different circumstances. If this is the reason for the observed heterogeneity in the use of HRM practices, there may be little scope for directly using HRM practices as a productivity enhancement policy.

However, it may also be the case that the observed variation of HRM practices across firms partly reflects different adjustment costs. According to this view, firms are trying to choose the optimal set of HRM practices but are hindered by some adjustment costs. This view provides a more optimistic perspective on the role of HRM practices as a productivity enhancement policy.

So clearly the scope of HRM practices to enhance productivity growth in the future crucially depends on the interpretation of the observed HRM practice variation. Which view is then the correct one? The truth is that we still know quite little about the reasons for the variation in HRM practices. Most of the empirical work in personnel economics has been done with data from individual firms, partly due to the natural reason that theories describe the internal organization of firms and partly due to data limitations. My view is that future research in personnel economics should devote more effort to explaining the trends and cross sectional variation in the choice of HRM practices across firms and, whenever possible, exploit data over multiple firms in doing this.

An example from the Finnish technology industry – the largest manufacturing industry in the country employing over 60 000 individuals annually – highlights the importance of understanding the heterogeneity in the choice of HRM practices, in this case the choice of incentive pay. Finnish wage register data provide detailed information on the use of different payment schemes on individual level. The collective agreement of the industry allows firms to choose between fixed time rate wages, piece rate wages solely determined by workers' individual output, and reward rate wages that are a more flexible form of performance pay where workers may be paid bonuses based on the output of more aggregate units.

**Figure 1. The use of different forms of performance pay in the Finnish technology industry 1990-2002**




Source*:* Author's own calculations based on the Wage Records of the Finnish Confederation of Industry and Employers.

Note: Hours worked under piece rates and reward rates as a share of total hours worked.

As shown by Figure 1, the use of performance pay – that is the sum of the share of total hours worked under both piece and reward rates – is pretty stable over the period 1990-2002. However, the composition of the performance pay changes dramatically. The use of piece rates declines throughout the period whereas the use of reward rates increases rapidly. It is precisely these kinds of changes we should understand if we want to assess the productivity enhancement scope of HRM practices. Clearly, the use of more flexible payment schemes is the trend in this industry but still, not all firms are adopting these payment schemes.

This is, of course, not the place for a detailed analysis of the reasons behind these changes, but simply looking at the role of firm dynamics may at least hint at a partial answer. Looking at firms that exit and enter the industry during the observation period reveals an interesting pattern. The firms that exit from the industry are using piece rate contracts more heavily than remaining firms. Also the new entrants to the industry are using reward rate contracts more heavily than the incumbent firms in the industry. These patterns may reflect some adjustment costs faced by older firms in the industry. If costly adjustment is really the reason behind the observed heterogeneity in the use of HRM practices across firms, then there may still be some scope for the spread of HRM practices to enhance productivity growth also in the future.

## References

Bloom, N. and van Reenen, J. (2011), Human resource management and productivity, in O. Ashenfelter and D. Card (eds), Handbook of Labor Economics vol 4b, Elsevier B.V., Amsterdam.

# Productivity and education: Benchmarking of elementary and lower secondary schools in Denmark

Peter Bogetoft* and Jesper Wittrup**

*Summary*

The performance of educational institutions has long been of interest to economists and politicians. This is not surprising because increasing academic skills may significantly boost economic development. School performance studies often focus on international comparisons, as in the popular PISA studies. However, we argue that national comparisons are more instrumental. Even in homogenous countries like Denmark, the ability of schools to create value for money varies considerably. Moreover, the comparability of schools and the ability to imitate peers are more apparent at a national scale. We use Data Envelopment Analysis (DEA) to evaluate schools and identify best practices. We show that even with additional restrictions on the comparisons, the potential for improvement is considerable. We also show how decompositions may support policy decisions on resource allocation and structure, and how Interactive Benchmarking IB may directly support learning and strategic planning at individual schools.

*Keywords:* Schools, performance, data envelopment analysis, multilevel models, interactive benchmarking.
*JEL classification numbers:* CO1, CO2, D24, I21.

* Department of Economics, Copenhagen Business School, CBS, pb.eco@cbs.dk.
** Danish Evaluation Institute for Local Government, KREVI, jw@krevi.dk.

A great number of studies address school effectiveness and the benchmarking of school performance. Commonly known as "school effectiveness research", these studies generally employ multilevel analysis (Goldstein, 1997). Multilevel modeling is well suited to address the problem of distinguishing the individual effects of different students (given their different abilities) from school effects and taking the impact of social variables into account.

Another important challenge to this line of research is to connect school services with the resources they employ and derive relevant policy conclusions from such connections. To this end, numerous recent studies have employed advanced benchmarking techniques, such as Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA).

Studies on school effectiveness within these two traditions (multilevel and DEA/SFA based modeling) tend to be carried out with limited or no cross-referencing to the other tradition. For example, in some DEA studies, researchers have adjusted school outputs using single level regression (Waldo, 2007a, 2007b; Borge and Naper, 2006). In cases when both approaches are addressed, they tend to be portrayed as competing or alternative analytical techniques.[1]

Our approach differs from earlier studies in that we utilize both models as complements to each other. Multilevel modeling is good at estimating school effects and the relevant confidence intervals of these estimates. DEA is good at relating performance outcomes, such as school effects, to resource use. Thus, we combine the best features of each of the models.

The potential applications of our analysis – as with most benchmarking studies – are threefold. First, the analysis can provide a general insight into the performance of an important sector of the economy and support the identification of model schools. Second, it can be used to support the allocation of resources and tasks within the educational sector and thus the structural development of that sector, e.g., school closings and school mergers. Finally, it can be used to motivate and regulate schools, either implicitly, through the publication of league tables, or explicitly, by linking budgets to performance measures. While we acknowledge all these potential applications of our research, we here focus on providing general insights and identifying model schools, or what we call a "learning perspective".

[1] See, e.g., Johnes (2006).

To examine the major issues related to schooling and highlight some important approaches to educational benchmarking, we conduct a detailed benchmarking of approximately 1 400 Danish elementary and lower secondary public schools, which are managed by 98 municipalities.

As inputs, we use time spent by teachers on teaching, time spent by teachers on other tasks (e.g., preparation for teaching), and the number of non-teaching staff. Outputs are “value added” by the schools in the sciences (Mathematics and Physics) and the humanities (Danish and English). The calculation of value added draws upon detailed data for individual students. For each student, we link achievements in standardized final examinations in ninth grade with information about the student’s parents, including parents’ income and their educational and ethnic background. We use a multilevel model to identify the impact of each school on academic achievement.

We compare an input-oriented model, where we estimate the potential for reducing inputs without reducing educational quality, with an output-oriented model, where we estimate the potential for increasing output (improving educational quality) without increasing inputs. We account for uncertainty in the estimation of educational quality by only allowing the comparison of resource use with another school when we are able to say that that other school has a better school effect with at least 95 percent certainty. By decomposing the estimated inefficiencies, we are able to identify some overall strategies for reducing inefficiency. We find that increasing average school size and changing the mix of inputs by encouraging teachers to spend more time on teaching may significantly contribute to reduce school inefficiency. The relative potential of these strategies differs between municipalities.

A large part of the estimated school inefficiency may be classified as ”pure technical inefficiency”, however, in the sense that it cannot be explained by either suboptimal school size or a suboptimal allocation of inputs. To cope with technical inefficiency, one possible strategy appears to be the adoption of a ”learning perspective”. That is, inefficient schools should learn from the practices of their efficient peers.

We have adapted our model to support a learning perspective. First, assuming that schools generally prefer to learn from other, “similar” schools, we have applied a number of restrictions on school comparisons. Thus, schools are only compared with other schools that are somewhat similar regarding the socioeconomic and ethnic background of students and the number of students with special needs.

Second, assuming that schools have different preferences, e.g., regarding the importance of different subjects or the relevance of the above mentioned restrictions, schools may perform their own interactive benchmarking, which allows them to vary relevant variables in their search for relevant peers.

We argue that the learning perspective, which is missing from most benchmarking efforts, may be a key approach in future attempts to move from theory to practice in the realm of educational benchmarking.

The outline of the paper is as follows. In the next section, we provide an introduction to the literature and we highlight our combination of two approaches, multilevel methods and DEA. In Section 2, we discuss the data used in the analysis. In Section 3, we provide the results of the multilevel analyses that are used to isolate the effects of the individual students. In Section 4, we discuss the two DEA models and the additional restrictions on relevant peers that may be introduced. Section 5 decomposes the efficiencies and discusses the strategic consequences of these decompositions. The learning perspective is further illustrated in Section 6 and Section 7 concludes the paper.

## 1. Literature

Some of the early attempts to assess the relative performance of schools in the United Kingdom used a "residual deviation score" based on regression analyses of aggregate data. Aggregate performance data (for example, average test scores for each school) were used as the outcome variable, while the independent variables included a measure of average social status. The difference between actual and "expected" performance was then interpreted as a measure of school efficiency. An important accomplishment of school effectiveness research has been to show that this use of aggregate data may provide unstable or misleading results (Woodhouse and Goldstein, 1988).

Furthermore, using a similar residual deviation score strategy on student-level data alone, and thus ignoring intra-school correlations or "clustering", may provide reasonable estimates if each school has a large number of students, but the effects of schools with only few students will be poorly estimated (Goldstein, 1997). Another consequence of ignoring clustering is the underestimation of standard errors for explanatory variables defined at the school level (e.g., school resources).

For these reasons, multilevel analytical techniques are considered as the only satisfactory approach to the assessment of school effectiveness by many people. One common result of multilevel analyses of schools is that the estimates of school effects have rather large confidence intervals, which brings into question the value of ranking schools in league tables.

