INSTITUTE FOR STRUCTURAL RESEARCH

www.ibs.org.pl

- preliminary version -

Decomposing productivity growth in OECD countries: domestic R&D vs. international technology diffusion

Jakub Growiec, Łukasz Marć, Dorota Pelle

IBS WORKING PAPER #01/2009

Decomposing productivity growth in OECD countries: domestic R&D vs. international technology diffusion

Jakub Growiec* Łukasz Marć† Dorota Pelle‡

May 31, 2009

– Preliminary draft –

Abstract. This paper decomposes productivity growth across highly developed OECD countries in the period 1972-2000 into components attributable to domestic R&D output, a catch-up effect, and international technology diffusion via imports of hi-tech products.

Two alternative specifications of "productivity growth" are considered: (i) increments of the conventional residual measure of total factor productivity (computed under the assumption that the production function takes both physical and human capital as inputs), and (ii) the Malmquist productivity index, computed non-parametrically from a DEA-based growth accounting exercise.

The methodology consists in using panel data techniques to regress productivity growth against our constructed measures of domestic R&D output, technological catch-up and technology diffusion. Domestic R&D output is derived using the Schumpeterian specification taken from fully endogenous R&D-based growth models while the catch-up and technology diffusion terms are based upon accumulation of the human-capital augmented and weighted sum of "technology flows", proxied by appropriate measures of technological distance between the source and destination country in each pair.

Our results indicate that all of these components have been important for productivity growth in OECD countries in the considered period. The estimated relative magnitudes vary significantly across countries though. On average, productivity growth has been fueled mainly by technology diffusion and technological catch-up, whereas domestic innovations account on average for only 11% of productivity growth.

*Warsaw School of Economics, Institute of Econometrics, Poland, and National Bank of Poland. Address for correspondence: Instytut Ekonometrii SGH, Al. Niepodległości 162, 02-554 Warszawa, Poland. E-mail: jakub.growiec@gmail.com.

†Institute for Structural Research, Warsaw, Poland.

‡Institute for Structural Research, Warsaw, Poland, and Warsaw School of Economics, Poland. E-mail: pelle.dorota@gmail.com.

Keywords and Phrases: productivity growth, decomposition, R&D, technology diffusion, OECD

JEL Classification Numbers: O11, O33, O47

1 Introduction

Economic growth may be driven by a wide range of factors. Within the supply side approach, one may group them into three broad categories: (i) physical capital accumulation,(ii) human capital accumulation and population growth and (iii) total factor productivity (TFP) growth. Even though each of these three categories poses certain difficulties to adequate measurement, it is certainly the last, productivity category, encompassing all developments in available production technologies as well as changes in the efficiency of their usage in actual economic environments, which is the least straightforward to capture.

R&D-driven technological progress is the component of greatest importance to growth theorists because it offers the possibility of *sustained* technological progress. In reality, however, as opposed to closed-economy growth models, total factor productivity may grow both due to domestic R&D accomplishments and international technology diffusion wherein technological developments arrive from abroad. Only for technological leaders such as Japan or the US does it seem plausible that own R&D produces the majority of actually observed technological change; it is hardly possible for smaller and/or lagging economies, though. One can thus credibly hypothesize that technological diffusion must have played a major role in world's technological advancement throughout the history. Furthermore, technology diffusion has many faces: countries may advance due to their own technological catch-up by gradually learning to apply frontier technologies, or due to R&D-induced catch-up where this learning process is accelerated by domestic, "adaptive" or "imitative" R&D (cf. Griffith, Redding and Van Reenen, 2004). Finally, they may also advance due to the imports of high-tech products and services or foreign direct investment: both of these phenomena provide direct channels through which frontier knowledge could be passed to the converging economy.

The current article relates to four complementary strands of literature. First of all, it relates to the old tradition of growth accounting studies (cf. Solow, 1957; Barro, 1999; Caselli, 2005) where empirically observed growth in GDP per worker in the medium-to-long run is decomposed into factors attributable to physical capital accumulation, accumulation of other reproducible factors such as human capital, and the growth of TFP which is identified as a (Solow) residual value. The second related strand of literature applies deterministic non-parametric methods of productivity analysis to macroeconomic data (Kumar and Russell, 2002; Henderson and Russell, 2005; Jerzmanowski, 2007; Badunenko, Henderson and Zelenyuk, 2008; Growiec, 2008). This literature adds a new dimension to the aforementioned growth accounting procedures by removing the assumption that the production function is Cobb–Douglas, or has any firm, predefined functional form. From this literature, we draw the concept of the Malmquist productivity index which measures residual technological progress in a given country without convoluting it with production function misspecification. The third area of research which we extensively draw from is preoccupied with Bayesian Model Averaging (BMA) methods, useful for inferring about the most appropriate model specification when explanatory variables are unknown and possibly, strongly interrelated (cf. Fernandez, Ley, and Steel, 2001; Sala-i-Martin, Doppelhofer, and Miller, 2004). We shall, in particular, use these methods to derive the most appropriate lag structure for our considered variables. The fourth major strand of literature which we invoke deals with empirical measures of international technology transfer. Technology diffusion models, initialized by Nelson and Phelps (1966) and reviewed recently by Benhabib and Spiegel (2005), are juxtaposed here with standard approaches to modeling technological change in growth theory (cf. Jones, 1999) which assume that technological improvements arrive primarily due to R&D output. The articles most closely related to this one are therefore Ha and Howitt (2007) and, most importantly, Madsen (2008a), who test the empirical relevance of various alternative specifications of technological progress found across the theoretical literature.

Given this background, the contribution of the current article is threefold. First of all, we extend the growth accounting literature by providing a decomposition of TFP growth into the effects of domestic R&D, international technology diffusion, and two complementary measures of technological catch-up. We also provide a qualification to the results obtained hitherto by replacing TFP growth (computed using a Cobb–Douglas production function) with the Malmquist productivity index which measures technological progress without imposing a functional form on the production technology, and by comparing both cases. Second, we follow Madsen (2008a) in applying panel data techniques to estimate the relative impacts of R&D output and technology diffusion on total observed technological progress. We however complement his work by considering the Malmquist productivity index and measures of diffusion that are explicitly based on volumes of trade in *high-tech* goods and services. We also resolve the uncertainty with regard to the appropriate lag structure of the considered variables with the use of the BMA method. Third, we disentangle diffusion effects driven by international trade from effects stemming from a "pure" catch-up process, where each given country catches up with the current technology frontier as well as catch-up driven by adaptive/imitative R&D (Griffith, Redding, and Van Reenen, 2004). This decomposition could be obtained here only thanks to the use of a non-parametric approach to productivity analysis.

The remainder of the article is structured as follows. In Section 2, we lay out our empirical methodology. Of particular interest is the stepwise procedure which we used to construct the values of our measures of R&D output and technology transfer. In Section 3, we discuss our data sources as well as provide a first cursory look at the properties of our constructed variables. In Section 4, we present our main results, i.e., a decomposition of total factor productivity growth into components attributable to domestic R&D, catch-up, and import-driven international technology diffusion. In a separate subsection we contrast these results, based on the Malmquist productivity index, with ones attainable under the standard Cobb–Douglas function paradigm. Section 5 concludes.

2 Empirical methodology

Our empirical investigation is carried out in the following stages:

1. **Construction of the dependent variable: productivity growth.** This is done in two ways – first, as the Malmquist index computed from the non-parametric Data Development Analysis method (cf. Kumar and Russell, 2002; Henderson and Russell, 2005; Jerzmanowski, 2007; Badunenko, Henderson, and Zelenyuk, 2008), and second, as total factor productivity growth from the standard, Cobb–Douglas function-based Solow decomposition (Solow, 1957), for each country and year. It is calculated both upon the original data on annual GDP per worker (including the business cycle component) and upon HP-filtered data which capture medium- and long-term trends only (cf. Hodrick and Prescott, 1997). Should the countries across the sample differ significantly in terms of their business cycles, solely HP filtered data will be employed in the analysis.

