Are EU Candidate Countries Ready for the Single Market? Comparing Performance with an EU Benchmark.

Jan De Loecker*
Department of Economics, K.U.Leuven
LICOS Centre for Transition Economics, K.U.Leuven

May 12, 2003

Abstract

The process of the East-West integration has come to a point that enlargement appears inevitable. But are these countries really ready to join the EU in terms of competitive performance? Firstly, we construct a virtual best practice production frontier for firms active in a European Union member state (EU) and firms in Central and East European (CEE) countries. By means of this common benchmark, we make inferences about the relative technical efficiency of firms in candidate countries. We show that technical efficiency is higher in the EU, however, firms in CEE countries are converging towards the EU average. We show that efficiency is not only different between countries, but even within countries. The latter suggests the need to look at industry-level technical efficiency. Secondly, we combine two existing techniques to develop a methodology to estimate firm level total factor productivity (TFP) change, decomposed into technical change, technical efficiency change and a scale component. It is mainly technical change that drives the change in TFP, i.e. technological progress. Firms in EU candidate countries are reallocating resources under decreasing returns to scale, leading to an increase in TFP. We suggest interpreting efficiency estimates as measures of firm heterogeneity.

Keywords: EU enlargement, stochastic frontier, technical efficiency, total factor productivity.

JEL: P2, P5, L6, C33

---

*This paper has benefited from participants at the International Industrial Organisation Conference (2003, North Eastern University, Boston) and LICOS Seminar (2002, K.U.Leuven). I would like to thank Joep Konings, Ana Xavier, Greetje Everaert and Jo Swinnen for valuable comments. All errors are mine. To contact author: Jan De Loecker, LICOS-Centre for Transition Economics, K.U.Leuven, Deberiotstraat 34, 3000 Leuven, Belgium, tel +32 16 32 65 86, fax +32 16 32 65 99, e-mail: jan.deloecker@econ.kuleuven.ac.be.

1 Introduction

The enlargement of the European Union (EU) towards Central- and Eastern Europe is an ongoing debate, with far reaching implications for both EU members and EU candidates. The process of the East-West integration has come to a point that enlargement appears inevitable. But are these countries really ready to join the EU? One of the criteria set in the Copenhagen summit of 1993 is that a country has to be able to meet the competition in the single market. Whether they can or not is of course of great interest for policy makers and business men in both the EU and in the candidate countries. According to the European Commission (EC) the existence of a functioning market economy requires an economic policy that improves efficiency of the economy. This is evaluated in the EC's annual report on the progress of the EU candidates and for e.g. Poland the EC (2001) states: "Poland is a functioning market economy. Provided that it continues and intensifies its present reform efforts in a consistent policy environment, it should be able to cope with the competitive pressure and market forces within the Union, in the near term.". The latter is based on an overall evaluation of Poland's economic situation, however, we check whether efficiency at the firm level gives additional information to discuss these statements.

Comparing countries from both sides of the iron curtain is one of the most important methodological challenges related to the East-West integration (Osiewalski, Koop and Steel, 1997). In this paper we construct a virtual best practice production frontier for firms active in an European Union member state (EU) and Central and East European (CEE) countries. This is the first paper that we know of that creates such a common benchmark for both West and East European countries using firm level data. By means of this common benchmark, we can make inferences about the relative technical efficiency of firms in candidate countries, both at the sector and the subsectoral level. We use this framework as a workhorse model to estimate firm level total factor productivity (TFP) change, decomposed into technical change, technical efficiency change and a scale component. To this end we extend the approach of Battese and Coelli (1995), who estimate firm- level efficiency and its determinants simultaneously, and allow for a firm and time specific decomposition of TFP change using the framework of Kumbhakar and Lovell (2000). This methodology is used to estimate a firm-specific decomposition of TFP change.

We focus on the manufacturing sector of two EU candidates (Poland and the Czech Republic) and Belgium. Belgium is a small open economy and therefore it is subject to competitive pressure from international markets and we expect it to set the benchmark.[1] This permits us to compare the relative efficiency of Polish and Czech firms to an EU-benchmark (created by the Belgian firms). If these countries

are to be ready to join the EU and let their firms compete freely within the single market, efficiency differences between these countries and the EU should not be too big or at least become smaller. Of course the words not too big are a relative concept that is to be treated with care. However, as stated in Konings and Repkine (1994): "a base line estimate of efficiency would be useful starting point to be able to judge whether economic reforms can indeed achieve higher efficiency". Geared towards the EU enlargement, this can serve as an indicator to compare the potential new member states from Central and Eastern Europe with the current EU member states.

Productivity efficiency has two components: i) the purely technical or physical component that refers to using as little input as output production allows, ii) the allocative or price component that refers to the ability to combine inputs and outputs in optimal proportions in light of prevailing prices.[2] In this paper we look at firm performance measured by technical efficiency and we also check the share of it in firm TFP growth.