Although the multilevel modeling of school effectiveness has been broadly applied, especially in the United Kingdom, and even influenced thinking within the British Department for Education and Employment at one time (Goldstein and Woodhouse, 2000), the use of this approach in Denmark has been somewhat limited. Furthermore, the best known league table of school performance in Demark, which is issued annually by the liberal think tank CEPOS, uses residuals from a simple regression on student-level data to derive its results.

Turning to the other major approach, during the last 30 years, numerous studies have taken a production economics approach and relied on state-of-the art frontier methods, most notably the non-parametric Data Envelopment Analysis (DEA) approach and the parametric Stochastic Frontier Analysis (SFA) approach. Several contributions are surveyed in Worthington (2001), Ruggiero (2004) and Johnes (2004). Here, we will focus on the DEA approach, which is applied in this study.

The pros of cons of DEA and SFA have been discussed in several textbooks, including Bogetoft and Otto (2011). The advantage of using these methods relative to key performance indicators is that they can include a description of the substitution between different resources used and the services produced. Further, they can accommodate different return-to-scale properties instead of imposing an implicit constant return-to-scale assumption, like the usual key performance indicators. This is accomplished with a multiple-input, multiple-output description of the educational entities.

Modern frontier models have typically used schools or school districts as the analytical units (DMUs), but there are also examples of the use of educational data at the individual level and on the benchmarking of nation states.

There is a wide variation in the choices of model inputs and outputs, not least because of differences in the types of data that are available to the analysts.

In terms of inputs, the data typically include the number or costs of teachers. Additionally, inputs may include indicators of other employees, the number of computers, building costs, teachers' experience and education, etc. Finally, many studies include data on students' socioeconomic background as an input, as will be further discussed below.

In terms of outputs, the existing analyses can be roughly divided into two groups. A small number of studies, including Banker et al. (2004), simply count the number of students and use measures such as the number of students at different grade levels as the output. Most studies, however, use test scores or examination results as outputs, and these outcomes are often disaggregated by subject. Finally, a small number of studies examine longer-term outcomes, such as students' later education or integration into the labor market.

There are several Nordic applications of DEA in the area of schooling, for example, articles describing DEA investigations of Norwegian (Bonesrønning and Rattsø, 1994), Swedish (Waldo, 2007b) and Finnish upper secondary schools (Kirjavainen and Loikkanen, 1998).

Several recent studies address Norwegian and Swedish primary and lower secondary schools (Borge and Naper, 2006; Naper, 2010; Waldo, 2006a, 2006b, 2007a). Because the databases available on Norwegian and Swedish schools are comparable with those available in Denmark, these studies serve as a natural source of inspiration for a Danish DEA study of public schools. Like we do in this study, they use grades that are pre-adjusted for differences in student composition (i.e., socioeconomic factors, see below) as outputs. They do not, however, use multilevel modeling to accomplish this.

The Swedish and Norwegian studies have focused on identifying various causes of differences in school efficiency at the municipal level, e.g. the political orientation of local governments. In contrast, we have mainly constructed our models with the purpose of facilitating learning at the school level.

In addition to the analysis of elementary and secondary school, a series of studies have also examined efficiency in other parts of the educational sector. For example, Daraio (2011) conducted a survey of the European university landscape based on a recent EU project, Aquameth. Additionally, as part of the Aquameth project, one of the authors of the present study examined the potential use of interactive benchmarking in an educational context (see Bogetoft, 1997).

## 2. Data

We analyze the efficiency of approximately 1 400 Danish public schools, excluding boarding schools, schools with only 10$^{th}$ grade students, and schools only dedicated to teaching students with special needs. As a rule, private schools are not part of the benchmarking exercise. We do, however, have data for approximately 330 private schools and, in some instances, we include these data in the analysis to determine their effect.

The analysis is based on detailed data at the individual student level acquired from Statistics Denmark. The data include information about the marks attained on the 9$^{th}$ grade exit examinations. We restrict the analysis to grades from state-controlled examinations that employ external examiners because these grades are expected to be more comparable. The following grades are included:

- Danish reading
- Danish oral proficiency
- Danish penmanship
- Danish spelling
- Danish writing
- English oral proficiency
- Science oral proficiency
- Mathematical literacy
- Mathematical problem solving

Whenever grade point averages are used, the different subjects are weighted according to the hours provided to each subject by the Ministry of Education. The marking scale is a 7-point scale consisting of five marks designating a passing level (12, 10, 7, 4 and 02) and two marks designating a non-passing level (00 and -3).

In addition to information about the students' grade, we have been granted access to detailed information about each student's socioeconomic background by Statistics Denmark. This includes information about parental income, education and employment, student ethnicity, and a series of other variables.

In the benchmarking models, where we link resource usage with school outcomes, we have access to information about resource usage. These data are provided by The Danish IT Centre for Education and Research, UNI-C.

In general, we use data from 2007-2009, and we examine average resource usage and other results across these years. Thus, we eliminate some effects of random variations in resources and students over time.

## 3. Student background and multilevel analysis

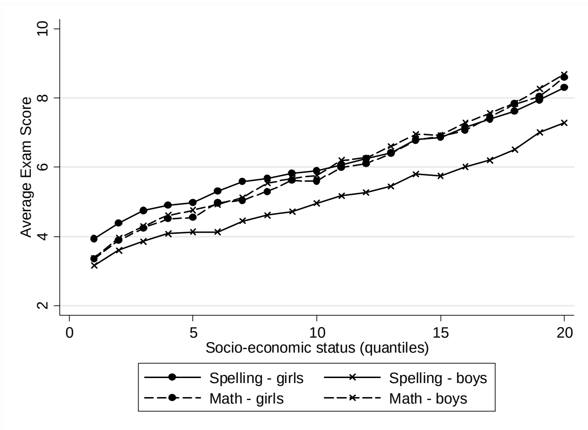
It is generally acknowledged that socioeconomic factors, which the schools cannot control, have a considerable impact on school outcomes. Students' grades do not only depend on their school but also on, for example, the educational, ethnic and economic characteristics of the parents. An important discussion in any efficiency analysis therefore pertains to how to account for such differences.

This issue is not unique to school models. For example, the patient mix is an important determinant of the resource needs of a hospital and the climate and topology of an electricity network are important determinants of the cost of providing distribution services. Several methods can be used to account for such non-controllable, non-discretionary factors, and these are discussed in general benchmarking textbooks like Bogetoft and Otto (2011). Here, we will specifically focus on a set of approaches that have been suggested in relation to school models.

There are at least three general approaches to correcting for socioeconomic conditions. One can make adjustments before, during or after the benchmarking. In the first case, the focus is on constructing inputs and outputs for benchmarking that are adjusted for individual socioeconomic conditions. That is, instead of using actual grades as outputs, we can use the difference between actual and expected grades given each individual's socioeconomic condition. Another approach is to include constraints in the benchmarking analysis such that a given school is only compared to schools with similar or more difficult socioeconomic conditions. The final approach is to benchmark across socioeconomic conditions and construct adjusted efficiency measures after the benchmarking by investigating how much of the inefficiency that can be explained by differences in socioeconomic conditions.

In this paper, we use a combination of all three approaches. The main adjustments are performed ex ante, but we include additional constraints in the search for relevant peers during the benchmarking. Additionally, we conduct an ex post investigation of the results in an attempt to derive insights that are relevant to policy.

**Figure 1. Examination score as a function of mother's income**




Source: Own calculations.

We will now explain the ex ante adjustment of grades to account for socioeconomic conditions.

Numerous studies have shown that students' social characteristics have a strong explanatory power with regard to student test scores. Therefore, if a school has high average test scores, this does not necessarily imply that it delivers exceptionally high-quality teaching, but may simply reflect that its students have parents with good incomes and high levels of education. Thus, to benchmark school performance, we must adjust for the students' social characteristics.

Figure 1 shows the final student examination scores for boys and girls in Math (problem solving) and in Danish spelling in $9^{th}$ grade as a function of an indicator for their socio-economic background (divided into quantiles). We have constructed the socio-economic index based upon detailed information about parents' income, educational level, the prestige score (ISEI score) associated with parents' job titles, as well as a number of other relevant variables.[2] We find that a student with a strong socio-economic background, on average, achieves a substantially higher examination score than a student with a weak socio-economic background. There is also a gender effect, with girls outperforming boys in Danish spelling.

In our analysis, we also included a number of other variables that have a significant impact on student test scores. This includes information about ethnic origin, parents' marital status, age of parents and student age.

We use a multilevel model to estimate school effects. The model assumes that student examination scores can be explained as a linear function of the socioeconomic variables (parent income, education, etc.), a student effect (depending upon individual abilities) and a school effect. Residuals at both levels are assumed to follow normal distributions with zero means. In mathematical terms, we may use the following model to explain the examination score, $y_{ij}$, for student $i$ at school $j$:

$$y_{ij} = \alpha + \beta x_{ij} + u_j + e_{ij}.$$

The model's first ("fixed") part $\left(\alpha + \beta x_{ij}\right)$ describes the examination score as a linear function of the student's social characteristics, $x_{ij}$. The model's other ("random") part $\left(u_j + e_{ij}\right)$ divides the additional variance into two parts: Between-school variance, $u_j$, based on deviations of school means from the overall mean, and within-school between-student variance, $e_{ij}$, based on individual deviations from school means.

[2] A detailed list of the variables is available from the authors upon request. They are also described (in Danish) in KREVI (2011).

