



Munich Personal RePEc Archive

**Productivity change using growth  
accounting and frontier-based approaches  
– Evidence from a Monte Carlo analysis**

Giraleas, Dimitris and Emrouznejad, Ali and Thanassoulis,  
Emmanuel

Aston University

July 2011

Online at <https://mpra.ub.uni-muenchen.de/37429/>  
MPRA Paper No. 37429, posted 18 Mar 2012 13:28 UTC

# Productivity change using growth accounting and frontier-based approaches – Evidence from a Monte Carlo analysis

Dimitris Giraleas<sup>1</sup>, Ali Emrouznejad and Emmanuel Thanassoulis

Operations and Information Management, Aston Business School, Aston University, Birmingham, UK

**Abstract:** This study presents some quantitative evidence from a number of simulation experiments on the accuracy of the productivity growth estimates derived from growth accounting (GA) and frontier-based methods (namely Data envelopment Analysis-, Corrected ordinary least squares-, and Stochastic Frontier Analysis-based Malmquist indices) under various conditions. These include the presence of technical inefficiency, measurement error, misspecification of the production function (for the GA and parametric approaches) and increased input and price volatility from one period to the next. The study finds that the frontier-based methods usually outperform GA, but the overall performance varies by experiment. Parametric approaches generally perform best when there is no functional form misspecification, but their accuracy greatly diminishes otherwise. The results also show that the deterministic approaches perform adequately even under conditions of (modest) measurement error and when measurement error becomes larger, the accuracy of all approaches (including stochastic approaches) deteriorates rapidly, to the point that their estimates could be considered unreliable for policy purposes.

**Keywords:** Data envelopment analysis, Productivity and competitiveness, Simulation, Stochastic Frontier Analysis, Growth accounting

---

<sup>1</sup> Corresponding Author: Email address: [giralead@aston.ac.uk](mailto:giralead@aston.ac.uk) (D Giraleas)

# 1 Introduction

The study of productivity is a very important topic. The UK Office of National Statistics ((ONS, 2007)) states that: 'Statistics relating to productivity are vital to understanding the economy and how it changes'. It also states that: 'it is crucial that both experts and the general public can depend on the accuracy and relevance of ONS productivity measures'. The Organisation for Economic and Social Development (OECD) also states that one of its major aims is to improve the measurement of productivity growth<sup>2</sup>.

The pursuit of productivity growth and productivity convergence is also one of the central goals of the European Union (EU). Probably the main instruments to achieve those goals are the so-called Structural funds, which are distributed based on Gross Domestic Product (GDP) per capital differentials between the various EU regions. Changes in GDP per capital are also used as simple measures of productivity growth and although probably sufficient for setting policy at this stage, a more refined productivity indicator is required to evaluate the effects of the funds and the degree of convergence. The issue of convergence is critical, since the underlying aim of the Structural funds is to increase GDP by providing the relatively poorer regions with the tools to achieve the productivity/efficiency potential of the more advanced regions, rather than raising GDP simply through factor accumulation.

More complex approaches that seek to estimate Total Factor Productivity growth (TFP) can provide the required granularity of information, by examining the sources of GDP growth that are not due to such factor accumulation. The EU seems to support the development and use of such approaches, given the emphasis the Directorate General for Economic and Financial Affairs (DG-ECFIN) has placed on the EU KLEMS project ((EU KLEMS, 2008)), an EU-wide research project that aims to provide estimates of aggregate TFP growth in the EU member states together with the data necessary for the estimation. The DG-ECFIN (Koszerek, 2007) states that the productivity indicators provided by EU KLEMS are '*essential for understanding recent EU productivity trends*', '*fundamental in assessing progress with the Lisbon Strategy*', '*can complement the "Structural Indicators" Programme*', and '*provide an additional data source for refining the potential growth rate estimates used in the EU's budgetary surveillance process*'.

---

<sup>2</sup> [http://www.oecd.org/about/0,3347,en\\_2825\\_30453906\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/about/0,3347,en_2825_30453906_1_1_1_1_1,00.html), accessed 14 January 2011

Productivity growth in the EU KLEMS database is estimated based on growth accounting (GA). GA is an index number-based methodology for measuring productivity growth which is based in the early work of Tinbergen (Tinbergen, 1942) and independently, Solow (Solow, 1957) and is the method of choice when measuring aggregate (ie country- or sector-wide) productivity growth for most interested agents, namely statistical agencies (national and international), central banks and government bodies (see for example the US Bureau of Labor Statistics technical note on multifactor productivity<sup>3</sup> and ONS (ONS, 2007)). A major factor in the widespread adoption of GA is the fact that estimates can be (relatively) easily produced using country- or sector-specific National Accounts data, without recourse to information from outside the country or the sector examined; on the other hand, GA requires the adoption of a number of simplistic (potentially unrealistic) assumptions, most notably those relying on the existence of perfect competition, which could lead to unreliable estimates.

Given the stated need for accurate productivity growth estimates, the first aim of this study is to assess the impact on the accuracy of the GA estimates when some of the assumptions central to the notion of perfect competition are violated. This is achieved by undertaking a number of simulation experiments, which utilise randomly generated data for which the parameters of interest (most importantly productivity change) are known a-priori; when GA (or any other productivity change measurement approach) is applied to the same dataset, a measure of the overall accuracy of the approach can be devised by comparing the estimate of productivity change to its true value.

Frontier-based methods offer an attractive alternative for the measurement of aggregate productivity change. Unlike the more traditional GA methods, they allow for the production to occur inside the frontier, thereby explicitly allowing for inefficiency in the production process and relaxing the stringent assumptions required when using growth accounting methods. In addition, frontier-based methods also allow for the decomposition of productivity growth, which could be of great interest to the users of productivity change estimates.

There are a number of applications of frontier based methods for the measurement of aggregate productivity growth in the academic literature. Färe et al. (Färe, Grosskopf, Norris, & Zhang, 1994) was one of the first studies that utilised Data Envelopment

---

<sup>3</sup> <http://www.bls.gov/mfp/mprtech.pdf>, accessed 14 January 2011

Analysis (DEA), the more widely-used non-parametric frontier based approach, to construct Malmquist indices of productivity growth; the approach has since been adopted in numerous other studies (for a comprehensive list of applications of DEA-based Malmquist indices see (Fried, Lovell, & Schmidt, 2008) and (Del Gatto, Di Liberto, & Petraglia, 2008)). Kumbhakar et al. (Kumbhakar & Lovell, 2000) introduced another way to construct a Malmquist index of productivity growth that relies on parametric frontier models, such as Corrected Ordinary Least Squares (COLS) and Stochastic Frontier Analysis (SFA); such models have also been widely used in the literature (see Sharma et al. (Sharma, Sylwester, & Margono, 2007) for a list of sample studies).

However, despite the adoption of such frontier-based methods in the academic literature and the theoretical advantages offered by frontier-based methods compared to the more traditional GA approach, there has been limited research on quantifying how these advantages translate into improved accuracy of the resulting productivity change estimates and under which conditions one frontier-based approach is more accurate than another. As such, the second aim of this study is to employ the aforementioned simulation experiments to provide quantitative evidence on the accuracy of the more widely adopted frontier-based approaches, namely DEA-, COLS- and SFA-based Malmquist indices, under a number of conditions that violate the assumptions made under perfect competition.

In more detail, this research aims to examine the accuracy of both GA and frontier-based productivity change estimates:

- when technical inefficiency, in various degrees of severity, is present,
- when inputs and prices are volatile from one period to the next,
- when the production function is miss-specified, and finally
- when the factors of production are measured inaccurately (again in various degrees of severity).

## **2 Methodology of the current research**

### **2.1 Productivity measurement approaches considered**

Each simulation experiment examines the performance of the following approaches:

- GA,
- DEA-based circular Malmquist indices,

- COLS-based Malmquist indices, and
- SFA-based Malmquist indices, (only when measurement noise is included in the experiment).

All frontier-based approaches examined in this analysis rely on the notion of what has come to be known as the Malmquist productivity index (Diewert, 1992), which has been used extensively in both the parametric (see for example Kumbhakar et al. (Kumbhakar & Lovell, 2000)) and the non-parametric (see for example Thanassoulis (Thanassoulis, 2001)) setting. Furthermore, the productivity index produced by GA can be considered as a special case of the Malmquist productivity index (see OECD (OECD, 2001)).

Given the nature of the approaches considered, the analysis focuses on the production side of the economic process.