2. **Construction of explanatory variables: R&D output, catch-up and technology diffusion.** Since neither of them is directly measurable, R&D output and the volume of technology transfer are constructed according to theoretical formulas obtained from state-of-the-art growth models (cf. Madsen, 2008a,b; Benhabib and Spiegel, 2005). We also take advantage of our non-parametric estimates of the world technology frontier to compute auxilliary non-standard measures of technological catch-up.

3. **Characterization of the dataset.** Having constructed the variables, we investigate their basic properties such as common trends, relative variability, and cross-correlations. This should provide an initial intuition on the relative importance of R&D, technology diffusion, and catch-up for productivity growth.

4. **Lag structure selection for our explanatory variables.** Since the impact of considered explanatory variables on productivity growth might be delayed, it is particularly important to take this lag into account in the empirical analysis. Due to high collinearity, the lag structure should not be determined within a panel econometric model, though. To resolve this problem we shall thus employ the Bayesian Model Averaging (BMA, cf. e.g. Sala-i-Martin, Doppelhofer, and Miller, 2004) framework to cross-sectional data to establish the lag structure and inclusion's probability of all explanatory variables. Subsequently, the results are applied to panel econometric analysis.

5. **Econometric analysis using panel data.** At this stage, we carry out an econometric procedure aimed at disentangling the respective impacts of R&D and technology diffusion on productivity growth in a 32-year long panel of 19 OECD countries.

6. **Robustness checks,** aimed at testing the sensitivity of our results to changes in the methodology such as substituting the non-parametric measure of productivity growth with the standard measure of TFP growth, changing lag lengths, and altering the formulas for R&D output and technology diffusion.

2.1 The parametric vs. non-parametric approach

As already mentioned above, productivity growth has been calculated in two ways. The main contribution of this paper is, however, to analyze the determinants of productivity growth computed nonparametrically, i.e. the Malmquist productivity index. This requires the use of Data Envelopment Analysis (DEA).

2.1.1 Data Envelopment Analysis

The idea behind the DEA method is to construct the best-practice production function, nonparametrically, as a convex hull of production techniques (input–output configurations) currently used in countries present in the data.

The production function is then inferred indirectly as a fragment of the boundary of this convex hull for which is output is maximized given inputs. More precisely, for each observation i, output y_i is decomposed as:

$$y_i = E_i f(\mathbf{x}_i) \tag{1}$$

i.e., into a product of the maximum attainable output given inputs $y_i^* \equiv f(\mathbf{x}_i)$ and the efficiency index $E_i \in (0,1]$. In other words, the efficiency index E_i measures (vertical) distance to the technology frontier, while the frontier itself is computed nonparametrically as $y_i^* = f(\mathbf{x}_i)$. It should be noted that the vector of inputs, $\mathbf{x}_i$, could in principle be of any length $n \in \mathbb{N}$, but if one distinguishes too many types of inputs then (i) the DEA could run into numerical problems due to the "curse of dimensionality" (cf. Färe et al., 1994), and (ii) the efficiency levels could be overestimated due to too small a sample size.

Formally, the (output-based) DEA method is a linear programming technique allowing one find the efficiency index E_j for each unit $j = 1, 2, ..., I$ in the sample such that its reciprocal is maximized given a series of feasibility constraints (cf. Fried, Knox

Lovell and Schmidt, 1993):

$$\begin{aligned}
& \max_{\{\theta_j,\lambda_1,\ldots,\lambda_I\}} \theta_j \\
\text{s.t.} \quad & \theta_j y_j \leq \sum_{i=1}^{I} \lambda_i y_i, \\
& \sum_{i=1}^{I} \lambda_i x_{1i} \leq x_{1j}, \\
& \sum_{i=1}^{I} \lambda_i x_{2i} \leq x_{2j}, \\
& \vdots \\
& \sum_{i=1}^{I} \lambda_i x_{ni} \leq x_{nj}, \\
& \lambda_i \geq 0, \qquad i = 1, 2, \ldots, I.
\end{aligned} \tag{2}$$

The (output-oriented Debreu–Farell) efficiency index E_j is computed as the reciprocal of θ_i (that is, $E_i = 1/\theta_i$).

Since the data contain a finite number of data points, one for each country and each year, by construction the DEA–based production function will be piecewise linear and its vertices will be the actually observed *efficient* input–output configurations (i.e. non-dominated by any linear combination of other observed input–output configurations).

2.1.2 Advantages and limitations of the approach

The DEA is a deterministic, data-driven approach to deriving the production function from observed input–output pairs. Its unquestionable strength lies in the fact that it does not require any particular functional form of the aggregate production function (provided that it has constant returns to scale and satisfies the free-disposal property), and provides testable predictions on its shape instead. Indeed, the usual assumption of a Cobb–Douglas production function may lead to marked biases within growth accounting or levels accounting exercises leading to an overestimation of the role of total factor productivity (TFP), as argued by Caselli (2005) and Jerzmanowski (2007). As for the predicted shape of the production function, one obvious fact is that due to the characteristics of the method, it will be piecewise linear for any finite data sample. With reasonably large data samples, however, certain parametric forms could be tested formally against the DEA-based nonparametric benchmark, such as the CES or the Cobb–Douglas (presumably leading to a rejection of the latter).

There are important limitations to the DEA approach as well. First, its deterministic character makes it silent on the estimation precision of the aggregate production function and of the predicted efficiency levels if inputs and outputs are subject to stochastic shocks.

Second, the DEA provides a biased proxy of the actual technological frontier. Certainly, even the most efficient units in the sample could possibly operate with some extra efficiency: they are themselves aggregates of smaller economic units and must therefore have some internal heterogeneity. Taking account of that, the frontier could easily be shifted upwards; efficiency in nevertheless normalized to 100% for the most efficient units in the sample.

Third, the DEA constructs the production function basing on the efficient data points. This makes it naturally sensitive to outliers and measurement error. On the one hand, outliers characterized by obvious errors are easily spotted because they spoil the whole subsequent analysis. Systematic mismeasurement associated with some units could be left unnoticed, however, if these units fall short of the frontier. The data have been carefully checked, though, so that one can be confident that the risk of errors has been minimized.

2.1.3 The Malmquist productivity index

To compute the Malmquist productivity index from data on the shape of the world technology frontier, we shall carry out a nonparametric growth accounting exercise (cf. Growiec, 2008). With three factors of production, as we have in our analysis, K, L^U, L^S (to be described later), the "Fisher-ideal" (cf. Henderson and Russell, 2005) decomposition of the year-on-year productivity ratio for each country i (index supressed for brevity) is the following (denoting the current year as t):

$$
\begin{aligned}
\frac{y_t(K_t, L_t^U, L_t^S)}{y_{t-1}(K_{t-1}, L_{t-1}^U, L_{t-1}^S)} &= \frac{E_t}{E_{t-1}} \cdot \frac{y_t^*(K_t, L_t^U, L_t^S)}{y_{t-1}^*(K_{t-1}, L_{t-1}^U, L_{t-1}^S)} = \\
= \underbrace{\frac{E_t}{E_{t-1}}}_{\text{efficiency}} &\cdot \underbrace{\sqrt{\frac{y_t^*(K_t, L_t^U, L_t^S)}{y_{t-1}^*(K_t, L_t^U, L_t^S)} \frac{y_t^*(K_{t-1}, L_{t-1}^U, L_{t-1}^S)}{y_{t-1}^*(K_{t-1}, L_{t-1}^U, L_{t-1}^S)}}}_{\text{techn. progress}} \cdot \\
&\cdot \underbrace{\sqrt{\frac{y_t^*(K_t, L_t^U, L_t^S)}{y_t^*(K_{t-1}, L_{t-1}^U, L_{t-1}^S)} \frac{y_{t-1}^*(K_t, L_t^U, L_t^S)}{y_{t-1}^*(K_{t-1}, L_{t-1}^U, L_{t-1}^S)}}}_{\text{factor accumulation}},
\end{aligned} \tag{3}
$$

where $y_t^* \equiv f(\mathbf{x}_t)$ denotes the maximum output attainable at time t given inputs contained in the vector $\mathbf{x}_t$.