**Figure 2. Multilevel modeling**

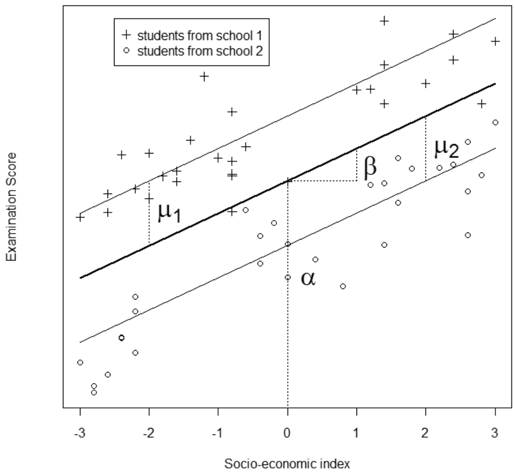


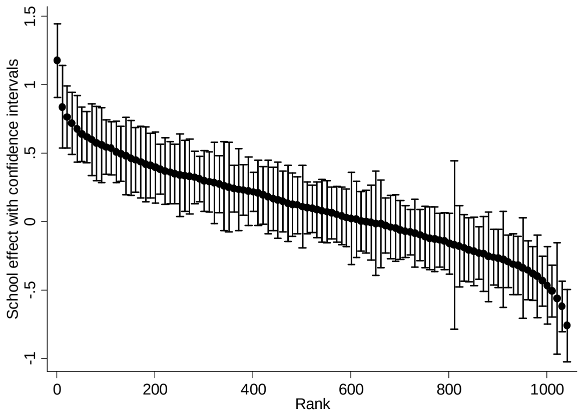

Figure 2 presents a simplified illustration of a school effect. For simplicity, this example only includes one socioeconomic variable, *x*. The bold line represents the examination score estimate derived from the fixed part of the model. For school 1, examination scores are distributed around the upper line such that the school can be said to have a positive school effect on scores of size $u_1$. In contrast, examination scores at school 2 are distributed around the lower line such that the school is said to have a negative effect of size $u_2$.

In addition to taking into account all individual socioeconomic characteristics mentioned in the appendix, our model also separates school effects from peer group effects. A number of studies have documented the existence of peer group effects in Danish schools (Andersen and Thomsen, 2011). The peer group effect implies that students' examination scores also depend upon the characteristics of the students' classmates. Students with a weak socioeconomic background tend to fare better if they are in a school with many students from a higher socioeconomic background. Thus, it should be assumed that the high examination scores that are found in schools with many students with a high socioeconomic status are partially caused by peers rather than by the school alone.

Our model estimates that approximately 28 percent of the variance in examination scores can be attributed to socioeconomic characteristics. Of the remaining variance, 93 percent are caused by variation in individual student characteristics (e.g., differences in abilities), and 7 percent of the variance are related to school effects. Overall, the impact of the school is significant.[3]

A major advantage of multilevel modeling is, as mentioned earlier, that it provides a more accurate assessment of the uncertainty of school effect estimates. Figure 3 shows the confidence intervals for the school effect estimates for a number of public schools in order of the effect estimate. For illustrative purposes, only the confidence intervals of one-tenth of the schools are included in the figure.

**Figure 3. School effects with confidence intervals**




Source: Own calculations.

The confidence intervals are constructed such that if the intervals of two schools do not overlap with at least 95 percent certainty, we will be able to conclude that their school effects are different.[4] Thus, we cannot be absolutely sure that the school with the highest estimated school effect indeed has a better school effect than the school ranked as, for example, number 40. This demonstrates that league tables are not an effective

[3] See KREVI (2011) for more details.

[4] See Schenker and Gentleman (2001) or Payton et al. (2003).

method of ranking schools. However, we are also able to conclude with great certainty that many schools have different school effects.

## 4. Resource usage and DEA models

In the previous section, we analyzed how students' performances depend on socioeconomic characteristics, and how to isolate the impact or value added of different schools. We will now link the school impact with the resource usage to evaluate the efficiency of different schools. A school is efficient if one cannot reduce the usage of one input without increasing the usage of some other input or reducing one or more outputs.

There are many ways of setting-up multiple inputs-multiple output benchmarking models. First of all, there are many ways of specifying the multiple inputs and outputs, and second there are many ways of estimating the empirical relationship between inputs and outputs. We suggest two model specifications that we consider to be of particular relevance in applications. Moreover, we suggest using Data Envelopment Analysis (DEA) to estimate the relationship between inputs and outputs.

DEA estimates a best practice technology from the actual observations of the inputs used and outputs produced in a group of schools using a minimal extrapolation principle. It finds the smallest set of input-output combinations that 1) contains the actual observations, and 2) satisfies some general properties of possible production sets. The base model, often referred to as the VRS (variable returns to scale) or BCC (Banker et al., 1984) assumes free disposability of inputs and outputs and convexity of the set of feasible input-output combinations. The first DEA model proposed, the CRS (constant returns to scale) or CCR (Charnes et al., 1978, 1979) model, in addition presumes constant returns to scale. The models used in this study are variants of these classical DEA models.

Technically, the estimation is done using mathematical programming. If we consider an analysis of $n$ schools transforming $M$ inputs, $x=(x_1,...,x_M)$, into $S$ outputs, $y=(y_1,...,y_S)$, then according to the VRS model the input-based Farrell efficiency, $E_i$, for school $i$ can be calculated as a solution to the following linear programming problem:

$$\min_{E_i,\lambda_1,\ldots\lambda_n} E_i$$

Subject to:

$$\sum_{j=1}^{n} \lambda_j y_{sj} \qquad s = 1,\dots,s$$

$$\sum_{j=1}^{n} \lambda_j x_{mj} \le E_i x_{mi} \qquad m = 1,\dots,M$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \ge 0 \qquad j = 1,\dots,n$$

It calculates the largest contraction of all inputs such that we can still find a convex combination of schools that produce at least the same outputs with at most the contracted inputs.

In the CRS model, the convexity constraint $\lambda_1+\lambda_2+\dots+\lambda_n=1$ is removed since any convex combination can be scaled up and down.

In an output oriented model, we remove *E* and instead multiply *F* on the output vector $y_j$ and maximize this. The interpretation of *F* is then as the largest proportional increase in all outputs that is feasible with at most the given inputs.

Our first model addresses the question: How much will a given school be able to reduce its resource use without having a negative impact on student achievement? To answer this question, we identify "peers" of the school that are characterized by delivering better or equivalent results while spending fewer resources. Furthermore, we might require that peer schools do not differ too much from the given school with regard to important variables, e.g., students' socioeconomic status.

With this purpose in mind, we apply the model (model 1) depicted in Table 1. The model uses a number of inputs: time spent by teachers on teaching; time spent by teachers on other tasks (e.g., preparation and meetings), and time spent by non-teaching staff. All inputs are measured in terms of full-time equivalents[5] and, for each school, the number is calculated as an average over a three-year period (2007-2009).

As outputs, in this model we use the number of students in each of three age groups: 0 to 3rd grade; 4th to 6th grade; and 7th to 10th grade. The

[5] 1 924 working hours equals 1 full-time equivalent.

model then identifies efficient schools or peers that are able to optimize the use of inputs in relation to the production of outputs.

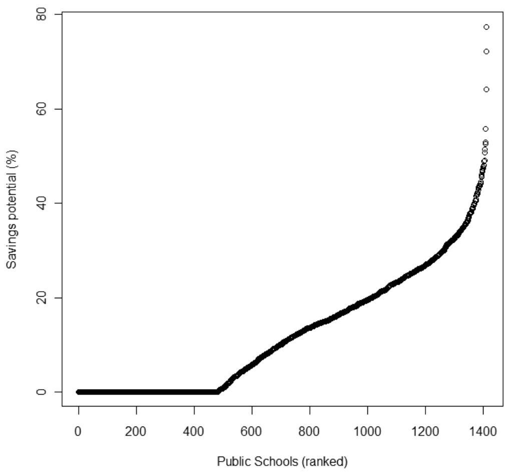
**Table 1. Model 1 – focus on savings**

| | |
|---|---|
| Inputs | 1. Number of teacher hours spent on teaching<br>2. Number of teacher hours spent on non-teaching activities<br>3. Number of working hours by other staff |
| Outputs | 1. Number of students in 0-3$^{rd}$ grade<br>2. Number of students in 4$^{th}$-6$^{th}$ grade<br>3. Number of students in 7$^{th}$-10$^{th}$ grade |
| Period of time | All inputs and outputs are taken as an average over the period from 2007 to 2009 |
| Restrictions | A school may only be compared to another schools if:<br>1. If this other school with at least 95 percent certainty has the same or a higher (more positive) general school effect<br>2. The difference between the average scores on the standardized socio-economic index for the two schools is less than one standard deviation<br>3. The difference between the scores on a standardized index for the share of students with a non-Danish origin for the two schools is less than one standard deviation<br>4. The difference between the scores on a standardized index for the share of students with special needs is less than one standard deviation |
| Other | The model is input-based and the base model presumes CRS |

To take quality into account, we apply some additional restrictions on the DEA comparisons. To be considered as a peer for another school, we require that the peer school's general (average over all subjects) school effect is better with at least 95 percent certainty. Thus, we use the estimates and the confidence intervals provided by the multilevel model. In this manner, we ensure that apparent peer school efficiency in resource use is not the result of lower educational quality.

Furthermore, we require that the peer schools of a given school should not differ too much from that school with regard to the social composition of its students. Therefore, we establish standardized indexes for students' socioeconomic status; the share of students with a non-Danish background; and the share of students that receive extra or specialized teaching because they are considered to have special needs. For each of these indexes, we require that peer schools do not deviate by more than one standard deviation.

**Figure 4. School saving potential (1-E) in model 1**




Source: Own calculations.

It could be argued that the restrictions with regard to the social composition of the students are not necessary because we have already taken these differences into account in estimating the school effects. It may be the case, however, that the optimal technology of education production varies across types of students. For instance, Lazear (2001) has argued that optimal class sizes are much larger for better-behaved students. Thus, since we here wish to promote learning, we use the additional restrictions based on the assumption that it will be easier for a school to learn from another school if the characteristics of their student populations are similar.

With our model, we are able to calculate the Farrell efficiency of each school. An efficiency of, e.g., 0.9 indicates that by learning from its peers, the school should be able to maintain its present school effect while reducing its resource use by 10 percent. The distribution of saving potentials ($1-E$) for the public schools is shown in Figure 4.