In some cases economic theory does not provide unambiguous guidance explaining efficiency differences between firms and concerning the impact of some forces on market performance. In those cases empirical analysis can provide both qualitative and quantitative evidence (Fried, Lovell and Schmidt, 1993). This appears to be the case in this analysis, i.e. is a firm in a CEE country less efficient or productive than an EU firm? If the question is whether producers in EU-firms outperform those in CEE-firms, ordinary least squares with dummies will do. However, these dummies categorize producers prior to estimation. In this case frontier techniques can be more informative.

Measurement of efficiency that is based upon a production frontier can essentially be done in two ways. The stochastic frontier approach implies creating a (stochastic) best-practice frontier based on an underlying production function and allows us to distinguish the effects of noise from the effects of inefficiency. This means that a fraction of the distance between the benchmark and a point under this benchmark results in a measure of technical efficiency (TE).[3] The programming approach, like D.E.A., is deterministic in nature and does not distinguish between noise and inefficiency effects, i.e. every deviation from the benchmark is attributed to technical inefficiency. This advantage of the stochastic frontier approach comes at a cost: the econometric approach is subject to specification errors (e.g. functional form). However, the size of the data used can reduce the potential problem substantially.

The theoretical framework is presented in the second section. We first discuss cross sectional models and we extend these models to allow for panel data. Furthermore, we discuss a methodology to estimate firm level TFP change decomposition. The third section deals with the data and issues of pooling several countries into one

data set. The estimated technical efficiencies are discussed in section 4 and we check whether technical efficiency changes over time. We also perform a sectoral analysis of technical efficiency within and over the three countries by estimating technical efficiency (TE) at a subsectoral level. We then estimate TFP change into its various components using firm level data. The penultimate section gives an interpretation to these estimated technical efficiencies and the last section concludes.

2 Theoretical Framework

In this section we give a brief overview of the stochastic frontier theory that started only 25 years ago. We first introduce the cross sectional model and then we look at different models that allow for the use of panel data. We discuss both the Battese and Coelli (1992) and (1995) specification more in detail since these are the core specifications we use in the empirical application.[4] We briefly comment on the different types of production functions and conclude this section with a method to estimate a firm level TFP change decomposition into technical change, technical efficiency change and a scale component. To this we extend the approach of Battese and Coelli (1995), where firm level efficiency and its determinants are estimated simultaneously, to allow for a firm and time specific decomposition of TFP change using the framework of Kumbhakar and Lovell (2000).

2.1 A Cross Sectional Model for Technical Efficiency

The 'stochastic frontier' production function was first proposed in 1977 independently by Aigner, Lovell and Schmidt (1977) and by Meeusen and van den Broek (1977). They came up with an original specification for a cross sectional model. The novelty of these models was that they constructed an error term consisting of two components: one that accounts for random effects and the other that accounts for technical inefficiency. This model can be represented as follows

$$y_i = f(x_i; \beta) \exp(v_i - u_i) \qquad (1)$$

with i the firm index, y the vector of outputs and $x = (x_1, \ldots, x_N)$ the vector of inputs. The vector of parameters is denoted by β. The error term consists of random variables v_i assumed to be $i.i.d$ with a zero mean and variance σ_v^2 and the elements in vector u_i which are crucial in this model. These non-negative random variables account for technical inefficiency in production and are often assumed to be $i.i.d$ or half-normal with mean μ and variance σ_u^2. A common criticism of stochastic frontier models is that there is no a priori justification for the selection of any particular distributional form for the technical inefficiency u_i. The common assumptions of half-normal or exponential distributions for the u are just arbitrary selections that

imply a high probability that technical inefficiencies are in the neighborhood of zero. This may lead to very high technical efficiency estimates.

A reaction to this critique led to different specifications like the truncated-normal distribution (Stevenson, 1980) and the 2-parameters gamma distribution (Green, 1990) for the u_is. The truncated-normal distribution is a generalisation of the half-normal distribution. It is obtained by truncation at zero of the normal distribution with mean μ and variance σ^2. It is easily seen that restricting μ to zero results in a half-normal distribution. The distribution depends strongly on the sign and size of μ. We can test whether this mean is significantly different from zero by means of a Wald or generalized likelihood-ratio test. The test-statistic has a slightly different distribution under the null hypothesis of no technical inefficiency in the model and thus other critical values are appropriate (Kodde and Palm, 1986).

2.2 Panel Data models

The previous discussion assumes that cross-sectional data on N firms are available for the estimation of the parameters of the stochastic frontier. With the growing availability of panel data, both on the micro and the macro level, techniques have been developed to incorporate panel data within the stochastic frontier framework. The firm-specific technical inefficiency could be captured by the firm-specific effect within the fixed effects framework. Strong distributional assumptions that are necessary in a cross-sectional setting were replaced by the single - and very strong - assumption that technical inefficiency is time-invariant. The last years, the literature developed techniques to allow for time-variant technical efficiency measures. The crucial question for this research was to relax the assumptions of time invariant technical efficiency measures without losing the advantages of panel data.