### Growth Accounting

Growth Accounting (GA) is an index number-based approach that relies on the neo-classical production framework, and seeks to estimate the rate of productivity change residually, ie by examining how much of an observed rate of change of a unit's output can be explained by the rate of change of the combined inputs used in the production process. There are many modifications that could be applied to the more general GA setting ((Balk, 2008); (Del Gatto, et al., 2008)); however, most applications still utilise 'traditional' growth accounting methods, as described in OECD (OECD, 2001) (see for example O'Mahony et al. (O'Mahony & Timmer, 2009)).

GA postulates the existence of a production technology that can be represented parametrically by a production function relating gross output ( $Y$ ), to primary inputs labour ( $L$ ) and capital services ( $K$ ) as well as intermediate inputs such as material, services or energy ( $M$ ).

$$Y = F(K, L, M) \tag{1}$$

If gross output is measured net of intermediate inputs, ie using a Gross Value Added (GVA) measure, (1) becomes:

$$Y_{GVA} = F(K, L) \tag{2}$$

GA assumes that productivity changes ( $TFP$ ) are Hicks-neutral type, i.e. they correspond to an outward shift of the production function, such that:

$$Y_{GVA} = F(K, L) \times TFP \quad (3)$$

A number of assumptions are required to parameterise (3), namely that:

- the production function is Cobb-Douglas and exhibits constant returns to scale;
- each assessed unit minimises the costs of inputs for any desired level of output and can adjust the level of primary inputs that it utilises at any moment and without additional costs;
- input markets are perfectly competitive and all production happens on the frontier;
- all relevant inputs and outputs are taken into account and measured without error.

For a more detailed discussion on the assumptions required for GA, see Annex 3 of the OECD manual (OECD, 2001).

If the above assumptions hold, once the Cobb-Douglas production function is differentiated with respect to time, the rate of change in output is equal to the sum of the weighted average of the change in inputs and the change in productivity. The input weights are the output elasticities of each factor of production; under perfect competition conditions, the marginal revenue generated by each factor is equal to its price and, as such, the output elasticity of each factor is equal to its share in the total value of production.

Therefore, productivity change is estimated by:

$$\frac{d \ln TFP_i^{GA}}{dt} = \frac{d \ln y_i}{dt} - S_i^L \frac{d \ln L_i}{dt} - S_i^K \frac{d \ln K_i}{dt} \quad (4)$$

where  $S_i^L$  is the average share of labour in periods  $t$  and  $t-1$ ,  $S_i^K$  is the average share of capital in  $t$  and  $t-1$  given by:

$$S_i^L = \left( \frac{w_{it}^L L_{it}}{p_{it} Y_{it}} + \frac{w_{it-1}^L L_{it-1}}{p_{it-1} Y_{it-1}} \right) / 2 \quad (5)$$

$$S_i^K = \left( \frac{w_{it}^{K,GA} K_{it}}{p_{it} Y_{it}} + \frac{w_{it-1}^{K,GA} K_{it-1}}{p_{it-1} Y_{it-1}} \right) / 2 \quad (6)$$

It should be noted that the price of capital is not observable; therefore analyses that utilise GA usually rely on an imputed price of capital. This is discussed in more detail in section 2.2 (Price data). The use of arithmetic averages for the input shares was

adopted for consistency with the EU KLEMS methodology. An alternative option would be to use geometric averages; this has been examined in some of the initial simulation experiments, but it had almost no impact on the resulting estimates. As such, results using geometric averages of shares are not reported in this paper.

As is apparent from the above, one of the major advantages of GA is that it does not require any information outside of the assessed unit to estimate productivity growth. To do so however, the analysis must adopt a number of restrictive assumptions as noted above and also have access to price information, which is necessary to parameterise the aggregate production function.

### **DEA-based Circular Malmquist index**

The most common non-parametric approach for productivity measurement utilises Data Envelopment Analysis (DEA) to construct Malmquist indices of productivity change. This approach was first proposed by Caves et al. (Caves, Christensen, & Diewert, 1982) and later refined by Färe et al. (Färe, et al., 1994).

This study utilises the notion of a circular Malmquist-type index (thereafter referred to as circular Malmquist), as first proposed by Pastor et al. (Pastor & Lovell, 2005) and refined by Portela et al. (Portela & Thanassoulis, 2010).

The circular Malmquist index is based on the observation that, although a measure of distance between two multidimensional points observed at two different time periods can be satisfactorily calculated directly, as per the 'traditional' Malmquist index, a similar measure of distance can be calculated indirectly, by comparing the multidimensional points of the two periods relative to a common reference point, or in this case, to a common frontier. This common frontier is defined as the 'meta-frontier' and since it envelopes all data points from all periods, it allows for the creation of a Malmquist-type index which is circular. To draw this 'meta-frontier', one must assume that convexity holds for all data points across different time-periods. This actually translates to the assumption that what was technologically feasible in a given time period will always be feasible in any future time period, a standard assumption in so-called sequential technology (Tulkens & Vanden Eeckaut, 1995). Distances can then be measured using the 'standard' DEA models.

The main advantages of the circular Malmquist index relative to the 'traditional' (Färe 1994) Malmquist index are the ease of computation and the ability to accommodate



unbalanced panel data. For a more detailed discussion, see Portela et al. (Portela & Thanassoulis, 2010).

### Corrected OLS

Corrected OLS is a deterministic, parametric approach and one of the numerous ways that have been suggested to 'correct' the inconsistency of the OLS-derived constant term of the regression when technical inefficiency is present in the production process.

Two different COLS model specifications were tested within these simulations. Both are based on a pooled regression model (ie all observations are included in the same model with no unit-specific effect). The first model assumes a Cobb-Douglas functional form and is used for those experiments where the data is generated using the Cobb-Douglas production function. In more detail, the functional form used is:

$$Y_{it} = L_{it}^{\alpha^*} K_{it}^{\beta^*} t^{\gamma^*} \exp(\varepsilon_{it}^*) \quad (7)$$

where  $\varepsilon_{it}^*$  are the estimated OLS residuals. The standard logarithmic transformation converts (7) into:

$$\ln Y_{it} = \alpha^* \ln L_{it} + \beta^* \ln K_{it} + \gamma^* t + \varepsilon_{it}^* \quad (8).$$

It should be noted that the above specification matches perfectly the data generating process, when measurement error is not included in the experiments.

The second COLS model specification assumes a translog functional form and is used, together with the Cobb-Douglas functional form specification, for those simulation experiments where the data is generated using the piecewise-linear production function. The translog COLS model is given by:

$$\begin{aligned} \ln Y_{it} = & a_i + \beta_L \ln L_{it} + \beta_K \ln K_{it} + \gamma_t t + \frac{1}{2} \beta_{LL} (\ln L_{it})^2 + \frac{1}{2} \beta_{KK} (\ln K_{it})^2 + \frac{1}{2} \gamma_{tt} t^2 \\ & + \beta_{KL} \ln K_{it} \ln L_{it} + \gamma_{Kt} \ln K_{it} t + \gamma_{Lt} \ln L_{it} t + \varepsilon_{it}^* \end{aligned} \quad (9)$$

Inefficiency estimates are derived by:

$$-u_{it}^* = \varepsilon_{it}^* - \max(\varepsilon_{it}^*) \quad (10)$$

Productivity change is calculated based on the same formula as used for the calculation of true productivity change, substituting the true parameters with the various parametric estimates (see section 2.2 and Kumbhakar et al. (Kumbhakar & Lovell, 2000)). So, productivity change is given by:

$$d \ln TFP_{it}^{COLS} / dt = d \ln EC_{it}^{COLS} / dt + d \ln TC_{it}^{COLS} / dt \quad (11)$$

where  $EC_{it}^{COLS}$  is the COLS-estimated efficiency change and  $TC_{it}^{COLS}$  is the COLS-estimated technical change. Note that (11) does not include a scale efficiency change component, since all experiments assume constant returns to scale (see section 2.2 for more details).

$$d \ln EC_{it}^{COLS} / dt = (u_{it-1}^* - u_{it}^*) \quad (12)$$

$$d \ln TC_{it}^{COLS} / dt = \gamma^* \quad (13)$$

for the Cobb-Douglas function and

$$d \ln TC_{it}^{COLS} / dt = \gamma_t + t\gamma_{tt} + \gamma_{Kt} \ln K_{it} + \gamma_{Lt} \ln L_{it} \quad (14)$$

for the translog function.