Approximately one-third of the schools appear to be fully efficient (input efficiency = 1). This means that it is not possible to identify any peers for these schools. A school can appear as fully efficient when it delivers a high level of teaching quality with few resources, but it can also appear as fully efficient when it has an unusual social composition of students such that there are only a few potential peers with which to compare it.[6]

At the aggregate level, our model indicates a potential for increasing overall efficiency in the public school sector by 13 percent. That is, by learning from relevant peers, inefficient schools should be able to reduce their use of inputs by 13 percent while maintaining their present educational quality.

This estimate is based on the assumption that we cannot compare public schools with private schools and that we have constant returns to scale; that is, we can compare schools of different sizes. We can vary these assumptions to obtain different overall results, as shown in Table 2. If we believe that private schools may serve as peers and provide inspiration to public schools, the total efficiency potential increases to 20 percent while it is reduced to 9 percent if we assume that schools cannot learn from other schools of a different size.

**Table 2. Saving potentials in different model variants**

| Overall savings potential with different assumptions | | Returns to scale: Constant % | Returns to scale: Variable % |
|---|---|---|---|
| Private schools part of the reference technology | Yes | 20.0 | 13.4 |
| | No | 13.1 | 9.4 |

Finally, we estimate the overall improvement potential if comparisons are restricted to schools within the same municipality. This may be a relevant restriction to consider because some local governments may have qualms about admitting that other municipalities have better performing

[6] Since the estimated technology is increasing in the number of observations, the smaller the groups, the larger the bias, i.e. the more we underestimate the improvement potential. In theory, we could correct for this by boot-strapping, e.g. Bogetoft and Otto (2011), but the results would then not have any simple interpretations as we could not point to any specific schools behind the benchmarks. From a learning perspective, we consider the bias problem to be a minor issue.

schools. This restriction on the model would reduce the overall improvement potential to less than 1 percent, however. Therefore, it is crucially important that local governments are willing to look outside their own municipality to identify peers as sources of inspiration.

To supplement model 1, which examines potential savings, we establish an alternative model 2, which estimates the potential to improve educational quality. The question addressed by this model is: To what extent can successful benchmarking improve school effects if we require that school resources remain constant?

**Table 3. Model 3 – focus on quality improvement**

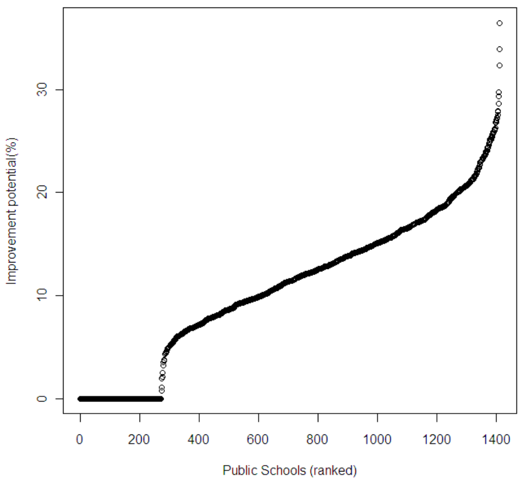
| | | |
|---|---|---|
| Inputs | 1. | Number of teacher hours spent on teaching per student |
| | 2. | Number of teacher hours spent on non-teaching activities per student |
| | 3. | Number of working hours by other staff per student |
| Outputs | 1. | School effect in humanities (Danish and English) |
| | 2. | School effect in sciences (Math and Science) |
| Period of time | All inputs and outputs are taken as an average over the period from 2007 to 2009 | |
| Restrictions | A school may only be compared to another schools if: | |
| | 1. | this other school with at least 95 percent certainty has the same or a higher (more positive) general school effect |
| | 2. | the difference between the average scores on the standardized socio-economic index for the two schools is less than one standard deviation |
| | 3. | the difference between the scores on a standardized index for the share of students with a non-Danish origin for the two schools is less than one standard deviation |
| | 4. | the difference between the scores on a standardized index for the share of students with special needs is less than one standard deviation |
| Other | The model is output-based and presumes VRS | |

It makes economic sense to improve student skills. According to an OECD estimate, an improvement of the PISA score of Danish students by one quarter of a standard deviation would result in additional economic growth valued at 586 billion USD over the next 80 years (OECD, 2010). If we make a parallel to average examination scores in the 9th grade, raising these scores by a quarter of a standard deviation would amount to an increase in examination scores (for the subjects included in this analysis) of approximately 10 percent.

The specifications for model 2 are provided in Table 3. The inputs are the same as in model 1 but are now in a ratio form. As outputs, we use the school effects estimated separately for science (Math and Science) and the humanities (Danish and English). We use the same restrictions that were applied in model 1.

According to this model, approximately one out of five schools is fully efficient. The majority of schools, however, have an improvement potential of between 5 percent and 20 percent, as shown in Figure 5. Therefore, less efficient schools should learn from their peers to obtain greater benefits from existing school resources.

**Figure 5. School improvement potential (F-1) according to model 2**




Source: Own calculations.

At the national level, model 2 indicates that benchmarking can increase student skills by 10.7 percent without increasing the resource use. This increase would enable the Danish school system to meet the milestone set by the OECD. Just as with model 1, private schools may potentially also be included as peers. In this case, the overall improvement potential would increase to 14 percent.

It should be noted here that our approach to accounting for socioeconomic variables as part of the benchmarking deviates from the approaches in previous DEA studies. Indeed, the appropriate approach to non-discretionary or environmental variables is a major issue within the DEA literature in general (Bogetoft and Otto, 2011) and in relation to the application of DEA to modeling educational production in particular (Ruggerio, 2004).

Consider the DEA model formulated as a linear programming problem above. That is, we consider $n$ schools transforming $M$ inputs, $x(x_1,...,x_M)$, into $S$ outputs, $y(y_1,...,y_S)$, and presumes variable returns to scale, VRS.

Now, let us assume that we also have $R$ environmental variables $x(z_1,...,z_R)$, which are non-discretionary in the sense that the school is not able to influence them (such as with the socioeconomic status of students). We assume here that a higher value of these non-discretionary variables represents more favorable conditions (e.g., a higher socioeconomic status). It is evident that we cannot add these variables to the model as ordinary inputs because the basic model assumes that the school can proportionally reduce all inputs.

The most common solution to this problem is probably the solution that was first suggested by Banker and Morey (1986), which added the following condition to the basic model:

$$\sum_{j=1}^{n} \lambda_j z_{rj} \leq z_{ri} \qquad r = 1,...,R$$

The intuitive appeal of this solution is that it requires school $i$ to have the same or a higher value (i.e., it is in a similar or better environmental position) as compared to its (potentially fictive) peers that define the production possibility.

As illustrated by Olesen and Petersen (2009), the above model may suffer from not taking into account differences in the size of the schools that are being benchmarked. For example, we can imagine a situation with a very large school, $A$, operating in a very favorable environment $(Z_A)$, and a very small school, $C$, operating under very unfavorable conditions *($Z_C$)*. If a third school, *B*, operates in an average environment $Z_B = ½\left(Z_A + Z_C\right)$, then Banker and Morey's model will allow the virtual combination ½*A*+½*C* to dominate *B* because this virtual combination per definition operates in a similar environment to school *B*. The problem with this is that when school *A* is much larger than school *C*, the virtual combination of *A* and *C* is almost identical to *A*. However, because school *A* operates in a very favorable environment, it might not be appropriate to compare *B* with *A*.

Dissatisfaction with the Banker and Morey model led Ruggiero (1996) to suggest that their additional condition should be replaced by the following:

$$\lambda_j = 0 \qquad \text{if } z_{rj} > z_{ri} \qquad r = 1,\dots,R$$

This solution avoids the problem that was just mentioned because it excludes any school with relatively more favorable conditions as being part of a virtual peer school. However, one problem with this model, which has been analyzed in numerous simulations (Ruggerio, 1998; Muniz et al., 2006; Waldo, 2006a; Olesen and Petersen, 2009), is that especially when more than one non-discretionary variable is involved, the model is extremely conservative. Many schools will appear to be efficient because they have a low score on one of the environmental variables.[7]

A third alternative is the two-stage model that was originally proposed by Ray (1991). The first stage of this model is to run the DEA program without environmental variables (e.g., not taking into account differences in students' socioeconomic status). The second stage is to carry out a Tobit regression[8] of the environmental variables upon the calculated efficiency:

$$E_{raw} = \alpha + \beta_1 z_1 + \cdots + \beta_R z_R + \varepsilon$$

[7] Olesen and Petersen (2009) have suggested an alternative that involves weighting the non-discretionary variables according to the size of the schools.

[8] Ray (1991) used OLS, but later applications of his basic approach more appropriately relied on a Tobit regression.

The beta values then represent an estimate of the effect of the environment and based upon this estimate, we can calculate an adjusted efficiency score from the residuals:

$$E_{raw} = E_{raw} - \alpha + \beta_1 z_1 - \cdots - \beta_R z_R$$

Although this approach is frequently applied, Barnum and Gleason (2008) empirically demonstrated that the method exhibits substantial bias and low precision and that the degree of bias and precision is affected by input variance and correlation. Furthermore, as noted by Cordero-Ferrera et al. (2008), the conventional two-stage model, and the more advanced three-stage model proposed by Ruggiero (1998), rely upon an ill-founded assumption, namely, that if output depends on non-discretional inputs, then a significant correlation will exist between these factors and efficiency. This does not have to be the case, however, because many other factors typically influence efficiency.

A simple solution to avoid this kind of bias is, as suggested by Barnum and Gleason (2008), to reverse the conventional two-stage model such that the regression of environmental variables on output is carried out before the DEA program is run with the adjusted outputs.