The decomposition of GDP growth defined by Eq. (3) singles out the dynamic changes in efficiency, shifts in the technology frontier given factor endowments, and factor accumulation holding the technological frontier fixed. The product of the "efficiency change" and "technological progress" factors is the (output-oriented) Malmquist productivity index (cf. Fried, Knox Lovell, and Schmidt, 1993) which we denote as M_{it}. It measures, for each country i and year t, the total change in productivity which resulted from anything but factor accumulation. In other words, the Malmquist productivity index captures the total productivity improvement under technologies *actually* *used* in the given country, whereas our "technological progress" index measures the total productivity improvement under *frontier* technology, given the country's factor endowments.

2.1.4 TFP growth

Compared with the nonparametric approach to measuring productivity growth, the standard Solow decomposition is very straightforward. It consists in computing the growth rate of TFP, A_{it}, computed as a residual from the Cobb–Douglas production function $Y_{it} = A_{it}K_{it}^{\alpha}(L_{it}^{S} + L_{it}^{U})^{1-\alpha}$. The usual choice of a Cobb–Douglas production function has two major limitations, though: (i) as opposed to the DEA approach, it does not allow one to distinguish between technological progress and efficiency growth, and (ii) the explicit assumption of a unitary elasticity of substitution between the factors of production is disputable empirically (cf. e.g. Duffy and Papageorgiou, 2000). What is also important here, we only included a single aggregate measure of human capital in our analysis, that is $H = L^S + L^U$. This implies a unitary elasticity of substitution between skilled and unskilled labor, an assumption which was not taken in the case of our nonparametric estimate of productivity growth.

2.2 Construction of the R&D output variable

The capacity of the economy to innovate might be evaluated on the basis of the number of patents submitted per R&D worker, the share of R&D employment, and the current "stock of technology". There are however a few qualifications to this argument, as far as empirical analyses are concerned, which ought to be explained. First, the "stock of technology" is not directly measurable, and proxying it with the residual TFP measure requires the assumption that the production function be Cobb–Douglas. If the elasticity of R&D output with respect to TFP is unitary, however, this problem might be alleviated by specifying all variables in "per technology unit" terms, i.e., divided by A_{it} (cf. Madsen, 2008a). Second, in panel data, hardly ever do the results of R&D efforts bring out immediate results. A certain number of lags in variables must therefore be included in the analysis. Third, there is fundamental uncertainty to the elasticity parameters in the specification,in particular with respect to the parameter θ used in the specification (4) below.

In our study, we shall therefore use the following empirical specification of R&D output:

$$R\&D_{it} = \sum_{q=0}^{Q} \omega_{i,q}^{RD} \left(\frac{Pat_{i,t-q}}{L_{i,t-q}^{A}} \right) (\ell_{i,t-q}^{A})^{\theta}, \tag{4}$$

where Q is the maximum lag length, Pat_{it} is the number of patents applied for in country i at time t, L_{it}^{A} is the total employment in the R&D sector, ℓ_{it}^{A} is the share of R&D in overall employment, and $\beta_{i,q}^{RD}$ is a weighting factor.

Intuitively, one should expect $R\&D_{it}$ to be stationary: in the long run, the share of R&D employment should stabilize around a fixed value, and so should do the per-researcher patent count. In the available data, however, we observe a general upward trend in the R&D employment share and a general downward trend in patenting frequency. These two trends offset each other so in the data, the R&D output measure is stationary for a wide range of values of θ.

The elasticity of R&D output with respect to the share of population employed in R&D, measured by θ, is a hidden parameter which cannot be directly measured. We are not aware of any systematic empirical investigation which would try to infer the value of θ from indirect evidence. For this reason, it would be of great importance to assess the sensitivity of our results with respect to changes in this parameter. As a base value for this parameter, we take $\theta = 0.5$.

2.3 Measures of technology diffusion and catch-up

In the literature, technology diffusion and the catch-up process have been captured empirically in a number of ways. On the one hand, the term may reflect a pure convergence process, allowing countries to cover their distance the technological frontier and thus catch up with the currently best performing countries (controlling for the difference in factor endowments). On the other hand, however, it may also capture a multi-directional import-driven diffusion process where technology improvements come together with imports of high-technology products, and may well diffuse from less productive to more productive countries as well.

We are going to distinguish between three different specifications of technology transfer: (i) the pure catch-up effect, (ii) the R&D-induced catch-up effect (i.e. "the second face of R&D", cf. Griffith, Redding and Van Reenen, 2004), and (iii) technology diffusion fueled by imports of hi-tech products (cf. Madsen, 2007, 2008a).

The pure catch-up effect can be specified as

$$PCU_{it} = \sum_{r=0}^{R} \omega_{i,r}^{PCU} \frac{E_{i,t-r}}{E_{i,t-r-1}} \Delta \frac{L_{i,t-r}^{S}}{H_{i,t-r}} \tag{5}$$

where E_{it} is the efficiency index of country i at time t. The first term therefore captures the distance of the country to the frontier *given production inputs.* We have however substituted the measure of *remaining* distance to the frontier with a measure of distance *actually covered* between $t-1$ and t. In theory, they ought to be proportional, but this regularity need not always hold in empirical data. The constant R represents the maximum lag in technological catch-up. The last term, $\Delta \frac{L_{i,t-r}^{S}}{H_{i,t-r}}$,[1] is the annual change in the ratio of total human capital within the skilled population to total human capital. It is included to capture the change in ease of technology adoption which is the higher, the more skilled the population becomes.

[1]Here and throughout the paper, we use the notation $\Delta X_t = X_t/X_{t-1}$ for any variable X.

The R&D-induced catch-up effect, capturing the "second face of R&D", is computed by interacting country's R&D with its distance to frontier given factor inputs:

$$R\&DCU_{it} = \sum_{s=0}^{S} \omega_{i,s}^{RDCU} \frac{E_{i,t-s}}{E_{i,t-s-1}} \Delta \frac{L_{i,t-s}^{S}}{H_{i,t-s}} \left(\frac{Pat_{i,t-s}}{L_{i,t-s}^{A}} \right) (\ell_{i,t-s}^{A})^{\theta}. \tag{6}$$

Technology diffusion fueled by imports of hi-tech products (cf. Madsen, 2008a) is, in turn, measured as:

$$DIFF_{it} = \sum_{p=0}^{P} \omega_{i,p}^{DIFF} \sum_{j \neq i} \frac{X_{ij,t-p}}{X_{i,t-p}} \cdot \frac{Y_{j,t-p}}{Y_{i,t-p}} \tag{7}$$

where $X_{ij,t}$ is the volume of hi-tech trade from country j to country i, X_{it} is the total volume of country i's imports, and $\frac{Y_{j,t-p}}{Y_{i,t-p}}$ represents GDP per worker in country j relative to i. The last factor is included to provide a proxy for product variety, so that technological diffusion between each particular pair of countries is rougly proportional to the "technological distance" between them, but allowing for positive diffusion from the richer to the poorer country as well.

2.4 Lag structure

Our approach to establishing the lag structure and solving the multicollinearity problem is to employ the BMA method to a cross-sectional database that consists of all explanatory variables in each period and their four subsequent lags. The explained variable is going to be either the Malmquist productivity index or TFP growth, in line with the dependent variable used in the subsequent panel data analysis.

Results of the BMA exercise should help us determine the weights of all lagged variables in constructing relevant aggregate variables for the four technological progress categories described above ($R\&D$, PCU, $R\&DCU$ and $DIFF$). Only these aggregates will be then subsequently used in our panel data study.