### Stochastic frontier analysis

The pre-eminent parametric frontier-based approach is Stochastic Frontier Analysis, which was developed independently by Aigner et al. (Aigner, Lovell, & Schmidt, 1977) and by Meeusen et al. (Meeusen & van Den Broeck, 1977). The approach relies on the notion that the observed deviation from the frontier could be due to both genuine inefficiency but also random effects, including measurement error. SFA attempts to disentangle those random effects by decomposing the residual of the parametric formulation of the production process into noise (random error) and inefficiency.

As is the case with the COLS approach, two separate SFA model specifications are used: a Cobb-Douglas functional form is employed for those experiments where the data are generated through a Cobb-Douglas production function, and both a Cobb-Douglas and translog functional form for those experiments where the data are generated through a piecewise linear production function. The models are very similar to those used under COLS; in fact, the only difference lies in the specification of the residual.

In more detail, the Cobb-Douglas model is given by:

$$\ln Y_{it} = \alpha^* \ln L_{it} + \beta^* \ln K_{it} + \gamma^* t + v_{it} - u_{it} \quad (15)$$

whereas the translog model is given by:

$$\ln Y_{it} = a_i + \beta_L \ln L_{it} + \beta_K \ln K_{it} + \gamma_t t + \frac{1}{2} \beta_{LL} (\ln L_{it})^2 + \frac{1}{2} \beta_{KK} (\ln K_{it})^2 + \frac{1}{2} \gamma_{tt} t^2 + \beta_{KL} \ln K_{it} \ln L_{it} + \gamma_{Kt} \ln K_{it} t + \gamma_{Lt} \ln L_{it} t + v_{it} - u_{it} \quad (16)$$

where  $u_{it}$  represents the inefficiency component (and as such  $u_{it} \geq 0$ ) and  $v_{it}$  represents measurement error ( $v_{it} \sim N(0, \sigma_v^2)$ ). The inefficiency component is estimated based on the JMLS (Jondrow, Knox Lovell, Materov, & Schmidt, 1982) estimator.

Two different distributions for the inefficiency component are tested:

- the exponential distribution,  $u_{it} \sim \text{Exp}(\sigma_u)$
- the half-normal distribution,  $u_{it} \sim N^+(0, \sigma_u^2)$

When the data is generated using the Cobb-Douglas production function, the exponential Cobb-Douglas SFA model is perfectly specified, since the data generation process also generates the inefficiency values from an exponential distribution. The estimates from the half-normal distribution are included in the experiments to examine the impact of misspecification in the inefficiency distribution to the SFA productivity change estimates. Productivity change is measured in exactly the same way as with COLS, ie using (11)-(14).

## 2.2 Data generating process

Since the focus on this analysis is on the production side of the economic process, information on inputs and output(s) is sufficient for the estimation of productivity change under the frontier-based approaches. However, GA also requires information on prices for both inputs and output(s) in order to parameterise the production function (see section 2.1), so price information that is consistent with the quantities of inputs used and outputs produced by each assessed unit also needs to be generated.

Given that the analysis includes both parametric and non-parametric approaches, the choice of the production function which would be used to generate the output values for the simulations can have a significant impact on the accuracy of the resulting estimates. If the functional form adopted by the parametric approaches matches the functional form of the underlying production function, it is expected that the resulting parametric-based estimates would be more accurate relative to when functional form misspecification is present. In addition, GA implicitly assumes that the underlying production function is Cobb-Douglas, and as such it would also be pertinent for the

analysis to examine what happens when this assumption is violated. To assess the effects of such functional form misspecification on the overall accuracy of the estimates, two sets of simulations are undertaken: the first set adopts a Cobb-Douglas production function and the second set adopts a piecewise-linear production function. In all cases, the production function assumes that two inputs are used to produce a single output, which is also the norm when measuring aggregate productivity change with value added as the output and labour and capital quantities as the inputs.

For the first set of simulations, which utilise a Cobb-Douglas production function, output is given by:

$$Y_{it} = L_{it}^{\alpha} K_{it}^{\beta} t^{\gamma} \exp(v_{it} - u_{it}) \quad (17)$$

where  $Y_{it}$  is the output of unit  $i$  in time  $t$ ,  $L_{it}$  is the labour input of unit  $i$  in time  $t$ ,  $K_{it}$  is the capital input of unit  $i$  in time  $t$ ,  $v_{it}$  is measurement error (noise) and  $u_{it}$  is the technical efficiency of unit  $i$  in time  $t$ . An element of technical change is also included in the form of the time trend  $t$ . Output elasticities are given by the parameters  $\alpha$  and  $\beta$  for labour and capital, while  $\gamma$  represents technical change. The values for the elasticity parameters in these experiments are set to  $\alpha=\beta=0.5$  and  $\gamma=0.02$ , ie the experiments assume constant returns to scale. Measurement error is normally distributed with zero mean and variance that changes according to aims of each simulation experiment (some simulation experiments assume no measurement error, while others assume varying degrees of measurement error).

For the second set of simulations, which utilise a piecewise-linear production function, output is given by:

$$y_{it}^* = \begin{cases} 1.84L_i + 0.13K_i & \text{for } L_i/K_i \leq 0.3 \\ 1.53L_i + 0.22K_i & \text{for } L_i/K_i > 0.3 \text{ and } L_i/K_i \leq 0.73 \\ 1.04L_i + 0.59K_i & \text{for } L_i/K_i > 0.73 \text{ and } L_i/K_i \leq 1.03 \\ 0.46L_i + 1.18K_i & \text{for } L_i/K_i > 1.03 \text{ and } L_i/K_i \leq 1.14 \\ 0.40L_i + 1.25K_i & \text{for } L_i/K_i > 1.14 \text{ and } L_i/K_i \leq 2.01 \\ 0.19L_i + 1.67K_i & \text{for } L_i/K_i > 2.01 \text{ and } L_i/K_i \leq 3.22 \\ 0.06L_i + 2.09K_i & \text{for } L_i/K_i > 3.22 \end{cases} \quad (18)$$

The labour coefficients, the number of 'pieces' (or facets) and the breakpoints in the above function were randomly generated, while the capital coefficients were calculated such that the above function would be convex in  $K$  and  $L$  (as in all

input/output correspondences belong to a convex set), monotonic, continuous and display constant returns to scale<sup>4</sup>.

The  $y_i^*$  parameter represents 'clean' output, ie before the effects of inefficiency, technical change and possible measurement error are included. The output value used in the simulation experiments includes all those elements and is given by:

$$y_{it} = y_i^* TE_{it} TC_t \exp(v_{it}) \quad (19)$$

where  $TE_{it}$  represents technical efficiency and is given by

$$TE_{it} = \exp(-u_{it}) \quad (20).$$

$TC_t$  represents technical change and is a function of time (t) and a constant  $\gamma$  and is given by:

$$TC_t = t^\gamma \quad (21),$$

and  $v_{it}$  represents measurement error, which is normally distributed with zero mean and variance that changes according to the aims of each simulation experiment.

All simulations rely on a panel dataset of 20 units, observed over five periods (ie the total number of observations is 100). The input sets for all units in the first period are randomly generated following a uniform distribution  $U[0,1]$ ; in subsequent periods, they are scaled by a random, normally-distributed number; the default assumption is that this scaling factor follows  $N(0,0.10)$ , but this paper also examines the condition of increased volatility, by setting the standard deviation to 0.25. It should be mentioned here that the same scaling factor is used to generate input prices (as discussed in the following section).

Efficiency is also randomly generated and follows the exponential distribution. Two cases, corresponding to different levels of average inefficiency are examined:

- for the 'average levels of inefficiency' experiments, the inefficiency term follows  $\text{Exp}(1/7)$ , which results in an average inefficiency of approximately 12%.
- for the 'higher levels of inefficiency' experiments, the inefficiency term follows  $\text{Exp}(1/2)$ , which results in an average inefficiency of approximately 32%.

---

<sup>4</sup> The function represents a production process under constant returns to scale since when input values are doubled, so does the total output, for all 'pieces' or facets of the function.