This final solution represents the approach we use to benchmark Danish schools, though our approach includes a few additional modifications. First, for the case provided above, we believe that there are major advantages to applying a multilevel model instead of a single level regression when adjusting outputs. Second, as noted by, e.g., Grosskopf et al. (2009), adjusting outputs for differences in the social composition of students may not allow us to ignore environmental variables in the second stage DEA because environmental variables may have an independent impact on production possibilities. Therefore, we retain additional environmental restrictions when performing the DEA.

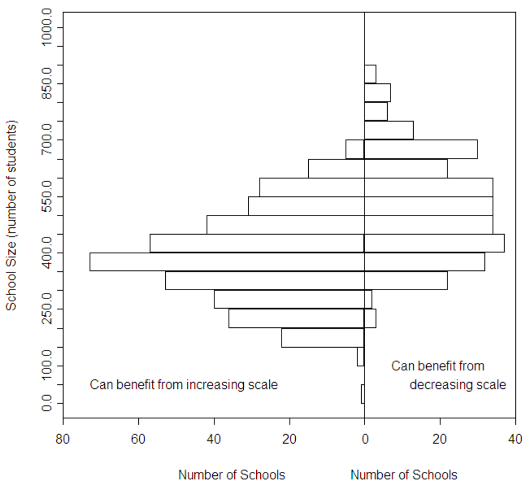
## 5. Decompositions and strategies

DEA does not only provide an assessment of the overall improvement potential within the school sector but may also be used for identifying potential improvement strategies. In the public debate on schools, issues related to school size and time spent teaching have featured prominently, so therefore we briefly focus on these topics.

As mentioned during the discussion of model 1 above, with regard to this model, we have the option to choose between an assumption of variable returns to scale (VRS) or of constant returns to scale (CRS). We may interpret the difference in results between these two models as representing a certain "scale inefficiency." If a school is efficient when we assume variable returns to scale but inefficient when we assume constant returns to scale, the current size of the school may be an obstacle to maximum improvement.

Overall scale inefficiency is 3.7 percent in model 1. Thus, one-third of the total savings potential in the CRS model (13 percent) is related to changes in school size. Furthermore, we can conclude that the majority of scale inefficient schools are too small.[9]

**Figure 6. Scale inefficiencies**




Source: Own calculations.

There is no simple method for characterizing optimal scale size, however. For two schools of similar size in terms of number of students, one

---

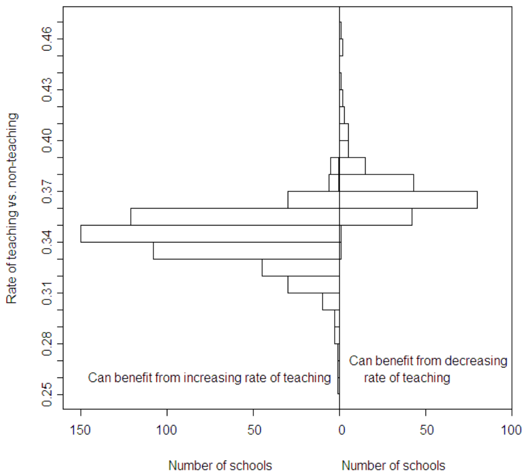
[9] We apply the method proposed by Färe and Grosskopf (1994) to determine whether a given scale inefficient school is either too small or too large.

may be too small while the other may be too large. Optimal scale appears to be conditional on the composition of resources, outcomes and students' socioeconomic characteristics. Figure 6 shows the distribution of "too small" and "too big" schools according to school size.

We find that schools with between 500 and 600 students strike an approximate balance between those that are deemed too small or too large. A cautious rule of thumb could therefore be that a reasonably efficient school size lies in that range. Because there are many small schools and few large schools in the analysis, however, we cannot rule out the possibility that the model underestimates production possibilities for large schools. The rule of thumb may then be that the efficient school size is above 500 students.

Another important debate involves the amount of time teachers spend on teaching as compared to time spent on other activities, such as preparation and meetings. Based on model 2, we calculate the allocative inefficiency as greater than 2 percent. Thus, schools might be able to improve student examination scores by 2 percent (a fifth of the total potential) if they optimize the allocation of inputs.[10]

**Figure 7. Allocative inefficiencies**




[10] Allocative efficiency is here calculated as the input efficiency in a model with aggregated hours compared to the efficiency in the model distinguishing between the three types of hours.

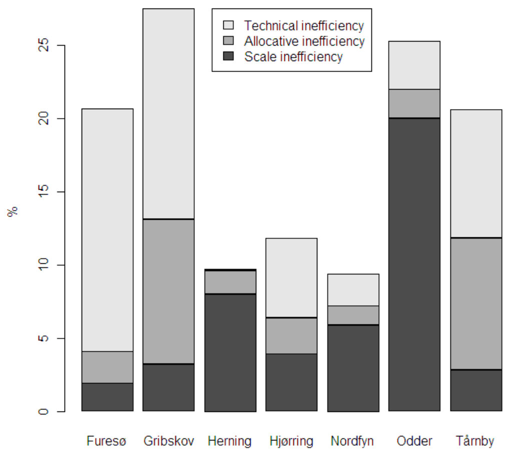
Source: Own calculations.

As might be expected, the model indicates that the vast majority of allocatively inefficient schools may improve their efficiency by increasing teacher time spent on teaching. More teaching and less teacher time spent on other tasks will generally increase school efficiency.

Just as with the discussion of scale, this finding does not allow us to determine a fixed optimal mix of teaching and non-teaching time. The above figure shows the distribution of schools that would benefit from increasing and decreasing teaching time according to the balance of teaching and non-teaching. We find no real support for the idea that a teaching rate below 36 percent is conducive to educational efficiency.

Besides casting some light upon the issues of scale size and rate of teaching, decomposition of inefficiency may serve to point out the most attractive policies for efficiency gains for each municipality. Figure 8 shows aggregated scale inefficiency, allocative inefficiency and technical inefficiency, calculated on the basis of our model 1, for seven Danish municipalities.

**Figure 8. Decomposition of school inefficiency for seven Danish municipalities**




Source: Own calculations.

The differences indicate that these municipalities should adopt different strategies in order to improve school efficiency. The municipalities of Odder and Herning might appropriately choose to focus on school size. Both these municipalities on average have schools with relatively few students. On the other hand, in particular allocative efficiency ought to be put on the agenda in the municipalities of Gribskov and Tårnby. In these two municipalities, teachers appear to spend less time teaching than in most other municipalities.

In many municipalities, however, a large part of the estimated school inefficiency may be classified as ”pure technical inefficiency”, in the sense that it cannot be explained by either suboptimal school size or a suboptimal allocation of inputs. To cope with technical inefficiency, an obvious strategy is for the schools to try to learn from the practices of their efficient peers. Facilitating such a learning perspective, as well as more school-specific analyses of scales and scope, is the subject of our next section.

## 6. Interactive benchmarking

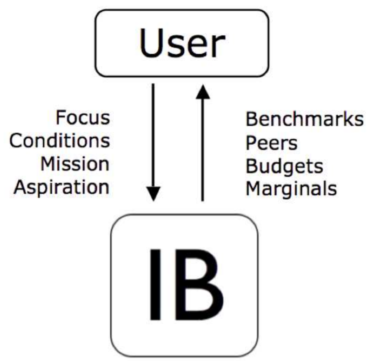
The DEA analyses indicate potential gains from school adjustments in terms of best practices, scale and the mix of resources and activities. To guide decision-making at individual schools, however, we need to provide more specific insights, and different methods exist to extract such information from a benchmarking model.

To eliminate technical inefficiency, schools must learn best practices. In the DEA approach, this is supported by the identification of a small number of peers from which a school can learn. The peers can be calculated using various standard software packages. For example, all DEA calculations in this report are performed using the free, open source R-framework found at www.r-cran.org and the “Benchmarking” package developed by Bogetoft and Otto (2010) to support a large number of benchmarking tasks, including the identification of peers.

The benchmarking model, however, allows schools to learn much more than potential improvements and their best practice direct peers. Interactive benchmarking supports individual learning based on a benchmarking model. This type of benchmarking was first proposed in Bogetoft and Nielsen (2005) and Bogetoft et al. (2006). An application to university benchmarking is described in Bogetoft et al. (2007). Several other applications are described at www.ibensoft.com. One software package that is based on this approach is Interactive Benchmarking IB®. We will illustrate some of the useful information that can be extracted from a benchmarking model using this software. This is particularly interesting because the Danish Evaluation Institute for Local Government (KREVI) plans to make the analyses in this paper available to individual users in this manner.

The idea of interactive benchmarking is to use a benchmarking model as a *learning lab*. Modern benchmarking estimates comprehensive multiple-input, multiple-output relationships derived from actual practices. The complexity that can be captured by such multiple-input, multiple-output benchmarking models vastly exceeds the complexity captured by mental models and textbook examples. To make these possibilities available to users, interactive benchmarking combines benchmarking techniques with multi-criteria decision theory, specifically with the multi-criteria approach that is generally known as the iterative articulation of preferences and possibilities, cf. Bogetot and Pruzan (1991). This is illustrated in Figure 9 below.

**Figure 9. Interactive Benchmarking IB**




In an interactive benchmarking session, the user (e.g., manager, headmaster, superintendent) selects a focus (model) for the analysis. The focus can be short run or long run, and it can involve the entire school or some parts of the school. Ideally, the model captures the relevant conditions because it accounts for inputs, outputs and environmental variables. However, the user can make supplementary assumptions about the analyzed school and its relevant peers. The evaluated unit can be a realized unit, a budgeted unit, a potential merger, etc. Similarly, comparisons with some other schools can be excluded using filters on the allowed peers. The user may, for example, only be interested in comparisons with local schools of a similar size. The user's specific strategy can be further specified by defining search directions. These directions reflect the user's interest in saving the different inputs (resources) and in expanding the different outputs (products and services). The aspiration level and performance of others can also be examined. Although best practices are of particular interest, the user may strive for less, e.g., 25 percent of best practices.