The BMA is a standard Bayesian solution to model uncertainty, where the prediction and inference are based on a weighted average over all possible models under consideration, rather than on one single regression model. One is required to assume a prior probability of each model and a prior probability distribution over the parameters of each model. Estimated parameters of the models and prior probabilities are then used to derive weights of all underlying regression models in the posterior model choice probability distribution. Thus BMA approach directly addresses a question that is central to our analysis: "how likely is it that a given regressor (i.e., given lag of a certain explanatory variable) has a significant effect on the dependent variable?"

In our reserach we employ the priors suggested by Sala-i-Martin, Doppelhofer, and Miller (2004) as well as Fernandez, Ley, and Steel (2001), properties of which have been widely discussed in the literature. Hence, in our BMA application we assume that a

Table 1: Weights of lags chosen on the basis of BMA.

Dependent variable: Malmquist productivity index					
Coefficients	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4
ω^{RD}	0.775	0.674	0.662	0.487	0.291
ω^{PCU}	0.837	0.738	0.517	0.168	0.000
ω^{RDCU}	0.725	0.641	0.473	0.139	0.000
ω^{DIFF}	0.618	0.584	0.496	0.391	0.277

Source: own computations.

prior variable's inclusion probability is random. Results of the BMA procedure taking the Malmquist productivity index as our dependent variable are contained in Table 1.

Results of the BMA exercise are very much in line with economic intuition. First of all, the impact of each particular variable decreases with the lag. In all considered cases, the largest observed impact is the instantaneous one (lag 0), but it is nearly as sizeable after one and two years as well. Secondly, we observe that the impacts of domestic R&D and international technology diffusion on productivity growth are much more extended in time than the impacts of catch-up effects. The coefficients on catch-up terms lagged by four years are already zero while for the other regressors, they are still larger than one third of their respective instantaneous impact.

2.5 Econometric methods

Our econometric approach is an application of panel data methods with country-specific fixed effects to estimate a number of models upon our datasets. Time-invariant country-specific fixed effects β_i^{SCE} have been introduced into the basic model to control for the heterogeneity of countries in the sample. These effects should be interpreted as encompassing the differences in the initial stock of human capital accumulated in the economies across the sample and/or the initial level of R&D output.

The general model which we estimate takes the form:

$$M_{it} = \beta_i^{SCE} + \beta^{RD} R\&D_{it} + \beta^{PCU} PCU_{it} + \beta^{RDCU} R\&DCU_{it} + \beta^{DIFF} DIFF_{it} + \varepsilon_{it}, \quad (8)$$

or in extensive notation:

$$\begin{aligned} M_{it} &= \beta_i^{SCE} + \quad \beta^{RD} \sum_{q=0}^{4} \omega_q^{RD} \left(\frac{Pat_{i,t-q}}{L_{i,t-q}^{A}} \right) (\ell_{i,t-q}^{A})^{\theta} + \\ &+ \qquad \beta^{PCU} \sum_{r=0}^{3} \omega_r^{PCU} \frac{E_{i,t-r}}{E_{i,t-r-1}} \Delta \frac{L_{i,t-r}^{S}}{H_{i,t-r}} + \\ &+ \qquad \beta^{RDCU} \sum_{s=0}^{3} \omega_s^{RDCU} \frac{E_{i,t-r}}{E_{i,t-r-1}} \Delta \frac{L_{i,t-s}^{S}}{H_{i,t-s}} \left(\frac{Pat_{i,t-s}}{L_{i,t-s}^{A}} \right) (\ell_{i,t-s}^{A})^{\theta} + \\ &+ \qquad \beta^{DIFF} \sum_{p=0}^{4} \omega_p^{DIFF} \sum_{j \neq i} \frac{X_{ij,t-p}}{X_{i,t-p}} \cdot \frac{Y_{j,t-p}}{Y_{i,t-p}} + \varepsilon_{it}, \end{aligned} \tag{9}$$

where the residual term ε_{it} is assumed to be spherical. The lag structure within each of the four technological progress measures is pre-determined at this stage as it has already been approximated within the BMA procedure.

2.6 Potential methodological problems

Three difficulties might arise while dealing with the aforementioned model.

1. Since highly aggregated data with high persistence will be employed in the analysis, the problems of **autocorrelation and heteroscedasticity** are likely to occur. To resolve this problem, a series of panel Wooldridge's tests likelihood-ratio tests will be carried out, so that the estimation technique could be tailored to their outcomes. Hence, the model will be estimated with the General Least Squares (GLS) method with a correction for autocorrelation and heteroscedasticity, should these be applicable.

2. At least some variables might by represented by **non-stationary processes**. If this is the case, the results of the estimation will be biased. Therefore a second-generation panel unit root test of Im, Pesaran and Shin (2003) will be applied and non-stationary variables will be transformed so that they do not endanger the analysis with spurious results anymore.

Eventually, the GLS estimator employed upon $I(0)$ panel data takes the following form:

$$\hat{\beta} = (\sum_{i=1}^{N} \tilde{X}' \hat{\Omega}^{-1} \tilde{X})^{-1} \sum_{i=1}^{N} \tilde{X}' \hat{\Omega}^{-1} \tilde{Y} \tag{10}$$

where $\hat{\Omega}$ is given by the orthogonal-deviation within-group residual intertemporal covariance matrix:

$$\hat{\Omega} = \frac{1}{N} \sum_{i=1}^{N} \hat{\tilde{v}}_i \hat{\tilde{v}}_i' \tag{11}$$

and $\tilde{v_i}$ is a $T-1\times 1$ vector of the form $\tilde{v_{i,t}} = \sqrt{\frac{T-t}{T-t+1}}(v_{i,t} - \frac{1}{T-t}(v_{i,t+1} + ... + v_{i,T}))$. The matrix $\tilde{X}$ and the vector $\tilde{Y}$ describe, respectively, explanatory variables and explained variables transformed with the orthogonal deviations method.

3 A look at the data

3.1 Data sources for the construction of our variables

International data on GDP and GDP per worker have been taken from the Penn World Table 6.2 (cf. Heston, Summers and Aten, 2006), available for 1960-2003. The unit of measurement is the PPP converted US dollar under constant prices as of year 2000.

The physical capital series have been calculated using the perpetual inventory method (cf. Caselli, 2005). We have taken country-level investment shares as well as government shares from the Penn World Tables 6.2. There are two polar standpoints as for the role of government in capital accumulation: one is that government spending is all consumption, and the other one is that it is all investment. We have taken an intermediate stance here, assuming that the government invests the same percentage of its GDP share as the private economy does. Under this assumption, the overall (private and public) investment share is $s/(1-g)$ where s is the private investment share and g is the government share. Furthermore, following Caselli (2005), we assumed an annual depreciation rate of 6%.

Country-level human capital data have been taken from de la Fuente and Doménech (2006) – D-D hereafter. The raw variables are shares of population aged 25 or above having completed primary, some secondary, secondary, some tertiary, tertiary, or post-graduate education. The considered dataset is of 5-year frequency only and it ends in 1995. Among all possible education attainment databases, the D-D dataset has been given priority due to our trust in its superior quality. The original D-D series has been extrapolated forward until the year 2000 using Cohen and Soto (2007) schooling data as a predictor for the trends. Neither Barro and Lee (2001) nor Cohen and Soto (2007) data could be used directly for this purpose because neither of them is (even roughly) in agreement with the D-D dataset – nor with each other – in the period where all datasets offer data points. Furthermore, the human capital data have been extrapolated to all intermediate years as well. This was possibly due to the fact that human capital variables are not susceptible to business cycle variations.