The definition of productivity change used for this analysis relies on the notion of what has come to be known as the Malmquist productivity index. The Malmquist productivity index is the product of the index of efficiency change  $EC_{it}$ , scale efficiency change  $SE_{it}$  and technical change  $TC_{it}$  (otherwise known as technological change or frontier shift). Taking logs and differentiating across time provides the definition of productivity change across time:

$$d \ln TFP_{it}^{true} / dt = d \ln EC_{it} / dt + d \ln SE_{it} / dt + d \ln TC_{it} / dt \quad (22)$$

All of the simulation experiments assume constant returns to scale and thus the change in scale efficiency measure can be ignored<sup>5</sup>. Thus, following Kumbhakar e al. (Kumbhakar & Lovell, 2000), the expression in (22) can be rewritten as:

$$\begin{aligned} d \ln TFP_{it}^{true} / dt &= d \ln EC_{it} / dt + d \ln TC_{it} / dt \\ &= (u_{it-1} - u_{it}) + \gamma \end{aligned} \quad (23)$$

Since all parameters in the right-hand side of the above equations are known in the generated dataset the calculation of true productivity change (in the context of the generated dataset) is trivial.

## Price data

To generate price information for the experiments consistent with each production function, the analysis relies on micro-economic production theory and assumes that each producer attempts to minimise costs (Samuelson, 1947), ie:

$$\begin{aligned} \min C_i &= w_i^L L + w_i^K K \\ s.t. & \\ y_i &= f(L_i, K_i) \end{aligned} \quad (24)$$

assuming that producers utilise two inputs, capital and labour, with prices  $w_i^K$  and  $w_i^L$  respectively to produce a given level of a single output. To explore the optimal solution for (24), the Lagrangian form is required, ie:

$$E = w_i^L L + w_i^K K + \lambda(y_i - f(L_i, K_i)) \quad (25)$$

where  $\lambda$  is the Lagrangian multiplier. Combining the first order conditions resulting from solving equation (25) yields:

---

<sup>5</sup> Scale efficiency for all units is always equal to one, and thus scale efficiency change is equal to zero.

$$\frac{w_i^L}{w_i^K} = \frac{\frac{\partial f(L_i, K_i)}{\partial L}}{\frac{\partial f(L_i, K_i)}{\partial K}} \quad (26)$$

which provides the structural relationship that links input prices to production characteristics. Note that due to the duality theory, the same relationship applies even if the producer is assumed to be output maximising.

For the simulation experiments that assume the Cobb-Douglas function specified in (17), equation (26) becomes:

$$\ln\left(\frac{L_{it}}{K_{it}}\right) = \ln\left(\frac{\alpha w_{it}^K}{\beta w_{it}^L}\right) \quad (27)$$

where  $\alpha$  is the output elasticity of labour,  $\beta$  is the output elasticity of capital,  $w_{it}^K$  and  $w_{it}^L$  are the prices of capital and labour respectively for unit  $i$  in time  $t$ .

For the simulation experiments that assume the piecewise-linear function specified in (18), equation (27) becomes:

$$\frac{w_i^L}{w_i^K} = \frac{\alpha_j}{\beta_j} \quad (28)$$

where  $\alpha_j$  is the output elasticity of labour for the  $j$ th 'piece' of the piecewise linear function,  $\beta_j$  is the output elasticity of capital for the  $j$ th 'piece' of the piecewise linear function and  $w_{it}^K$  and  $w_{it}^L$  are the prices of capital and labour respectively for unit  $i$  in time  $t$ .

Note that this analysis and (26) specifically assume that allocative inefficiency, is either zero or time-invariant for each assessed unit.

Given the above, input prices are generated using the following approach:

- First, prices for labour that are unique for each unit are generated for the first period of the analysis as random draws from a uniform distribution (U(0,0.1]).
- These values are then scaled by a random, normally-distributed number to generate values for the subsequent periods, similar to the approach used for the generation of the input quantities. As with the input quantities, the experiments test

- (27) or (28) are then used to calculate the true price of capital input, depending on whether the simulation experiments assume a Cobb-Douglas or a piecewise linear production function respectively. Note that the true price of capital is not observable by the researcher and as such, it is not used directly in the simulation experiments (this is discussed in more detail below).

Next, output prices are generated by equating total revenues to total costs:

$$p_{it}Y_{it}^* = w_{it}^K K_{it} + w_{it}^L L_{it} \quad (29)$$

$$p_{it} = \frac{w_{it}^K K_{it} + w_{it}^L L_{it}}{Y_{it}^*} \quad (30)$$

where  $p_{it}$  is the price of output of unit  $i$  in period  $t$  and  $Y_{it}^*$  is the efficient level of output of unit  $i$  in period  $t$ . By using the efficient level of output in equations (29) and (30), the analysis explicitly assumes that only the producers that operate on the frontier are able to fully recover their total costs; this way, the effects of technical inefficiency can be linked to total costs and revenues. Note that the above does not mean that an efficient producer achieves zero profits. Rather, the price of capital includes an element commonly referred to as the ‘user cost of capital’, which ensures that an efficient producer receives an ‘appropriate’ return on the capital invested (ie achieves a ‘normal’ level of profits). By extension, an inefficient producer will receive a lower return of capital. Also note that this analysis assumes that no producer has sufficient market power to generate ‘above normal’ profits.

The data generation process described above is fully consistent with the economic theory of production, but unfortunately produces data, and specifically data for the price of capital inputs, that are not available in the majority of real life applications and certainly within the bounds of the National Accounts data. In the majority of real-life situations, the ‘user cost of capital’, which is a component necessary for the calculation of the price of capital, is not observable; as such the real price of capital cannot be measured with certainty. So to calculate the full price of capital, most GA applications (see for example OECD (OECD, 2001) and O'Mahony et al. (O'Mahony & Timmer, 2009)) adopt an endogenous ‘user cost of capital’, which is calculated residually. This



is achieved by setting capital compensation (ie the cost of capital) to be equal to Value Added (which is equivalent to revenue in the setting of these simulations) minus the labour compensation (ie the cost of labour). Since the quantity of capital can be estimated using national account data, the price of capital based on an endogenous ‘user cost of capital’ can be derived by:

$$w_{it}^{K,GA} = \frac{P_{it}Y_{it} - w_{it}^L L_{it}}{K_{it}} \quad (31)$$

So, to ensure that the constructed data used for the simulations are similar to what is available in real-life applications, the analysis also uses this ‘GA-adjusted’ price of capital ( $w_{it}^{K,GA}$ ) to generate GA productivity change estimates. This modification ensures that input shares add up to one and thus allows the use of GA in such a way that is consistent with EU KLEMS and the methodology proposed by the OECD.

Finally, in a very few cases in the simulations, the cost of labour could exceed total revenue, and as such the GA-adjusted price of capital is negative. Although negative capital prices are not inconsistent with theory (Berndt & Fuss, 1986), they are incompatible with the standard GA framework, since they result in negative capital shares. To avoid this, the analysis follows the EU KLEMS practice of setting all instances of negative prices to zero.

### 2.3 The simulation experiments

As mentioned in the introduction, the aim of this analysis is to assess the impacts to the accuracy of the produced productivity change estimates from GA and a number of frontier-based approaches when the underlying data do not adhere to the perfect competition assumptions. To do so, elements of both technical inefficiency and measurement error (noise) are gradually introduced to the production function used to generate the simulated output. More specifically, different experiments are undertaken assuming two different levels of technical inefficiency: ‘average’ levels ( $u_{it} \sim Exp(1/7)$ ) and ‘higher’ levels ( $u_{it} \sim Exp(1/2)$ ) and three different levels of noise: zero noise ( $v_{it} = 0$  for all  $i$  and  $t$ ), ‘modest’ noise relative to inefficiency ( $v_{it} \sim N(0,0.05)$ ) and ‘extensive’ noise relative to inefficiency ( $v_{it} \sim N(0,0.2)$ ).

This paper also examines the impact of functional form misspecification in the estimates derived from the parametric approaches and GA; as noted in section 2.2,

this is achieved by subdividing the simulation experiments into two sets. Those in the first set (denoted as S1) use the data generating process that assumes a Cobb-Douglas production function, while those in the second set (denoted as S2) use the data generating process that assumes a piecewise linear production function.

Lastly, this paper examines the effects of input and price volatility from one period to the next; this is achieved by generating a new set of data for each simulation experiment which is based on a more volatile scaling factor (default scaling factor is randomly generated and follows  $N(0,0.10)$ , while the more volatile scaling factor follows  $N(0,0.25)$ ). The way all of the above parameters enter into the production function is described in detail in section 2.2. As a reminder, all data generated come from production functions that display constant returns to scale and also include an element of time-invariant technical change (which corresponds to approximately 1% p.a. increase in output).