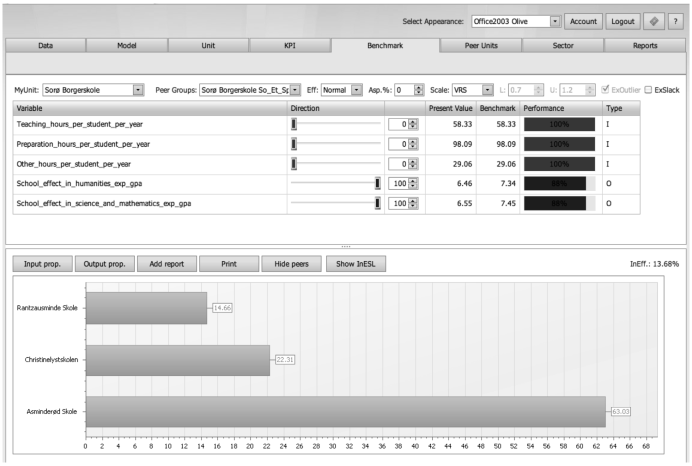
Using a benchmarking model as a learning lab allows the user to easily support a series of operational, tactical and strategic analyses. Among other things, the user can use the benchmarking model to:

- Budget and set reasonable targets that reflect the user's strategy
- Account for differences in responsibilities
- Find peers to imitate based on the idea of learning from the best
- Evaluate the school's strategic position
- Find marginal costs and trade-offs to support marginal decisions
- Optimize scale and scope
- Analyze non-marginal decisions about outsourcing and insourcing activities, and about collaborating or possibly merging schools

We will now illustrate some of these applications.

To do so, consider a given school, here Sorø Bogerskole. The central benchmarking screen for this school in a model with the inputs and outputs of model 2 is illustrated in Figure 10 below. In this illustration and the following, we have not taken into account the uncertainty of the school effects. The present situation for the school is summarized in the "MyUnit" column. For pedagogical reasons, we use hours per students, i.e. a simple rescaling of the original fte per students, as inputs. A best practice combination of other schools leads to the input and output values in the "Benchmark" column, which correspond to a target budget. For example, we see that with the given budget, i.e. without changing the hours spent by teachers on teaching, by teachers on non-teaching activities and by other staff, the grades could be increased from 6.46 to 7.34 in the humanities and from 6.55 to 7.45 in Science and Mathematics.

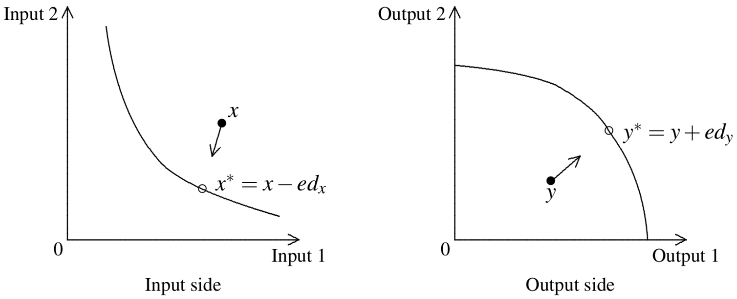
**Figure 10. Benchmarking of Sorø Borgerskole**




Source: Own calculations.

The benchmark budget depends on the presumed strategy. The strategy expresses in general terms the goals of the school and how it plans to accomplish them. The school may be particularly interested in reducing some parts of its cost base and expanding some of its services. In a benchmarking context, we can consider this as the improvement direction. Strategic prioritizing between saving on different inputs and expanding different outputs is controlled by the "Direction" handles. These handles provide numbers, usually between 0 and 100, that express how eager the user is to save on the associated inputs or to expand the associated output. The use of directional efficiency is illustrated in Figure 11 below. The direction $d\left(d_x, d_y\right)$ interacts with the present situation $\left(x, y\right)$ to form the benchmark. For the user of benchmarking, the "Direction" can be viewed as a wheel that is used to steer the school in different directions to produce different projections on the efficient frontier. The general idea is that the more an input or output dimension is emphasized, the more we save (expand) the associated input (output). Note also that the traditional Farrell efficiencies that we relied on in the empirical analyses above correspond to special cases of the directional approach. Input efficiency corresponds to the use of $d = (x, 0)$ and output efficiency corresponds to the use of $d = (0, y)$.

**Figure 11. Directional efficiency approach**




Therefore, the important managerial task of linking strategy with budgets

Strategy → Budget

can be supported by a benchmarking model. The user manipulates the direction to express his strategy and the benchmarking system then calculates the resulting benchmark budget

Strategy → Direction → Benchmark → Budget.

It is important to understand that this particular budget takes into account the school's particular historical mix of resources and services, and the environmental constraints it might face to the extent that these aspects have been modeled. Moreover, the budget is derived using all available information about the complex multiple-inputs, multiple-outputs relationships as they have been estimated using data from several schools, and it takes into account the strategy and preferences of the school in question.

One advantage of such benchmark-based planning and budgeting is that it is relatively easy to find the relevant targets because we can directly explore the strategy-budget relationship. That is, managers can experiment with different goals and strategies and get immediate feedback on the likely consequences. This means that the iterative process of trying different strategies, predicting their consequences, modifying the plan and updating the consequences is simplified when using a benchmarking framework.

Another common issue in applications is to take into account different conditions without making managers responsible for aspects they cannot control. This can also be supported by the choice of direction. The environmental factors that may either facilitate or complicate production, e.g., a high socioeconomic index or a required minimum number of teaching hours, can be included as non-controllable inputs or outputs by letting the corresponding dimensions of the direction vector equal zero.

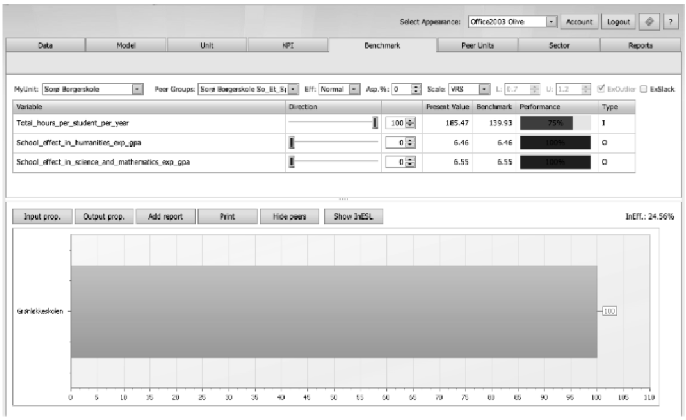
One of the advantages of non-parametric benchmarking models like DEA is that they provide explicit peers to learn from, as discussed earlier in the paper. In Figure 10, the peers are illustrated in the lower part of the screen shot. In particular, we see that Sorø Borgerskole should try to learn from Rantzauminde Skole, Christinelystskolen, and Asminderød Skole. We can also see the relative importance of these peers from the length of the horizontal bars.

It is noteworthy that the peers take into account the characteristics of the evaluated school. Specifically, and assuming that we do not use negative directions, the combination of peers produces more of each output using less of each input than the evaluated school. This explains why peers often seem intuitively sensible to industry. The peers also take into account the strategy that is reflected by the "Direction." If we change the "Direction", the relative importance of the peers will change and some new peers will become relevant

In practical applications of benchmarking, peers are studied with interest by managers and researchers alike. First, the relevance of the benchmarking approach is intuitively evaluated by the relevance of the peers. Moreover, practical managers may have additional requirements that they want to impose on the comparisons. There can be many reasons for this. Some of these reasons are rational, for example, comparing one school with another school that operates under a similar regulation because this makes their conditions more comparable. Other reasons are more emotional or at least based on "softer" arguments. A manager may not trust the data from some schools or may already have established good relationships with the managers of some other schools and may thus be particularly interested in learning from these. From a learning perspective, it is acceptable to introduce additional restrictions if the managers recognize that potentially attractive learning possibilities are foregone when such restrictions are introduced. In Interactive Benchmarking IB, additional restrictions on potential peers are easy to introduce. One possibility is to define filters on the set of potential peers. Another is to eliminate peers that are suggested by the mathematical optimization but that the user dislikes. In the software, this is done with a single click on the horizontal bar that shows the importance of a peer. A new set of peers will be calculated based on the remaining set of potential peers.

If we are mainly interested in the cost savings and want to keep the school responsible for allocation of the total time on teaching, preparation and other time, we could move to a simple model with one input only, total hours used per student per year as in Figure 12 below. We see that by adjusting to best practice, Sorø Borgerskole is now estimated to have a saving potential of 24.6 percent and should learn from Grønløkkeskolen.

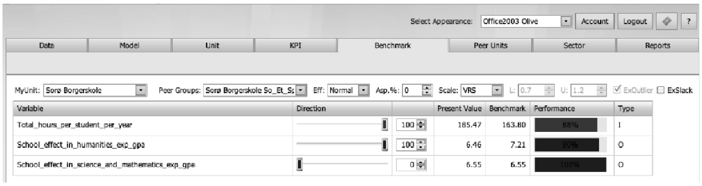
**Figure 12. Sorø Borgerskole with cost saving objective**




Source: Own calculations.

If the manager is also concerned about the exam grades in the humanities in particular, he could try to lower the hours spent and, at the same time, increase the school effect in the humanities. The outcome is shown in Figure 13 below. Note that increasing the average grade from 6.46 to 7.21 comes at the marginal costs of increasing the benchmark hours from 139.4 to 163.8 per student per year.

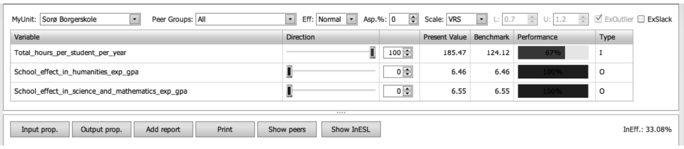
**Figure 13. Cost saving and increased school effect in humanities**




Source: Own calculations.