Data on the number of patents applied for, used as a measure of domestic R&D intensity, have been provided to us by Madsen (2007, 2008a,b). The same applies to the dataset on the number of R&D workers in each country and year within the 1970–2000 range. These datasets seem to be much more reliable than any other data available.[2]

For data on bilateral trade of hi-tech goods, we have used the freely available OECD dataset (International Trade by Commodity Statistics), available for 1988-2007. These

[2]We are grateful to Professor Madsen for providing these data to us.

series have been extrapolated backwards using the data on total imports. The list of commodities classified as hi-tech products is provided in the Appendix.

3.2 Stationarity concerns

To test stationarity, the t-test proposed by Im, Pesaran and Shin (2003) for unit roots in heterogenous panels with cross-section dependence was used. The test is based on the mean of individual denisity function t-statistics of each unit in the panel, with the null hypothesis assuming that all series are non-stationary. The null hypotheses were rejected in case of all variables used in further analysis: $R\&D, PCU, R\&DCU, DIFF$ (for all these variables their respective p-values were equal to 0.00), the Malmquist productivity index M (0.00) and TFP growth ΔA (0.00). Therefore, all variables used in the analysis are stationary.

4 Main results

The Data Envelopment Analysis, Bayesian Model Averaging, and econometric exercise described above allow us to evaluate the relative significance of four basic components of productivity growth across OECD countries: research and development (R&D) output, pure and R&D-augmented catch-up effects, as well as technology diffusion fueled by high-tech imports. Our results indicate that these components have contributed to the productivity growth in OECD countries within the considered time slot, but the estimated relative magnitudes vary significantly across countries.

4.1 Sources of the productivity growth under DEA decomposition

The final specification of the model (9) is satisfactory from the statistical perspective and takes the form of:

$$\begin{aligned} M_{it} &= \beta_i^{SCE} + \beta^{RD} \sum_{q=0}^{4} \omega_q^{RD} \left(\frac{Pat_{i,t-q}}{L^A_{i,t-q}} \right) (\ell^A_{i,t-q})^{\theta} + \\ &+ \beta^{PCU} \sum_{r=0}^{3} \omega_r^{PCU} \frac{E_{i,t-r}}{E_{i,t-r-1}} \Delta \frac{L^S_{i,t-r}}{H_{i,t-r}} + \\ &+ \beta^{RDCU} \sum_{s=0}^{3} \omega_s^{RDCU} \frac{E_{i,t-r}}{E_{i,t-r-1}} \Delta \frac{L^S_{i,t-s}}{H_{i,t-s}} \left(\frac{Pat_{i,t-s}}{L^A_{i,t-s}} \right) (\ell^A_{i,t-s})^{\theta} + \\ &+ \beta^{DIFF} \sum_{p=0}^{4} \omega_p^{DIFF} \sum_{j \neq i} \frac{X_{ij,t-p}}{X_{i,t-p}} \cdot \frac{Y_{j,t-p}}{Y_{i,t-p}} + \varepsilon_{it}, \end{aligned} \tag{12}$$

Table 2: Estimation results.

Method: Generalized Least Squares Characteristics: heteroskedastic with cross-sectional correlation Dependent variable: Malmquist productivity index				
	Coefficients	Standard Error	t-stat	$Prob > \|t\|$
β^{RD}	2.923	0.621	4.71	0.000
β^{PCU}	0.396	0.020	19.74	0.000
β^{RDCU}	0.385	0.006	6.07	0.000
β^{DIFF}	0.020	0.005	3.95	0.000
$const$	0.983	0.005	207.30	0.000
Wald test $\chi^2(4) = 482.50$ Log likelihood = 1688.702 Prob> $\chi^2 = 0.000$				

Source: own computations.

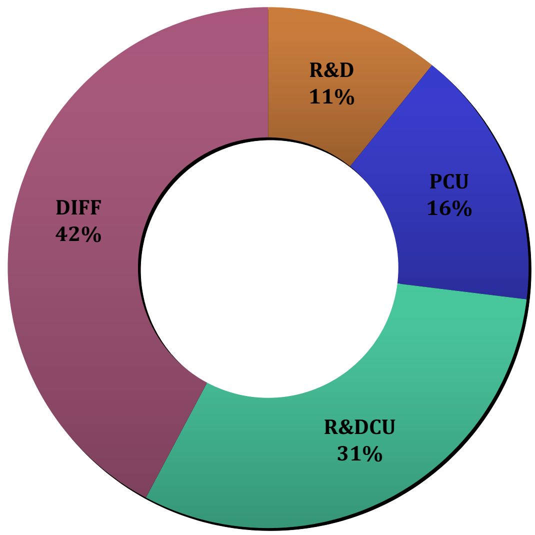
where the estimates of respective β parameters, after correction for autocorrelation and heteroskedasticity, take the values presented in Table 2.

Clearly, all variables have a positive and significant impact on the productivity growth of highly developed OECD countries within the 1972–2000 time range. The imposed lag structure implies that domestic innovations are, on average, more time-consuming than technological catching-up, no matter if supported by domestic R&D effort or not.

Thus, international technology diffusion might lead to a decrease in a country's distance to the world technology frontier. However, the stimulating impact of foreign technology absorption on the productivity gap is likely to shrink over time because the frontier keeps getting away. Furthermore, the international technology diffusion channel is further strengthened by countries' direct catch-up to the frontier, partially supported by their domestic adaptive/imitative R&D.

Set aside estimation error and country fixed effects, the results obtained above can be presented in the appealing form of a decomposition exercise for averaged data. It turns out (Figure 1) that on average, domestic R&D output explains around 11% of total OECD-wide productivity growth, pure catching-up effects capture around 16%, R&D-supported catching-up effects capture around 31%, and technology diffusion explains around 42% of productivity growth. It should be emphasized though that R&D effort has had a major contribution to the catching-up of lagging countries (cf. Griffith, Redding and Van Reenen, 2004), as shown on Figure 2. Clearly, diffusion and catch-up have thus been the major forces behind productivity growth, even though our sample contains only highly developed OECD countries, for which own R&D could be the most important. The contribution of domestic R&D to productivity growth of each particular country is relatively small, but it actually *is* quite sizeable even in arguably small countries (in comparison to the US) such as Switzerland, Sweden, Finland, and Austria. This result contrasts with earlier results of Comin (2004).

Figure 1: Decomposition of DEA-based productivity growth averaged across OECD countries



Source: Own calculations.

Given the large differences across OECD countries in terms of R&D effort made in the period of 1972-2000 (cf. OECD, 2003; Madsen, 2008a,b) it comes as no surprise that the contribution of each productivity growth component varies significantly across countries in the sample, as shown in Figure 2.

Japan, Italy, Switzerland, Sweden, Australia, and the United States could be found among the most efficiently innovating countries, in which domestic R&D output explained more than 14% of productivity growth. Domestic R&D was least productivity-improving in Portugal, Belgium, Canada, the Netherlands, Norway, and Spain with less than 6%.

Pure catching-up effects, on the other hand, have provided their strongest impacts on productivity growth in Portugal, the Netherlands, and Norway, whereas least impact has been observed in Switzerland, Canada, and the United States. This mirrors the countries' distance to the DEA-based technology frontier (cf. Growiec, 2008) – the smaller the actual distance to cover, the smaller is the potential for catching-up effects.