**Table 1: Simulation experiments**

|             | Production function | Technical inefficiency | Noise       | Input and price Volatility assumptions |
|-------------|---------------------|------------------------|-------------|--|
| <b>S1.1</b> | Cobb-Douglas        | 'average' levels       | zero        | both 'default' and 'higher' volatility |
| <b>S1.2</b> | Cobb-Douglas        | 'higher' levels        | zero        | both 'default' and 'higher' volatility |
| <b>S1.3</b> | Cobb-Douglas        | 'average' levels       | 'extensive' | both 'default' and 'higher' volatility |
| <b>S1.4</b> | Cobb-Douglas        | 'average' levels       | 'modest'    | both 'default' and 'higher' volatility |
| <b>S2.1</b> | Piece-wise linear   | 'average' levels       | zero        | both 'default' and 'higher' volatility |
| <b>S2.2</b> | Piece-wise linear   | 'higher' levels        | zero        | both 'default' and 'higher' volatility |
| <b>S2.3</b> | Piece-wise linear   | 'average' levels       | 'extensive' | both 'default' and 'higher' volatility |
| <b>S2.4</b> | Piece-wise linear   | 'average' levels       | 'modest'    | both 'default' and 'higher' volatility |
| <b>S2.5</b> | Piece-wise linear   | 'higher' levels        | 'extensive' | both 'default' and 'higher' volatility |
| <b>S2.6</b> | Piece-wise linear   | 'higher' levels        | 'modest'    | both 'default' and 'higher' volatility |

It should also be mentioned that the original analysis also tested whether the inclusion of fully efficient units would have any impact on the summary accuracy measures<sup>6</sup>; the analysis found that the accuracy measures from the simulations which included fully efficient units are almost indistinguishable from the base case and thus these results are not reported in this paper.

<sup>6</sup> The data generation methodology implemented for these simulations ensures that no unit is fully, ie 100%, technically efficient.

## 2.4 Measures of accuracy

The productivity change estimates produced by each approach are compared to the true rate of productivity change (derived by (23)). Three different measures are employed to judge the accuracy of the estimates under each approach:

The **mean absolute deviation (MAD)** of productivity change, given by:

$$MAD = \sum_{i=1, t=1}^{n, 5} |TFP_{it}^{TRUE} - TFP_{it}^{EST}| / N \quad (32)$$

where  $TFP_{it}^{TRUE}$  is true productivity change and  $TFP_{it}^{EST}$  is the estimated productivity change derived from the approach under examination. The MAD measure provides a robust central estimate of the overall accuracy of each approach, regardless of the sign of the deviation between the true and the estimated value. Lower MAD scores represent better overall accuracy.

The **mean square error (MSE)** of productivity change, given by:

$$MSE = \sum_{i=1, t=1}^{n, 5} (TFP_{it}^{TRUE} - TFP_{it}^{EST})^2 / N \quad (33)$$

The MSE measure plays a complementary role to the MAD measure, since it gives more weight to larger deviations and thus provides a better picture of 'extreme' deviation. Lower MSE scores represent better overall accuracy.

The **mean absolute deviation of the 25<sup>th</sup> percentile ('top' MAD or TMAD)** of productivity change, which is the MAD of the top 25% of observations when sorted in descending order according to the absolute deviation from the true value. In other words, the analysis calculates the absolute deviation of all observations and then takes into account only the top 25% of those, in order to calculate the TMAD measure. This results in a measure that is quite similar to the MSE measure, with the notable exception that it uses the same units as the MAD measure (absolute deviations rather than squared deviations), and is thus easier to interpret and the fact that it provides a clearer indication of the maximum deviation. As is the case with both MAD and MSE measures, lower TMAD scores represent more accurate estimates.

In addition to calculating the above measures, the analysis also included statistical testing to determine whether the pair-wise differences in those measures between

approaches are statistically significant, for all combinations<sup>7</sup>. Both standard pair-wise Student's t-tests (assuming unequal variance) and the signed-rank tests were used for this purpose.

### 3 Results

Tables 1-8 provide a summary of the three main accuracy measures for all of the assessed approaches, as well as the relative accuracy rankings of each approach, taking into account the results of the statistical tests for the difference in mean accuracy estimates<sup>8</sup>.

[insert tables 1-4 here]

#### 3.1 S1 simulation experiments

In general, the analysis found that the most accurate approaches in the simulation experiments that adopted a Cobb-Douglas production function are the parametric approaches, ie COLS when measurement error was not included in the analysis and SFA when measurement error was present (with one exception discussed below). This is not an unexpected result, since the parametric models that are ranked highest in each experiment are perfectly specified, in that they utilise the same functional form as the adopted production function and, in the case of the best-performing SFA models, assume the correct distribution for the inefficiency term.

Regarding the two non-parametric, determinist approaches, the overall performance of GA was perhaps surprisingly robust, even if the approach displayed the worst (or joint worst accuracy) in the majority of the experiments. In most cases, the difference in accuracy scores between GA and DEA was quite small and for the experiments that included measurement error, the difference was statistically insignificant. The analysis however identified some conditions where the accuracy of the GA quickly deteriorates:

- As technical inefficiency becomes more prevalent in the data that include no measurement error, the accuracy of the GA estimates rapidly deteriorates. In the S1.2 experiment, GA was ranked last, while both COLS and DEA were assessed to be substantially more accurate.

---

<sup>7</sup> For example, the average MAD score of the DEA estimates over all simulation runs in a single experiment is tested against the average MAD score of the GA, COLS and SFA (where applicable) estimates.

<sup>8</sup> In order to put the various MAD and TMAD measures into context, note that the data generation process adopted, both for the Cobb-Douglas and the piecewise-linear function, results in an average true productivity change of 2% p.a. but with a standard deviation of approximately 20%.

- When volatility in inputs and input prices increases from one period to the next, the accuracy of the GA deteriorates at a faster rate than the other approaches. This is the case in almost all of the S1 experiments (the exception is experiment S1.4, which includes both relatively high technical inefficiency and measurement error levels, where the overall accuracy of all approaches considered does not change when volatility is increased).

As for the performance of DEA, the analysis raises three major points:

- DEA is the most accurate approach based on the MAD measure when technical inefficiency is found at relatively high levels in the data that also do not include any measurement error. This is a rather surprising result, since as was mentioned above, the COLS model that is also assessed in the relevant experiment (S1.2) is perfectly specified, given that the S1.2 data are constructed using a Cobb-Douglas functional form and contain no measurement error. And indeed, the COLS approach is more accurate than DEA in this experiment based on the TMAD measure and equally accurate based on the MSE measure, which suggests that the performance of COLS improves for the units that occupy outlying positions in the dataset.
- The accuracy of the DEA-based estimates decreases at a lower pace relative to the accuracy of the other deterministic approaches when inputs and input prices become more volatile from one period to the next, in the experiments that do not include any measurement error (ie S1.1 and S1.2).

In addition to the points made above, some more general comments can be made when considering the analysis as a whole:

- When technical inefficiency is modest, there is no measurement error and the input levels and prices between subsequent periods are relatively stable (S1.1 experiment), all approaches provide quite accurate estimates of true productivity change.
- Increased volatility in inputs and prices in subsequent periods adversely affects accuracy of all approaches, when no measurement error is included in the constructed dataset. The DEA estimates are the least affected, while the GA estimates are the most affected. Interestingly, when measurement error is introduced in the analysis, the increased volatility appears to have very little

- When measurement error is present, the SFA approaches provide the most accurate estimates. However, when measurement error is more severe, even the best performing SFA model demonstrates quite large deviations from the true productivity change values (MAD scores of approximately 12.5%). In addition, when measurement error is more moderate, the gains in accuracy achieved by the SFA models are quite modest compared to the deterministic approaches (eg GA and DEA MAD scores are 5.8%, while the best performing SFA model has a MAD score of 5% in S1.3).