If the manager wants to extend his search for interesting peers, he can make a comparison for example to all peers, as shown in the Peer Group field. In this case, the saving potential increases to 33 percent, cf. Figure 14.

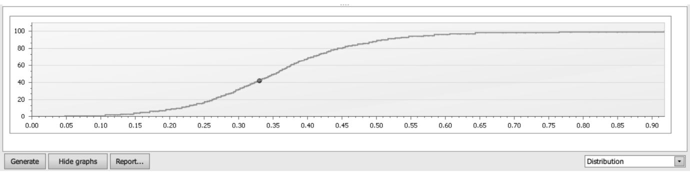
**Figure 14. Extending the search for peers**




Source: Own calculations.

This may seem like a dramatic saving and it may therefore be interesting to make a parallel analysis of all the schools under the same conditions. The cumulative distribution from such a sector analysis is shown in Figure 15 below. The graph shows that Sorø Borgerskole is ranked among the 33 percent most efficient schools.

**Figure 15. Sector analysis**




Source: Own calculations.

Interactive benchmarking can also support more structural decisions. It is easy for a user to evaluate the scale efficiency of a given school. As discussed above, the optimal scale depends on the composition of the resources, outcomes and environmental factors, including students' socioeconomic characteristics. This means that there is no simple answer to the question of optimal scale size. Rather, the specific characteristics of the school, its environment and its strategy must be taken into account, which is exactly what the interactive benchmarking approach allows the user to do. Moreover, the user can easily change the so-called returns to scale assumption in this framework.

Although the idea of optimal scale is useful, the concept may not fully reflect the reality faced by a manager, who may not be able to adjust to optimal scale size. If a town has two schools and both are too small, we cannot simply increase the size of each school because there are too few students to do so in practice. The available option may be to merge the two schools. This action may lead to a school that is too large, but the losses from the merger may be smaller than the benefits of increasing the size of the original schools. To support such evaluations of more specific options for restructuring, one may once more use benchmarking, cf. Bogetoft and Wang (2005). In fact, the Interactive Benchmarking IB® software contains a mechanism for evaluating this option by simply selecting which schools to merge and then analyzing the potential merger. Similar facilities are available in the R package "Benchmarking". By comparing the costs and results of the tentatively merged school with the individual costs and services of the original schools, we can obtain a specific evaluation of the actual possibilities that are available in a given town. Other possibilities can also be analyzed. If, for example, one would like to consider outsourcing some services, it is possible to calculate the costs with and without those services. If the difference exceeds the payment to the provider of the outsourced services, then outsourcing is economically attractive.

These few examples suffice to demonstrate how individual managers, headmasters, superintendents and politicians can use the benchmarking model to support a series of managerial decisions. The individual learning perspective stands in sharp contrast to the use of benchmarking to provide league tables. While the latter often – and sometime rightfully so – creates organizational resistance to benchmarking since they serve as public pillory, the learning perspective may be of interest to all parties involved.

## 7. Conclusion

Public schools are not all alike. The social composition of the student body varies substantially. Similarly, school size varies, schools prioritize the use of teachers' time differently, and schools have different strengths and weaknesses in teaching different subjects.

Given these differences, a fair and useful benchmarking system of school performance that takes into account the use of resources must incorporate many different aspects of schools. In this analysis, DEA is applied in combination with multilevel modeling as a tool for "realistic benchmarking", which is designed to identify relevant peers for each public school. These peers should be very similar to the given school but should perform better. This approach can strengthen the efforts for school improvement by identifying the most relevant sources of inspiration.

The major causes of variation in examination scores are individual-level student characteristics, such as abilities, motivation, and socioeconomic background. However, differences with regard to school quality also play a significant role. Therefore, there is reason to consider how less efficient schools can learn from the best schools. In addition, the considerable costs of school resources make it relevant to optimize resource use.

With this purpose in mind, this paper introduced two models. The first model was designed to answer the question: How much can public schools save without reducing quality, when they learn from the most relevant peers? We found that by applying realistic benchmarking, public schools could save approximately 13 percent compared to the current level of spending on school resources.

The other model addressed the problem: How much can public schools that are inspired by learning from relevant peers improve student skills without increasing the use of resources? We found that realistic benchmarking, which disseminates best practices among schools, ought to result in an increase in student skills by more than 10 percent. Based on the OECD model for estimating the economic impact of improved student skills, the value for society of this increase would be more than 3 000 billion Danish kroner over 80 years.

In relation to both models, we have examined the relevance of various general strategies for improving school efficiency. In some municipalities, changing school size has a large savings potential. In other municipalities, there is a major potential for improvement related to a better allocation of resources such that teachers spend relatively more time teaching. In this way, the model may allow managers to determine priorities among the various strategies.

To support managerial decisions, it is also useful to "allow the manager back in". Interactive benchmarking that combines multiple criteria decision making with benchmarking supports this. The benchmarking model serves as a learning lab that the manager can use to estimate the potential impacts of alternative decisions, both marginal and non-marginal ones. In particular, using a directional distance approach, the manager can easily experiment with different preferences and plans to see the possible consequences.

There is not much doubt that detailed data about school performance will increase in the future. If we are to benefit from the increased data, there is a need for strong tools to address multiple performance dimensions. By combining DEA with multilevel modeling, we have introduced such a tool and shown some of its applications. As stated by a recent government task force on public schools, if our public schools are to be world class, decisions have to be made; money and resources have to be reallocated; and functions, competencies and institutions have to be cut, moved or rebuilt. Advanced benchmarking techniques may help ensure that we make the appropriate choices.

## References

Andersen, S.C. and Thomsen, M.K. (2011), Policy implications of limiting immigrant concentration in Danish public schools, Scandinavian Political Studies 34, 27-51.

Banker, R.D., Charnes, A. and Copper, W.W. (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, Management Science 30, 1078-1092.

Banker, R.D., Janakiraman, S. and Natarajan, R. (2004), Analysis of trends in technical and allocative efficiency: An application to Texas public school districts, European Journal of Operational Research 154, 477-491.

Banker, R.D. and Morey, R.C. (1986), Efficiency analysis for exogenously fixed inputs and outputs, Operations Research 34, 513-521.

Barnum, D.T. and Gleason, J.M. (2008), Bias and precision in the DEA two-stage method, Applied Economics 40, 2305-2311.

Bogetoft, P. (1997), DEA-based yardstick competition: The optimality of best practice regulation, Annals of Operations Research 73, 277-298.

Bogetoft P., Bramsen, J.M. and Nielsen, K. (2006), Balanced benchmarking, International Journal of Business Performance Management 8, 274-289.

Bogetoft, P., Fried, H.O. and Vanden Eeckaut, Ph. (2007), The university benchmarker: An interactive computer approach, in A. Bonaccorsi and C. Daraio (eds.), Universities and Strategic Knowledge Creation, Edward Elgar Publishing, Cheltenham.

Bogetoft, P. and Nielsen, K. (2005), Internet based benchmarking, Journal of Group Decision and Negotiation 14, 195-215.

Bogetoft, P. and Otto, L. (2010), Benchmarking with DEA and SFA, R package, available from www.r-cran.org.

Bogetoft, P. and Otto, L. (2011), Benchmarking with DEA, SFA, and R, Springer, Berlin.

Bogetoft, P. and Pruzan, P. (1991), Planning with Multiple Criteria: Investigation, Communication, and Choice, North Holland, Amsterdam.

Bogetotf, P. and Wang, D. (2005), Estimating the potential gains from mergers, Journal of Productivity Analysis 23, 145-171.

Bonesrønning, H. and Rattsø, J. (1994), Efficiency variation among the Norwegian high schools: Consequences of equalization policy, Economics of Education Review 13, 289-304.

Borge, L. and Naper, L. (2006), Efficiency potential and efficiency in Norwegian lower secondary schools, FinanzArchiv 62, 221-249.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units, European Journal of Operational Research 2, 429-444.

Charnes A. Cooper, W.W. and Rhodes, E. (1979). Short communication: Measuring the efficiency of decision making units, European Journal of Operational Research 3, 339.

Cordero-Ferrera, J.M., Pedraja-Chaparro, F. and Salinas-Jiménez, J. (2008), Measuring efficiency in education: An analysis of different approaches for incorporating non-discretionary inputs, Applied Economics 40, 1323-1349.

Daraio, C. et al. (2011), The European university landscape: A micro characterization based on evidence from the Aquameth project, Research Policy 40, 148-164.

Färe, R. and Grosskopf, S. (1994), Estimation of returns to scale using data envelopment analysis: A comment, European Journal of Operational Research 79, 379-382.

Goldstein, H. (1997), Methods in school effectiveness research, School Effectiveness and School Improvement 8, 369-395.

Goldstein, H. and Woodhouse, G. (2000), School effectiveness research and educational policy, Oxford Review of Education 26, 353-363.

Grosskopf, S., Hayes, K. and Taylor, L.L. (2009), The relative efficiency of charter schools, Annals of Public and Cooperative Economics 80, 67-87.

Johnes, J. (2004), Efficiency measurement, in G. Johnes and J. Johnes (eds), International Handbook on the Economics of Education, Edward Elgar, Cheltenham.

Johnes, J. (2006), Measuring efficiency: A comparison of multilevel modeling and data envelopment analysis in the context of higher education, Bulletin of Economic Research 58, 75-104.

Kirjavainen, T. and Loikkanen, H. (1998), Efficiency differences of Finnish senior secondary schools: An application of data envelopment analysis and Tobit analysis, Economics of Education Review 17, 377-394.

KREVI (2011), Folkeskolens faglige kvalitet: Analyse af skolernes undervisningseffekt, www.krevi.dk.

Lazear, E.P. (2001), Educational production, Quarterly Journal of Economics 116, 777-803.