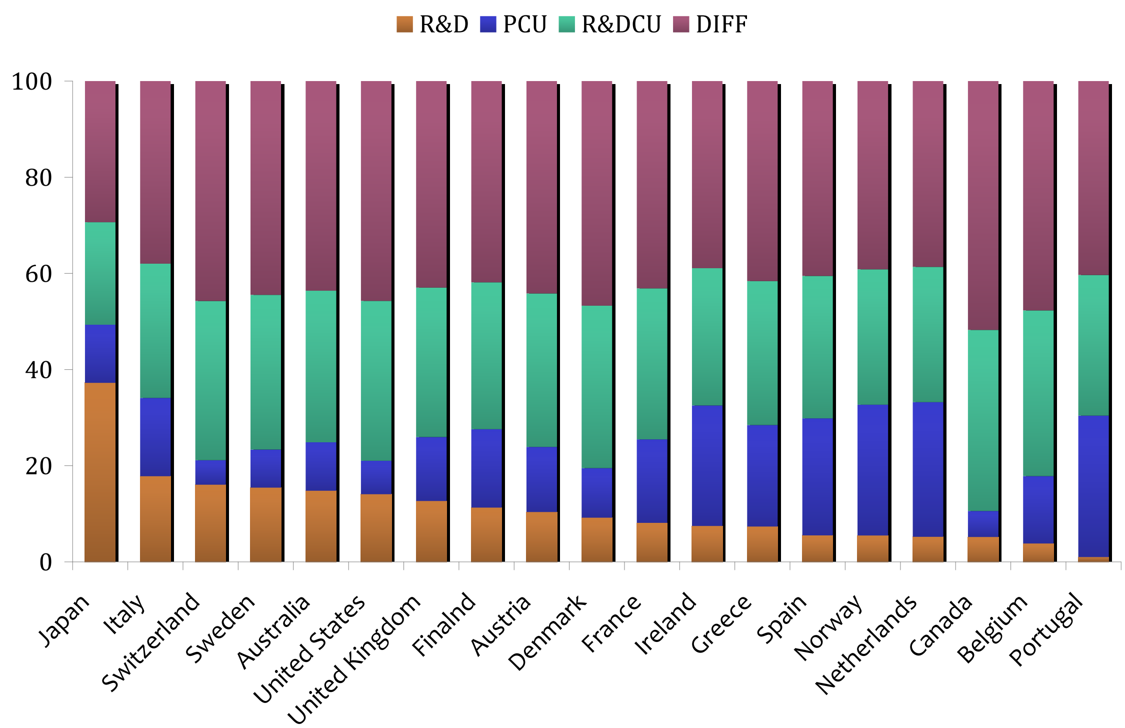
R&D-induced catching-up effects were most effective in Canada and Belgium – i.e., two relatively small countries with relatively high in-house R&D outlays. Japan and

Italy, on the other hand, have benefited least from this form of technological progress: their R&D output has been mostly classified as direct, i.e. "frontier" R&D, and not indirect "adaptive" or "imitative" R&D.

Technology diffusion has been a major force behind productivity growth in all countries. Analogously to R&D-supported catching-up, its effects have been felt most strongly in Canada and Belgium, and least strongly in Japan and Italy.

It was however Austria, Canada, Ireland and Norway which witnessed the highest productivity growth rates in the sample, mainly due to the R&D enhanced technological catch-up and technology diffusion.

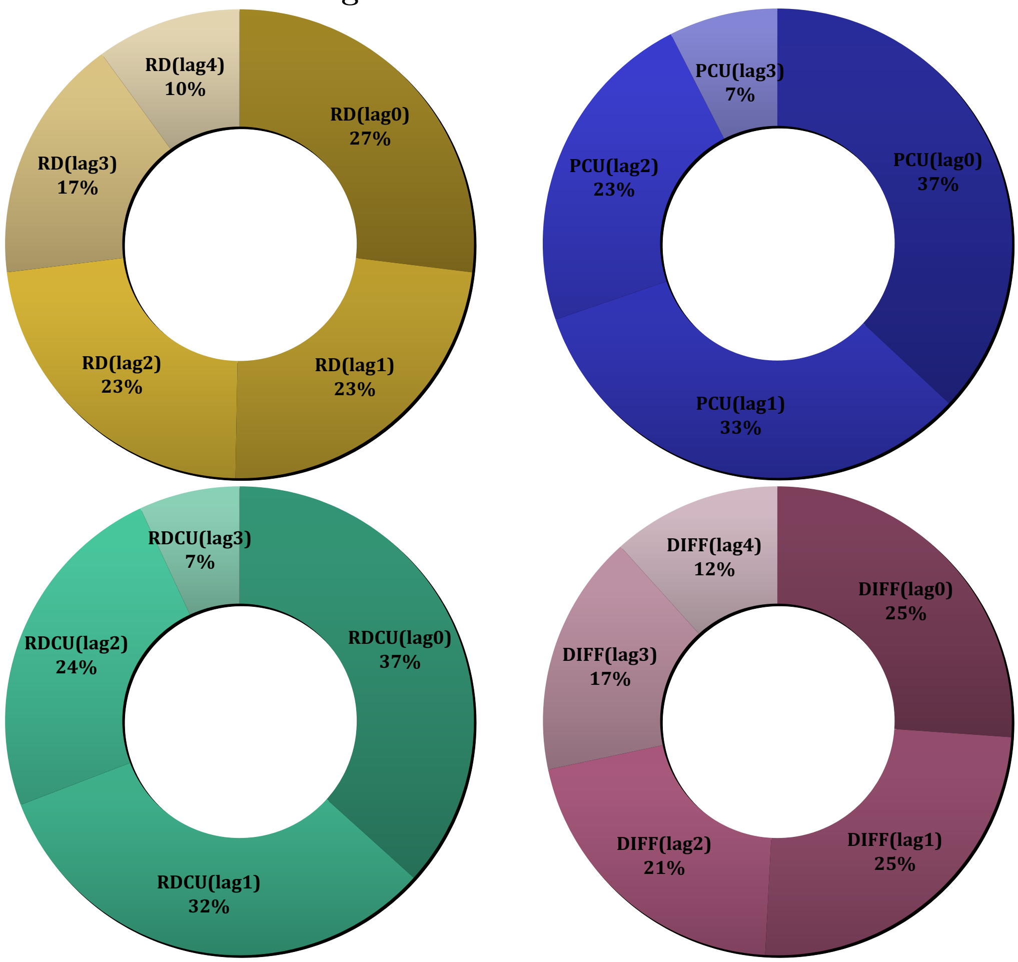
Figure 2: DEA-based productivity growth breakdown across OECD countries



Source: Own calculations.

It is also interesting to analyze the lags after which those total effects, described in above paragraphs, take place. A decomposition of these total effects between all consecutive lags has been presented in Figure 3.

Figure 3: Decomposition of DEA-based productivity growth averaged across OECD countries: lag structure



Source: Own calculations.

4.2 Malmquist productivity index vs. TFP growth

Let us now compare our benchmark results, where we decomposed the Malmquist productivity index, i.e., a DEA-based productivity growth measure, into components attributable to $R\&D$, PCU, $R\&DCU$ and $DIFF$, to an exercise based on the standard measure of TFP growth.

The general model which we shall estimate in the current section takes the form:

$$\Delta A_{it} = \beta_i^{SCE} + \beta^{RD} R\&D_{it} + \beta^{PCU} PCU_{it} + \beta^{RDCU} R\&DCU_{it} + \beta^{DIFF} DIFF_{it} + \varepsilon_{it}, \quad (13)$$

which differs from the model 8 only in the assumption about the dependent variable. The estimation results are contained in Table 3. Again, all parameters have expected signs and are strongly significant.

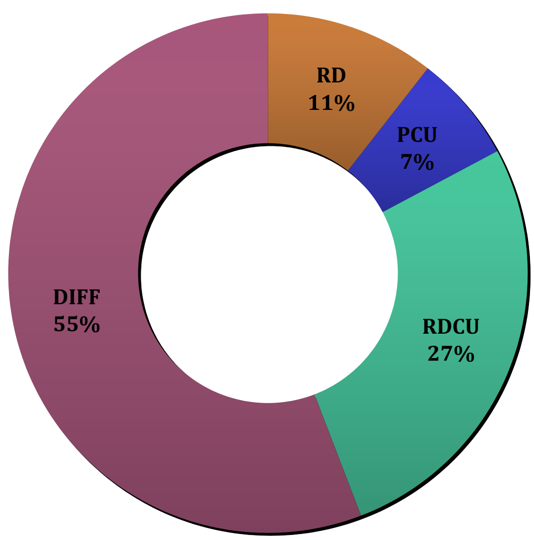
Table 3: Estimation results.