**Table 5: Summary accuracy results for the S1 experiments, default volatility**

|   | Measure | GA     | COLS   | DEA    | SFA (exponential) | SFA (half-normal) |
|---|---------|--------|--------|--------|-------------------|-------------------|
| S1.1: CRS CD, 12% average inefficiency, no noise        | MAD     | 0.90%  | 0.40%  | 0.70%  |                   |                   |
|   | MSE     | 0.21   | 0.04   | 0.2    |                   |                   |
|   | TMAD    | 4.10%  | 1.30%  | 4.10%  |                   |                   |
| S1.2: CRS CD, 32% average inefficiency, no noise        | MAD     | 2.80%  | 1.70%  | 1.20%  |                   |                   |
|   | MSE     | 1.57   | 0.62   | 0.46   |                   |                   |
|   | TMAD    | 10.10% | 4.80%  | 5.90%  |                   |                   |
| S1.3: CRS CD, 12% average inefficiency, noise~N(0,0.2)  | MAD     | 22.50% | 22.40% | 22.50% | 12.50%            | 13.70%            |
|   | MSE     | 79.52  | 78.87  | 79.08  | 27.8              | 32.13             |
|   | TMAD    | 62.00% | 61.70% | 61.40% | 40.00%            | 41.30%            |
| S1.4: CRS CD, 12% average inefficiency, noise~N(0,0.05) | MAD     | 5.80%  | 5.70%  | 5.80%  | 5.00%             | 5.40%             |
|   | MSE     | 5.3    | 5.06   | 5.34   | 4.04              | 4.64              |
|   | TMAD    | 16.30% | 15.90% | 16.20% | 14.20%            | 15.10%            |

**Table 6: Accuracy rankings for the S1 experiments, default volatility<sup>9</sup>**

|   | Measure | GA | COLS | DEA | SFA (exponential) | SFA (half-normal) |
|---|---------|----|------|-----|-------------------|-------------------|
| S1.1: CRS CD, 12% average inefficiency, no noise        | MAD     | 3  | 1    | 2   |                   |                   |
|   | MSE     | 3  | 1    | 2   |                   |                   |
|   | TMAD    | 2  | 1    | 2   |                   |                   |
| S1.2: CRS CD, 32% average inefficiency, no noise        | MAD     | 3  | 2    | 1   |                   |                   |
|   | MSE     | 3  | 1    | 1   |                   |                   |
|   | TMAD    | 3  | 1    | 2   |                   |                   |
| S1.3: CRS CD, 12% average inefficiency, noise~N(0,0.2)  | MAD     | 4  | 4    | 4   | 1                 | 2                 |
|   | MSE     | 4  | 4    | 4   | 1                 | 2                 |
|   | TMAD    | 4  | 4    | 4   | 1                 | 2                 |
| S1.4: CRS CD, 12% average inefficiency, noise~N(0,0.05) | MAD     | 4  | 3    | 4   | 1                 | 2                 |
|   | MSE     | 4  | 3    | 4   | 1                 | 2                 |
|   | TMAD    | 4  | 3    | 4   | 1                 | 2                 |

<sup>9</sup> The rankings take into consideration the results of the statistical tests for the difference in mean accuracy scores

**Table 7: Summary accuracy results for the S1 experiments, increased volatility**

|   | Measure | GA     | COLS   | DEA    | SFA (exponential) | SFA (half-normal) |
|---|---------|--------|--------|--------|-------------------|-------------------|
| S1.1: CRS CD, 12% average inefficiency, no noise        | MAD     | 2.50%  | 0.80%  | 1.20%  |                   |                   |
|   | MSE     | 1.7    | 0.16   | 0.5    |                   |                   |
|   | TMAD    | 11.80% | 2.70%  | 6.10%  |                   |                   |
| S1.2: CRS CD, 32% average inefficiency, no noise        | MAD     | 7.60%  | 3.00%  | 2.20%  |                   |                   |
|   | MSE     | 12.92  | 1.92   | 1.99   |                   |                   |
|   | TMAD    | 30.60% | 9.30%  | 12.70% |                   |                   |
| S1.3: CRS CD, 12% average inefficiency, noise~N(0,0.2)  | MAD     | 22.90% | 22.50% | 23.10% | 12.50%            | 13.90%            |
|   | MSE     | 82.76  | 79.75  | 83.16  | 28.11             | 33.43             |
|   | TMAD    | 63.90% | 62.50% | 63.20% | 40.00%            | 42.00%            |
| S1.4: CRS CD, 12% average inefficiency, noise~N(0,0.05) | MAD     | 6.30%  | 5.80%  | 6.00%  | 5.00%             | 5.40%             |
|   | MSE     | 6.53   | 5.17   | 5.66   | 3.96              | 4.67              |
|   | TMAD    | 18.5%  | 15.9%  | 16.7%  | 14.1%             | 15.2%             |

**Table 6: Accuracy rankings for the S1 experiments, increased volatility<sup>10</sup>**

|   | Measure | GA | COLS | DEA | SFA (exponential) | SFA (half-normal) |
|---|---------|----|------|-----|-------------------|-------------------|
| S1.1: CRS CD, 12% average inefficiency, no noise        | MAD     | 3  | 1    | 2   |                   |                   |
|   | MSE     | 3  | 1    | 2   |                   |                   |
|   | TMAD    | 3  | 1    | 2   |                   |                   |
| S1.2: CRS CD, 32% average inefficiency, no noise        | MAD     | 3  | 2    | 1   |                   |                   |
|   | MSE     | 3  | 1    | 2   |                   |                   |
|   | TMAD    | 3  | 1    | 2   |                   |                   |
| S1.3: CRS CD, 12% average inefficiency, noise~N(0,0.2)  | MAD     | 4  | 4    | 4   | 1                 | 2                 |
|   | MSE     | 4  | 4    | 4   | 1                 | 2                 |
|   | TMAD    | 4  | 4    | 4   | 1                 | 2                 |
| S1.4: CRS CD, 12% average inefficiency, noise~N(0,0.05) | MAD     | 5  | 3    | 4   | 1                 | 2                 |
|   | MSE     | 5  | 3    | 4   | 1                 | 2                 |
|   | TMAD    | 5  | 3    | 4   | 1                 | 2                 |

<sup>10</sup> The rankings also take into consideration the results of the statistical tests for the difference in mean accuracy scores



[insert tables 5-8 here]

### **3.2 Summary results for the S2 simulation experiments**

In the S2 experiments, it is the non-parametric approaches that are generally assessed as more accurate, with the exception of the simulations that assume 'extensive' noise. This was not unexpected, given that the underlying production function (piecewise linear) is not a perfect match to the functional form adopted by the parametric approaches. The experiments however demonstrate that the effect of this functional form misspecification can be quite severe. For example, when no measurement error is present, the COLS Cobb-Douglas specification displays MAD scores that are at least twice as large as those displayed by the DEA estimates and the discrepancy in MSE scores is significantly larger (at least three times higher).

Furthermore, the overall accuracy of the COLS specification that adopts a translog functional form is even worse; in almost all experiments, the COLS translog specification was ranked last in terms of overall accuracy. The SFA translog models however do perform relatively better, compared to their COLS counterparts, especially in the simulations that assume 'modest' noise levels. In the 'extensive' noise simulations do the Cobb-Douglas SFA models perform relatively better, but the difference in MAD scores is quite small (albeit statistically significant), relative to the overall accuracy of the estimates.

Another important issue revealed by the S2 simulations is the relative underperformance of the SFA models under conditions of 'modest' noise; in both such experiments (ie S2.3 and S2.6), GA and DEA, both deterministic approaches, perform better than the stochastic models. In addition, the Cobb-Douglas SFA models that (incorrectly) assume that the inefficiency is half-normally distributed is more accurate than the correctly specified Cobb-Douglas SFA exponential model, again under conditions of 'modest' noise. Only when measurement error is more severe (is the standard deviation is increased from 0.05 to 0.20), is the correctly specified SFA (exponential) model deemed to be more accurate. It should be noted however that the overall accuracy of all approaches examined under these conditions (ie 'extensive' noise) is relatively low (the better performing SFA models display MAD scores of 14.6% and 22.7% in experiments S2.4 and S2.5 respectively). This was also the case for the S1 experiments and suggests that additional research would be required to identify approaches that can produce robust estimates under these seemingly adverse conditions.

It should also be mentioned that the analysis encountered some difficulties when estimating the SFA models using datasets that included relatively large technical inefficiency levels. The problem was that the skew of the residual of the affected models was wrong, which meant that the SFA estimation could not proceed. Although the analysis circumvented this issue by discarding the problematic datasets, this would not be a possible solution in a productivity measurement exercise utilising real data.

Overall, the S2 experiments revealed that GA and DEA provide reasonably accurate estimates under various conditions. In the case of GA, this is a somewhat surprising result, given that the approach does not acknowledge the presence of technical inefficiency, which is a not inconsiderable component of productivity change in these experiments. However, DEA is revealed to be the more accurate approach of the two, and, one could argue, the more accurate approach overall, in the S2 experiments.