Muniz, M., Paradi, J., Ruggiero, J. and Yang, Z. (2006), Evaluating DEA models used to control for non-discretionary inputs, Computers and Operations Research 33, 1173-1183.

Naper, L.R. (2010), Teacher hiring practices and educational efficiency, Economics of Education Review 29, 658-668.

OECD (2010), The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes, OECD, Paris.

Olesen, O.B. and Petersen, N.C. (2009), Target and technical efficiency in DEA: Controlling for environmental characteristics, Journal of Productivity Analysis 32, 27-40.

Payton, M.E., Greenstone, M.H. and Schenker, N. (2003), Overlapping confidence intervals or standard error intervals: What do they mean in terms of statistical significance?, Journal of Insect Science 34, 1-6.

Ray, S. (1991), Resource-use efficiency in public schools: A study of Connecticut data, Management Science 37, 1620-1628.

Ruggiero, J. (1996), On the measurement of technical efficiency in the public sector, European Journal of Operational Research 90, 553-565.

Ruggiero, J. (1998), Non-discretionary inputs in data envelopment analysis, European Journal of Operational Research 111, 461-469.

Ruggiero, J. (2004), Performance evaluation in education: Modeling educational production, in William W. Cooper et al. (red.), Handbook on Data Envelopment Analysis, Kluwer Academic Publishers, Hingham, MA.

Schenker, N. and Gentleman, J.F. (2001), On judging the significance of differences by examining the overlap between confidence intervals, The American Statistician 55, 182-186.

Waldo, S. (2006a), Competition and public school efficiency in Sweden – An empirical evaluation of second stage regression results for different models of nondiscretionary inputs in Data Envelopment Analysis (DEA), Department of Economics Working Papers 2006:7, Lund University.

Waldo, S. (2006b), School vouchers and public school productivity – The case of the Swedish large scale voucher program, Department of Economics Working Papers 2006:8, Lund University.

Waldo, S. (2007a), Efficiency in Swedish public education: Competition and voter monitoring, Education Economics 15, 231-251.

Waldo, S. (2007b), On the use of student data in efficiency analysis – Technical efficiency in Swedish upper secondary school, Economics of Education Review 26, 173-185.

Woodhouse, G. and Goldstein, H. (1988), Educational performance indicators and LEA league table, Oxford Review of Education 14, 301-320.

Worthington, A.C. (2001), An empirical survey of frontier efficiency measurement techniques in education, Education Economics 9, 245-268.

# Comment on Bogetoft and Wittrup: Productivity and education: Benchmarking of elementary and lower secondary schools in Denmark

Timo Kuosmanen*

There is a large and growing stream of literature on the productivity of educational institutions that overlaps such disciplines as educational economics, educational statistics, pedagogy, and operations research. The simplest thinkable productivity measure of a school is the pupil-teacher ratio, which is frequently referred to in public discussion. However, such simple ratios overlook such factors as the quality of educational outcomes, family background of pupils, economic resources of the school and its operating environment.

Bogetoft and Wittrup approach this topic from a novel perspective that combines *multi-level modeling* (MLM) from educational research (e.g., Raudenbush and Bryk, 2002) with *data envelopment analysis* (DEA, Charnes et al., 1978) used in applied economics and operations research. This allows the authors to take into account both the qualitative and quantitative factors suggested above. I find this to be a very useful contribution towards further integration of this scattered field of research that is too much divided by disciplinary boundaries. However, I also see the scope for a more systematic and rigorous integration of the methods. The purpose of this note is to provide some thoughts for further development

* School of Economics, Aalto University, Helsinki, Finland, timo.kuosmanen@aalto.fi.

of this line of research. I start from methodological issues and proceed to practical implementation via scale efficiency.

## 1. Further integration of methodology

MLM is mainly confined to educational research, but it is analogous to the standard *fixed effects* (FE) and *random effects* (RE) models in econometrics (e.g., Greene, 2012). While the FE and RE models are usually applied to panel data where $n$ schools are observed over $T$ time periods, they could equally well be applied to a cross-section of $n$ schools that are observed from the perspectives of $T_i$ individual pupils (assume an unbalanced panel where $T_i$ differs across schools). The fixed effects of the FE model can be used to estimate the "value added" of the school, analogous to the MLM approach. Interestingly, in the literature on *stochastic frontier analysis* (SFA, e.g., Kumbhakar and Lovell, 2003), the econometric panel data methods are similarly used for estimating productivity differences across schools. Such SFA models could readily incorporate both pupil-level and school-level information in a coherent and fully-integrated framework.

The main advantage of DEA to SFA is that it does not require any particular functional form for the production frontier. While DEA originated as a mathematical programming approach, the statistical properties of the DEA estimator have been established (Banker, 1993; Simar and Wilson, 2000). Recently, Kuosmanen and Johnson (2010) have shown that DEA is a sign-constrained variant of convex regression (Hanson and Pledger, 1976). This link between DEA and regression analysis enables us to integrate a DEA-style nonparametric frontier to the econometric models of panel data (see Kuosmanen and Kortelainen, 2011 for details). A practical challenge with such an integrated modeling approach is the enormous computational burden when the sample size (number of pupils) is very large. However, more efficient computational algorithms are being developed. Moreover, a clever use of modeling tools can significantly reduce the computational burden.

The main disadvantage of DEA is that it does not take into account an explicit disturbance term and hence, any measurement errors, omitted variables and other noise that is attributed to inefficiency. Incorporating a

DEA-style axiomatic frontier in an econometric framework as in Kuosmanen and Kortelainen (2011) can effectively solve this problem.

Finally, the two-stage method, where DEA efficiency scores are regressed on environmental variables $z$ (which Bogetoft and Wittrup consider at the end of Section 5) has been sharply criticized by Simar and Wilson (2007) who argue that the statistical inferences in the second-stage regression are invalid. Integrating the DEA-style nonparametric frontier to the econometric model also solves this problem (see Johnson and Kuosmanen, 2012).

## 2. Scale efficiency

An important issue from the policy perspective is the optimal school size, which is closely related to the optimal pupil-teacher ratio: small schools tend to have smaller classes than large schools. Small groups obviously require more teacher resources than large groups, but there seems to be a tradeoff between group size and the quality of educational outcomes. Further, if small schools are closed and pupils are reallocated to larger units, the transportation costs can increase significantly, especially in sparsely populated areas. Estimating the economically optimal school size is a challenging task involving several tradeoffs.

Bogetoft and Wittrup address this issue using the DEA-based scale efficiency measure. From a conceptual point of view, scale efficiency is perfectly appropriate for evaluating the most productive school size. However, the input variables considered in this study are restricted to labor inputs (teacher and staff hours). If the data so permits, the scale efficiency analysis could easily be extended to include the capital stock and other relevant inputs. However, to estimate the potential cost saving that can be obtained by scale efficient operation of schools from a social point of view, one should look beyond the school-specific input resources and take into account external costs (e.g., transportation cost) that are not included in the school budget.

## 3. Realizing the cost saving potential

Achieving the potential cost saving estimated by frontier methods is a challenge in practice. In this respect, Bogetoft and Wittrup devote considerable attention to interactive benchmarking where inefficient schools can learn from their efficient peers. Benchmarking is widely recognized as a useful strategy to facilitate the diffusion of best practices. However, benchmarking does involve an opportunity cost (consider the time devoted to searching the appropriate peers, visiting the schools and organizing meetings concerning school practices, etc.), whereas the benefits are uncertain. It is not always self-evident that the benefits outweigh the cost. To my knowledge, there is little evidence that DEA-based benchmarking is more cost-effective than some simpler and cheaper alternative for the diffusion of best practices. I would like to suggest this to be an interesting topic for further research.

The approach that I would propose is to first identify good practices that schools have (e.g., pedagogical approach, extracurricular activities, the use of ICT), then collect data regarding the practices, and finally apply an efficiency analysis to assess how different practices affect performance. The effects of school practices (or local or national policy) on productivity could be estimated similar to the $z$-variables that represent the operating environment (see Johnson and Kuosmanen (2012) for a more detailed discussion). The results of such estimation could provide valuable information for both educators and policy makers about the productivity improving practices.

## References

Banker, R.D. (1993), Maximum likelihood, consistency and data envelopment analysis: A statistical foundation, Management Science 39, 1265-1273.

Charnes A., Cooper, W.W. and Rhodes, E. (1978), Measuring the efficiency of decision making units, European Journal of Operational Research 2, 429-444.

Greene, W.H. (2012), Econometric Analysis, 7th edition, Prentice Hall, Boston.

Hanson, D.L. and Pledger, G. (1976). Consistency in concave regression, Annals of Statistics 4, 1038-1050.

Johnson, A.L. and Kuosmanen, T. (2012), One-stage estimation of the effects of operational conditions and practices on productive performance: Asymptotically normal and efficient, root-*n* consistent StoNEZD method, Journal of Productivity Analysis, forthcoming, DOI: 10.1007/s11123-011-0231-5.

Kumbhakar, S. and Lovell, C.A.K. (2003), Stochastic Frontier Analysis, Cambridge University Press, Cambridge.
Kuosmanen, T. and Johnson, A.L. (2010), Data envelopment analysis as nonparametric least squares regression, Operations Research 58, 149-160.
Kuosmanen, T. and Kortelainen, M. (2011), Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints, Journal of Productivity Analysis, forthcoming, DOI: 10.1007/s11123-010-0201-3.
Raudenbush, S.W. and Bryk, A.S. (2002), Hierarchical Linear Models: Applications and Data Analysis Method, Sage, Newbury Park, CA.
Simar, L. and Wilson, P.W. (2000), Statistical inference in nonparametric frontier models: The state of the art, Journal of Productivity Analysis 13, 49-78.
Simar, L. and Wilson, P.W. (2007), Estimation and inference in two-stage, semi-parametric models of production processes, Journal of Econometrics 136, 31-64.