Method: Generalized Least Squares				
Characteristics: heteroskedastic with cross-sectional correlation				
Dependent variable: TFP growth				
	Coefficients	Standard Error	t-stat	$Prob > \|t\|$
β^{RD}	1.808	0.350	5.16	0.000
β^{PCU}	0.094	0.018	5.10	0.000
β^{RDCU}	0.020	0.006	3.57	0.000
β^{DIFF}	0.020	0.005	4.17	0.000
$const$	0.992	0.005	221.08	0.000
Wald test $\chi^2(7) = 64.340$				
Log likelihood=1627.004				
Prob> $\chi^2 = 0.000$				

Source: own computations.

In Figure 4, we present the decomposition of cross-country average TFP growth into components attributable to $R\&D$, PCU, $R\&DCU$ and $DIFF$.

Figure 4: Decomposition of TFP growth averaged across OECD countries

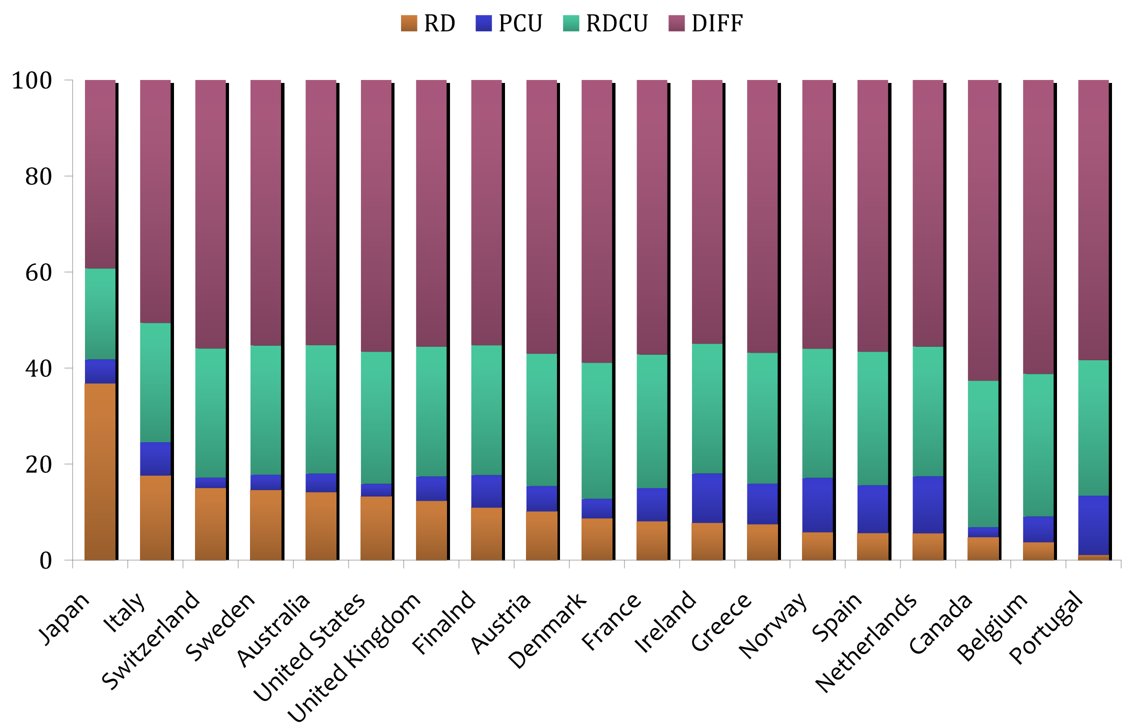


Source: Own calculations.

The results parallel the ones obtained for the Malmquist productivity index, but a few striking similarities and differences should be mentioned. First of all, the impact of domestic R&D on technological progress measured as TFP change is 11% which is exactly the same number as before. Secondly, there is less room for catching-up with the frontier if a Cobb–Douglas technology is imposed: this is reflected in lower shares of both our measures of catching-up. Thirdly, this difference is captured by a respective increase in the share of hi-tech import-based technology diffusion.

Turning to the question of the TFP growth breakdown across considered countries (Figure 5), we see no qualitative differences between the two specifications: it is again the case that countries which are the furthest from the world technology frontier benefit most from catching-up but least from domestic R&D (Portugal, the Netherlands, Spain, Norway; also Canada and Belgium, but for these two countries this is mostly R&D-supported catching-up).

Figure 5: TFP growth breakdown across OECD countries

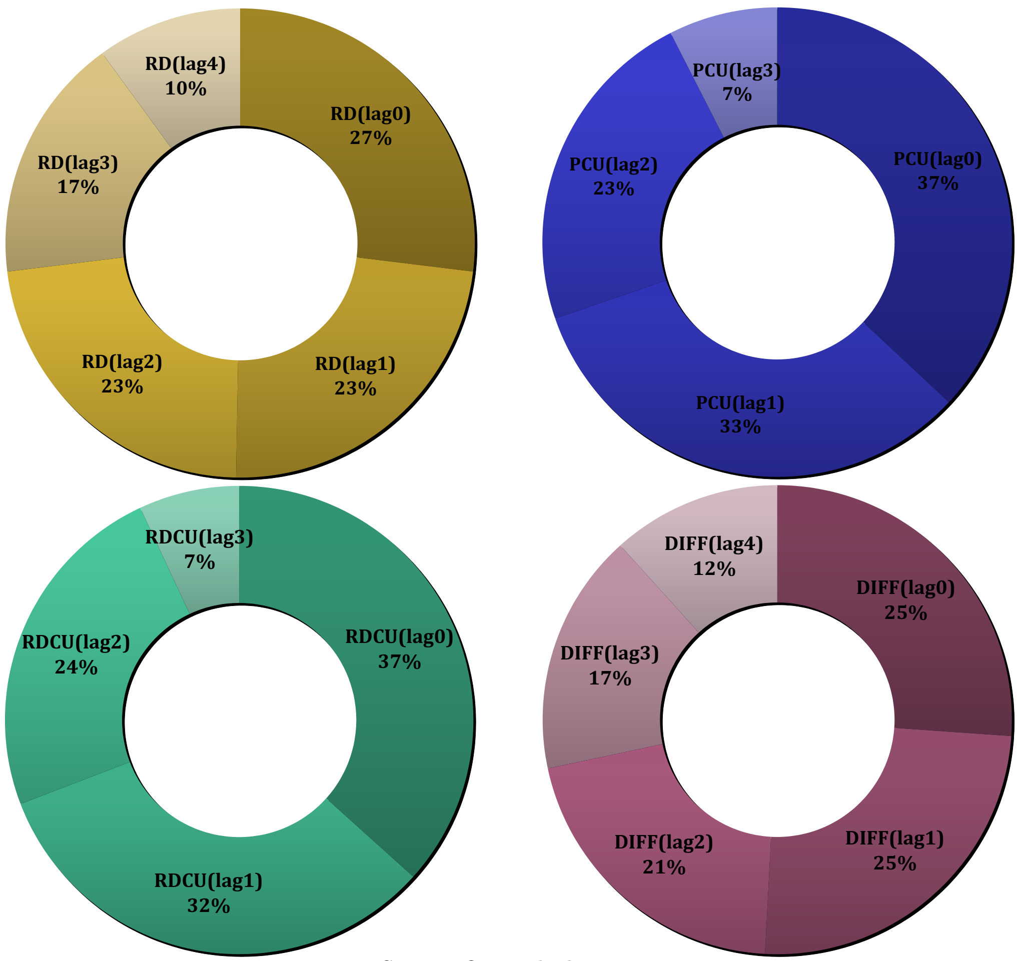


Source: Own calculations.