Another advantage of the DEA-derived estimates is their apparent robustness under conditions of increased volatility in inputs and prices. The S2 simulations showed that increased volatility reduces the accuracy of all estimates, but to a different degree for each approach; the same experiments also showed that the DEA-based estimates are in the majority of cases the ones that are least affected.

**Table 5: Summary accuracy results for the S2 experiments, default volatility**

|   | Measure     | GA    | COLS  | COLS (translog) | DEA   | SFA   | SFA (translog) | SFA (half-normal) |
|---|-------------|-------|-------|-----------------|-------|-------|----------------|-------------------|
| <b>S2.1: CRS PWL, 12% average inefficiency, no noise</b>        | <b>MAD</b>  | 0.9%  | 2.4%  | 2.9%            | 0.8%  |       |                |                   |
|   | <b>MSE</b>  | 0.20  | 1.08  | 70.31           | 0.22  |       |                |                   |
|   | <b>TMAD</b> | 4.1%  | 8.2%  | 21.6%           | 4.1%  |       |                |                   |
| <b>S2.2: CRS PWL, 32% average inefficiency, no noise</b>        | <b>MAD</b>  | 2.2%  | 2.4%  | 4.9%            | 1.1%  |       |                |                   |
|   | <b>MSE</b>  | 1.09  | 1.20  | 132.12          | 0.39  |       |                |                   |
|   | <b>TMAD</b> | 8.8%  | 8.8%  | 32.1%           | 5.4%  |       |                |                   |
| <b>S2.3: CRS PWL, 12% average inefficiency, noise~N(0,0.05)</b> | <b>MAD</b>  | 5.8%  | 6.3%  | 6.5%            | 5.8%  | 7.1%  | 6.1%           | 6.4%              |
|   | <b>MSE</b>  | 5.33  | 6.24  | 7.81            | 5.23  | 9.11  | 6.12           | 6.96              |
|   | <b>TMAD</b> | 16.2% | 17.6% | 20.6%           | 16.0% | 21.6% | 17.7%          | 18.5%             |
| <b>S2.4: CRS PWL, 12% average inefficiency, noise~N(0,0.2)</b>  | <b>MAD</b>  | 22.9% | 23.0% | 23.4%           | 22.8% | 14.3% | 15.4%          | 16.6%             |
|   | <b>MSE</b>  | 86.21 | 86.93 | 95.96           | 85.38 | 38.41 | 43.70          | 49.76             |
|   | <b>TMAD</b> | 67.7% | 67.8% | 71.3%           | 67.4% | 48.8% | 51.4%          | 52.2%             |
| <b>S2.5: CRS PWL, 32% average inefficiency, noise~N(0,0.2)</b>  | <b>MAD</b>  | 23.6% | 23.6% | 24.8%           | 23.4% | 21.5% | 22.2%          | 22.9%             |
|   | <b>MSE</b>  | 91.33 | 91.50 | 148.54          | 89.94 | 76.76 | 82.49          | 86.35             |
|   | <b>TMAD</b> | 69.1% | 69.5% | 82.0%           | 68.7% | 63.6% | 66.4%          | 67.4%             |
| <b>S2.6: CRS PWL, 32% average inefficiency, noise~N(0,0.05)</b> | <b>MAD</b>  | 6.2%  | 6.2%  | 8.7%            | 5.8%  | 7.1%  | 6.5%           | 6.2%              |
|   | <b>MSE</b>  | 6.09  | 6.12  | 287.07          | 5.33  | 8.56  | 6.77           | 6.16              |
|   | <b>TMAD</b> | 17.5% | 17.4% | 47.2%           | 16.3% | 20.4% | 18.6%          | 17.4%             |

**Table 6: Accuracy rankings for the S2 experiments, default volatility<sup>11</sup>**

|  | Measure | GA | COLS | COLS (translog) | DEA | SFA | SFA (translog) | SFA (half-normal) |
|--|---------|----|------|-----------------|-----|-----|----------------|-------------------|
| S2.1: CRS PWL, 12% average inefficiency, no noise        | MAD     | 1  | 3    | 4               | 1   |     |                |                   |
|  | MSE     | 1  | 3    | 4               | 1   |     |                |                   |
|  | TMAD    | 1  | 3    | 4               | 1   |     |                |                   |
| S2.2: CRS PWL, 32% average inefficiency, no noise        | MAD     | 2  | 3    | 4               | 1   |     |                |                   |
|  | MSE     | 2  | 3    | 4               | 1   |     |                |                   |
|  | TMAD    | 2  | 2    | 4               | 1   |     |                |                   |
| S2.3: CRS PWL, 12% average inefficiency, noise~N(0,0.05) | MAD     | 1  | 5    | 5               | 1   | 7   | 3              | 5                 |
|  | MSE     | 2  | 4    | 6               | 1   | 7   | 3              | 4                 |
|  | TMAD    | 1  | 4    | 6               | 1   | 7   | 4              | 4                 |
| S2.4: CRS PWL, 12% average inefficiency, noise~N(0,0.2)  | MAD     | 5  | 5    | 7               | 4   | 1   | 2              | 3                 |
|  | MSE     | 5  | 5    | 7               | 4   | 1   | 2              | 2                 |
|  | TMAD    | 5  | 5    | 7               | 5   | 1   | 2              | 2                 |
| S2.5: CRS PWL, 32% average inefficiency, noise~N(0,0.2)  | MAD     | 5  | 5    | 7               | 3   | 1   | 2              | 3                 |
|  | MSE     | 5  | 5    | 7               | 3   | 1   | 2              | 3                 |
|  | TMAD    | 5  | 5    | 7               | 3   | 1   | 2              | 3                 |
| S2.6: CRS PWL, 32% average inefficiency, noise~N(0,0.05) | MAD     | 3  | 3    | 7               | 1   | 6   | 5              | 3                 |
|  | MSE     | 3  | 3    | 7               | 1   | 6   | 5              | 3                 |
|  | TMAD    | 3  | 3    | 7               | 1   | 6   | 5              | 3                 |

**Table 7: Summary accuracy results for the S2 experiments, increased volatility**

|  | Measure | GA    | COLS  | COLS (translog) | DEA   | SFA   | SFA (translog) | SFA (half-normal) |
|--|---------|-------|-------|-----------------|-------|-------|----------------|-------------------|
| S2.1: CRS PWL, 12% average inefficiency, no noise        | MAD     | 2.5%  | 6.8%  | 9.0%            | 1.6%  |       |                |                   |
|  | MSE     | 1.77  | 8.64  | 20.57           | 0.96  |       |                |                   |
|  | TMAD    | 12.4% | 23.4% | 35.5%           | 7.1%  |       |                |                   |
| S2.2: CRS PWL, 32% average inefficiency, no noise        | MAD     | 5.7%  | 6.8%  | 17.4%           | 2.2%  |       |                |                   |
|  | MSE     | 8.22  | 9.00  | 73.56           | 1.62  |       |                |                   |
|  | TMAD    | 25.6% | 24.2% | 63.6%           | 10.8% |       |                |                   |
| S2.3: CRS PWL, 12% average inefficiency, noise~N(0,0.05) | MAD     | 6.3%  | 8.9%  | 11.4%           | 5.9%  | 8.9%  | 11.4%          | 8.6%              |
|  | MSE     | 6.54  | 13.40 | 27.97           | 5.47  | 14.59 | 29.43          | 13.08             |
|  | TMAD    | 18.7% | 27.6% | 38.8%           | 16.5% | 29.0% | 39.8%          | 27.6%             |
| S2.4: CRS PWL, 12% average inefficiency, noise~N(0,0.2)  | MAD     | 23.5% | 24.4% | 26.6%           | 23.3% | 14.7% | 19.3%          | 17.3%             |
|  | MSE     | 90.67 | 97.77 | 116.72          | 89.43 | 41.05 | 67.99          | 54.33             |
|  | TMAD    | 68.9% | 71.1% | 78.3%           | 68.4% | 50.5% | 61.7%          | 54.5%             |
| S2.5: CRS PWL, 32% average inefficiency, noise~N(0,0.2)  | MAD     | 23.9% | 23.9% | 31.8%           | 22.9% | 21.7% | 28.1%          | 23.3%             |
|  | MSE     | 94.84 | 94.30 | 177.16          | 87.49 | 78.90 | 140.91         | 89.70             |
|  | TMAD    | 72.1% | 71.4% | 95.4%           | 69.1% | 65.8% | 87.6%          | 70.0%             |
| S2.6: CRS PWL, 32% average inefficiency, noise~N(0,0.05) | MAD     | 8.4%  | 9.1%  | 19.1%           | 6.2%  | 9.8%  | 14.6%          | 9.0%              |
|  | MSE     | 13.00 | 13.72 | 81.15           | 6.56  | 16.34 | 49.48          | 13.71             |
|  | TMAD    | 28.5% | 27.7% | 66.1%           | 18.9% | 29.6% | 53.1%          | 28.2%             |

<sup>11</sup> The rankings take into consideration the results of the statistical tests for the difference in mean accuracy scores

**Table 6: Accuracy rankings for the S2 experiments, increased volatility<sup>12</sup>**

|   | Measure     | GA | COLS | COLS<br>(translog) | DEA | SFA | SFA<br>(translog) | SFA (half-<br>normal) |
|---|-------------|----|------|--------------------|-----|-----|-------------------|-----------------------|
| <b>S2.1: CRS PWL,<br/>12% average<br/>inefficiency, no<br/>noise</b>        | <b>MAD</b>  | 2  | 3    | 4                  | 1   |     |                   |                       |
|   | <b>MSE</b>  | 2  | 3    | 4                  | 1   |     |                   |                       |
|   | <b>TMAD</b> | 2  | 3    | 4                  | 1   |     |                   |                       |
| <b>S2.2: CRS PWL,<br/>32% average<br/>inefficiency, no<br/>noise</b>        | <b>MAD</b>  | 2  | 3    | 4                  | 1   |     |                   |                       |
|   | <b>MSE</b>  | 2  | 3    | 4                  | 1   |     |                   |                       |
|   | <b>TMAD</b> | 3  | 2    | 4                  | 1   |     |                   |                       |
| <b>S2.3: CRS PWL,<br/>12% average<br/>inefficiency,<br/>noise~N(0,0.05)</b> | <b>MAD</b>  | 2  | 4    | 6                  | 1   | 4   | 6                 | 4                     |
|   | <b>MSE</b>  | 2  | 4    | 6                  | 1   | 4   | 6                 | 4                     |
|   | <b>TMAD</b> | 2  | 4    | 6                  | 1   | 4   | 6                 | 4                     |
| <b>S2.4: CRS PWL,<br/>12% average<br/>inefficiency,<br/>noise~N(0,0.2)</b>  | <b>MAD</b>  | 5  | 6    | 7                  | 4   | 1   | 3                 | 2                     |
|   | <b>MSE</b>  | 5  | 6    | 7                  | 4   | 1   | 3                 | 2                     |
|   | <b>TMAD</b> | 4  | 6    | 7                  | 4   | 1   | 3                 | 2                     |
| <b>S2.5: CRS PWL,<br/>32% average<br/>inefficiency,<br/>noise~N(0,0.2)</b>  | <b>MAD</b>  | 4  | 4    | 7                  | 2   | 1   | 6                 | 4                     |
|   | <b>MSE</b>  | 4  | 4    | 7                  | 2   | 1   | 6                 | 4                     |
|   | <b>TMAD</b> | 4  | 4    | 7                  | 2   | 1   | 6                 | 4                     |
| <b>S2.6: CRS PWL,<br/>32% average<br/>inefficiency,<br/>noise~N(0,0.05)</b> | <b>MAD</b>  | 2  | 3    | 7                  | 1   | 5   | 6                 | 3                     |
|   | <b>MSE</b>  | 3  | 3    | 7                  | 1   | 5   | 6                 | 3                     |
|   | <b>TMAD</b> | 3  | 3    | 7                  | 1   | 5   | 6                 | 3                     |

<sup>12</sup> The rankings also take into consideration the results of the statistical tests for the difference in mean accuracy scores

## 4 Summary and recommendations

The major findings of this simulation study on the accuracy of the productivity change estimates are summarised below:

- Deterministic approaches perform adequately even under conditions of (modest) measurement error.
- When measurement error becomes larger, the accuracy of all approaches (including SFA) deteriorates rapidly, to the point that their estimates could be considered unreliable for policy purposes.
- Functional form misspecification has a severe negative impact on the accuracy of all parametric approaches.
- The SFA models that adopt a translog specification appear to be more accurate in general than the Cobb-Douglas SFA models when the underlying (true) production function is piecewise linear. The opposite is true for the COLS models, is the Cobb-Douglas COLS models are more accurate than their translog counterparts in such circumstances.
- Misspecification of the inefficiency distribution in the SFA models does not appear to have a significant effect on the overall accuracy of said approach.
- Increased volatility in inputs and prices from one period to the next adversely affects the accuracy of all approaches, in almost all experiments. The DEA estimates are the least affected, while the GA estimates are the most affected.

This analysis demonstrates that no productivity change measurement approach has an absolute advantage over another, but rather under some specific circumstances, a specific approach is likely to be more accurate than its counterparts. The analysis also demonstrates that frontier-based approaches can usually produce at least as accurate, and in most cases more accurate, productivity change estimates than the more traditional GA approach. And given that high quality databases on measures of economic growth, employment creation and capital formation are becoming increasingly available (EU KLEMS, 2008), the adoption of frontier-based approaches when measuring aggregate productivity growth can only help improve our understanding of this elusive and complex topic.

## References

- Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21-37.
- Balk, B. (2008). Measuring Productivity Change Without Neoclassical Assumptions: A Conceptual Analysis. In *ERIM Report Series Reference No. ERS-2008-077-MKT*.
- Berndt, E. R., & Fuss, M. A. (1986). Productivity measurement with adjustments for variations in capacity utilization and other forms of temporary equilibrium. *Journal of Econometrics*, 33, 7-29.
- Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica*, 50, 1393-1414.
- Del Gatto, M., Di Liberto, A., & Petraglia, C. (2008). Measuring Productivity. In: Centre for North South Economic Research, University of Cagliari and Sassari, Sardinia.
- Diewert, W. E. (1992). The Measurement of Productivity. *Bulletin of Economic Research*, 44, 163-198.
- EU KLEMS. (2008). EU KLEMS Database, see Marcel Timmer, Mary O'Mahony & Bart van Ark, The EU KLEMS Growth and Productivity Accounts: An Overview, University of Groningen & University of Birmingham. In.
- Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review*, 84, 66-83.
- Fried, H. O., Lovell, C. A. K., & Schmidt, S. S. (2008). The Measurement of Productive Efficiency and Productivity Growth. In: Oxford University Press.
- Jondrow, J., Knox Lovell, C. A., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19, 233-238.
- Koszerek, D., Havik, K., Mc Morrow, K., Röger, W., & Schönborn, F. . (2007). An overview of the EU KLEMS growth and productivity accounts. In *Economic paper no. 290*: European Commission.

- Kumbhakar, S. C., & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*: Cambridge University Press.
- Meeusen, W., & van Den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review*, 18, 435-444.
- O'Mahony, M., & Timmer, M., P. (2009). Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database. *Economic Journal*, 119, 374-403.
- OECD. (2001). Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth. In: OECD.
- ONS. (2007). The ONS productivity Handbook: A Statistical Overview and Guide In D. Camus (Ed.): Palgrave Macmillan.
- Pastor, J. T., & Lovell, C. A. K. (2005). A global Malmquist productivity index. *Economics Letters*, 88, 266-271.
- Portela, M. C. A. S., & Thanassoulis, E. (2010). Malmquist-type indices in the presence of negative data: An application to bank branches. *Journal of Banking & Finance*, 34, 1472-1483.
- Samuelson, P. A. (1947). *Foundations of Economic Analysis*: Harvard University Press.
- Sharma, S. C., Sylwester, K., & Margono, H. (2007). Decomposition of total factor productivity growth in U.S. states. *The Quarterly Review of Economics and Finance*, 47, 215-241.
- Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics*, 39, 312-320.
- Thanassoulis, E. (2001). *Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software*: Springer.
- Tinbergen, J. (1942). Theorie der Langfristigen Wirtschaftsentwicklung, (On the Theory of Long-Term Economic Growth). *Weltwirtschaftliches Archiv*, 55, 511-549.
- Tulkens, H., & Vanden Eeckaut, P. (1995). Non-parametric efficiency, progress and regress measures for panel data: Methodological aspects. *European Journal of Operational Research*, 80, 474-499.