It is also again true that Japan, Italy, Switzerland, Sweden, Australia and the United States constitute the group of countries which benefited most from own R&D. Again Japan is the country which benefited least from international technology diffusion which might be due to geographic or other non-economic differences between Japan and other countries in our sample.

In Figure 6, we present the breakdown of the TFP growth decomposition across the temporal dimension: we assess the role of respective lagged values of each particular explanatory variable.

Figure 6: Decomposition of TFP growth averaged across OECD countries: lag structure



Source: Own calculations.

4.3 Questions for further research

The conclusions resulting from the DEA-based decomposition and the econometric exercise presented above enable us to formulate two additional issues to be raised in further research on the subject.

1. The basic limitation of the research described above consists in employing highly aggregated data. In particular, within this frame it is impossible to study sectoral reallocation effects, which as proved by Caselli and Tenreyro (2005) and OECD (2003) might have largely contributed to the productivity growth of some European economies.

2. R&D is a very broad economic category that encompasses inputs both from the public and the business sector. The distinction between these two origins of R&D efforts has not been made in the paper due to lack of consistent and reliable data for the time slot selected in the research. Nevertheless, further research in this field should possibly be extended so as to resolve this problem.

5 Conclusion

This paper decomposes productivity growth across highly developed OECD countries in the period 1972-2000 into components attributable to domestic R&D output, pure and R&D-driven catch-up effects, and international technology diffusion via imports of hi-tech products.

The methodology employed in the current research consisted of decomposing productivity growth into appropriately constructed measures of domestic R&D output, technological catch-up and technology diffusion. Domestic R&D output has been derived using the Schumpeterian specification taken from fully endogenous R&D-based growth models while the catch up and technology diffusion terms are based upon accumulation of human-capital augmented and weighted sums of "technology flows", proxied by appropriate measures of technological distance between the source and destination country in each pair.

To sum up, all of the variables considered had a positive and significant impact on productivity growth of highly developed OECD countries within the period of 1972-2000.

On average, own R&D output explains around 11% of this process, a pure catching-up process – between 7% and 16%, an R&D-supported catching-up process – between 27% and 31%, and technology diffusion driven by hi-tech imports – between 42% and 55%.

The contribution of each productivity growth component varied significantly in the sample, though. The most innovative countries – in terms of R&D output share in the overall productivity growth – were Japan, Italy, Switzerland, Sweden, Australia, and the United States. R&D output is however by no means the sole source of fast productivity growth. The front-runners in this field: Austria, Canada, Ireland and Norway, achieved such high increments in their productivity only because they successfully combined their efforts to simultaneously innovate new technologies and adopt foreign ones.

6 Appendix

List of hi-tech products from OECD's International Trade by Commodity Statistics (number refers to the reference number in the database):

- 28: Inorganic chemicals, precious metal compound, isotope
- 29: organic chemicals
- 30: pharmaceutical products
- 37: Photographic or cinematographic goods
- 38: Miscellaneous chemical products
- 39: Plastics and articles thereof
- 85: Electrical, electronic equipment
- 86: Railway, tramway locomotives, rolling stock, equipment
- 87: Vehicles other than railway, tramway
- 88: Aircraft, spacecraft, and parts thereof
- 89: Ships, boats and other floating structures
- 90: Optical, photo, technical, medical, etc apparatus
- 91: Clocks and watches and parts thereof
- 92: Musical instruments, parts and accessories

References

[1] Badunenko, O., D.J. Henderson, V. Zelenyuk (2008), "Technological Change and Transition: Relative Contributions to Worldwide Growth During the 1990's", *Oxford Bulletin of Economics and Statistics* 70(4), 461-492.

[2] Barro, R.J. (1999), "Notes on Growth Accounting", *Journal of Economic Growth* 4, 119-137.

[3] Barro, R.J., J.-W. Lee (2001), "International Data on Educational Attainment: Updates and Implications", *Oxford Economic Papers* 53(3), 541-563.

[4] Benhabib, J., M. M. Spiegel (2005), "Human Capital and Technology Diffusion" [In:] P. Aghion, S. Durlauf (eds.), *Handbook of Economic Growth*. Elsevier, Amsterdam.

[5] Caselli F. (2005), "Accounting for Cross-Country Income Differences" [In:] P. Aghion, S. Durlauf (eds.), *Handbook of Economic Growth*. Elsevier, Amsterdam.

[6] Caselli F., S. Tenreyro (2005), "Is Poland the Next Spain?", NBER Working Paper W11045.

[7] Cohen, D., M. Soto (2007), "Growth and Human Capital: Good Data, Good Results", *Journal of Economic Growth* 12(1), 51-76.

[8] Comin, D. (2004), "R&D: A Small Contribution to Productivity Growth", *Journal of Economic Growth* 9(4), 391-421.

[9] de la Fuente, A., R. Doménech (2006), "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?", *Journal of the European Economic Association* 4(1), 1-36.

[10] Duffy, J., C. Papageorgiou (2000), "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification", *Journal of Economic Growth* 5, 87-120.

[11] Färe, R., S. Grosskopf, M. Noriss, Z. Zhang (1994), "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries", *American Economic Review* 84(1), 66-83.

[12] Fernandez, C., E. Ley, M.F.J. Steel (2001), "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics* 100(2), 381-427.

[13] Fried, H.O., C.A. Knox Lovell, S.S. Schmidt, eds. (1993), *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford University Press.

[14] Griffith, R., S. Redding, J. Van Reenen (2004), "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Countries", *Review of Economics and Statistics* 86(4), 883-895.

[15] Growiec, J. (2008), "Productivity Differences Across OECD Countries, 1970–2000: The World Technology Frontier Revisited", IBS Working Paper 1/2008.

[16] Ha, J., P. Howitt (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.

[17] Henderson, D. J., R. R. Russell (2005), "Human Capital and Convergence: A Production–Frontier Approach", *International Economic Review* 46(4), 1167-1205.

[18] Heston, A., R. Summers, B. Aten (2006), "Penn World Table Version 6.2", Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

[19] Hodrick, R. J., E. C. Prescott (1997), "Postwar U.S. Business Cycles: An Empirical Investigation", *Journal of Money, Credit and Banking* 29(1), 1-16.

[20] Im, K. , M. H. Pesaran, Y. Shin (2003) "Testing for Unit Roots in Heterogeneous Panels",*Journal of Econometrics* 115, 53-74.

[21] Jerzmanowski, M. (2007), "Total Factor Productivity Differences: Appropriate Technology Vs. Efficiency", *European Economic Review* 51, 2080-2110.

[22] Jones, C. I. (1999), "Growth: With or Without Scale Effects", *American Economic Review Papers and Proceedings* 89, 139–144.

[23] Kumar, S., R. R. Russell (2002), "Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence", *American Economic Review* 92(3), 527-548.

[24] Madsen, J. B. (2007), "Technology Spillover Through Trade and TFP Convergence: 135 Years of Evidence for the OECD Countries," *Journal of International Economics* 72(2), 464-480.

[25] Madsen, J. B. (2008a), "Semi-Endogenous Versus Schumpeterian Growth Models: Testing the Knowledge Production Function Using International Data", *Journal of Economic Growth* 13(1), 1-26.

[26] Madsen, J. B. (2008b), "Economic Growth, TFP Convergence and the World Export of Ideas: A Century of Evidence", *Scandinavian Journal of Economics* 110(1), 145-167.

[27] Nelson, R. R., E. S. Phelps (1966), "Investment in Humans, Technological Diffusion, and Economic Growth", *American Economic Review* 256, 69-75.

[28] OECD (2003), *Sources of Economic Growth*, Paris.

[29] Sala-i-Martin, X., G. Doppelhofer, R.I. Miller (2004), "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review* 94(4), 813-835.

[30] Solow, R.M. (1957), "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics* 39, 312-320.

[31] Wooldridge, J.M. (2007), “Econometric Analysis of Cross Section and Panel Data”, MIT Press.