Panel data have some advantages over cross-sectional data in the estimation of the stochastic frontier. A larger number of degrees of freedom for the estimation of parameters is available, and more importantly simultaneous investigation of both technical change and technical efficiency change over time is possible.[5]

Pit and Lee (1981) specified a panel data version of the Aigner, Lovell and Schmidt (1977) half-normal model and is given by the following expression

$$\log y_{i;t} = (\log X_{i;t})'\beta + v_{i;t} - u_{i;t} \quad (2)$$

with $i = 1;\ldots;N$ and $t = 1;\ldots;T$ and same definitions as before.[6] The $v_{i;t}$s are random errors and assumed to be $i.i.d.$ and have a $N(0;\sigma_v^2)$ distribution and are distributed independently of the technical inefficiency effect u.[7] The latter were first assumed to be $i.i.d.$, but this resulted in no particular advantages in obtaining additional observations on a given firm versus obtaining observations on more firms

at particular time periods. Thus, the second basic model assumed $u_{i,t} = u_i$, i.e. time-invariant technical inefficiencies.

Battese and Coelli (1988) extended this model and assumed u_i to be truncated-normally distributed and later on permitted unbalanced panels (Battese, Coelli and Colby, 1989). The assumption of time-invariant technical inefficiency becomes, however, more and more difficult to justify as T becomes larger. As Battese and Coelli (1998) stated "one would expect that managers learn from their previous experience in the production process and so their technical inefficiency effects would change in some persistent pattern over time".

Kumbhakar (1990) suggested a stochastic frontier model for panel data in which technical inefficiency effects vary systematically with time according to the time-varying specification $u_{i,t} = \left[1 + \exp(bt + ct^2)\right]^{-1} u_i$, where b and c are to be estimated. However, no empirical application has yet been attempted. Battese and Coelli (1992) suggest an alternative to this approach in which the $u_{i,t}$ are assumed to be an exponential function of time involving only one unknown parameter.

Schmidt and Sickles (1984) observed that when panel data are available there is no need to specify a particular distribution for the technical inefficiency effects, because the parameters of the model can be estimated by means of standard panel data techniques (fixed effects) or error-components estimation (see also Hill and Jugde, 1993). In the fixed effects approach, the largest estimated firm intercept is used to estimate the intercept parameter so that all firm effects are estimated to be zero or negative. An estimate of the technical inefficiency of the various firms are obtained relative to the most efficient firm(s). An advantage of traditional panel data is that fixed effects approach are preferred when one assumes that the technical inefficiency effects $u_{i,t}$ are correlated with the production inputs $X_{i,t}$. Yet, this approach implies that efficient and inefficient firms have the same influence on the frontier. Cornwell, Schmidt and Sickles (1990) and Lee and Schmidt (1993) came up with generalizations of Schmidt and Sickles (1984).[8] The advantage of the Schmidt and Lee (1993) approach is that it allows for inclusion of observables that are time-invariant or invariant over firms. The Schmidt and Lee (1993) model is represented by $y_{i,t} = X'_{i,t}\beta + \theta_t\delta_i + v_{i,t}$. The $v_{i,t}$ are assumed to be independently and identically distributed with mean zero and variance σ^2_v (note that we do not have to assume normality). This type of estimation of technical efficiency is essentially based on weighted averages of residuals (with the average taken over time for a given firm). In the presence of a small time period T these could be very noisy. This may lead to overestimation of the differences between the firms' intercepts, and therefore in the underestimation of the technical efficiency levels.

A number of authors allowed for time-variant efficiencies, however in rather structured ways. The Battese and Coelli (1992) approach is an illustration of such a

structured model whereas Schmidt and Lee (1993) brought more flexibility[9] in the way these change over time. Battese and Coelli (1992) propose a time varying model for the technical inefficiency effects in the stochastic frontier production function for panel data. The technical inefficiency effects are assumed to be defined by

$$u_{i,t} = [\exp(-\eta(t - T))]u_i \qquad (3)$$

where the u_is are the random variable defined as before and assumed to be $i.i.d$ according to the generalized truncated-normal and η is an unknown parameter to be estimated. If the i-th firm is observed in the last period of the panel (T) then $u_{i,T} = u_i$ because the exponential function in (3) has value one when $t = T$. Thus the random variable can be considered as the technical inefficiency effect for firm i in the last period of the panel. For earlier periods in the panel, the technical inefficiency effects are the product of the technical inefficiency at the last period and the value of the exponential function $[\exp(-\eta(t - T))]$, whose value depends on the parameter η and the number of periods before the last period of the panel ($-(t - T) \equiv T - t$). If this parameter is positive then the value of the exponential function is no less than one, which implies that $u_i \leq u_{i,t}$. The converse is true if the parameter that accounts for time-variance of technical inefficiency - η - is negative.[10] This approach implies that the ordering of the firms according to the magnitude of the technical inefficiency effects is the same at all time periods.[11] The Battese and Coelli (1992) specification does not account for situations in which some firms may be relatively inefficient but become relatively more efficient in subsequent periods. The Cornwell, Schmidt and Sickles (1990) and Lee and Schmidt (1993) specifications do allow for this possibility.[12]

One of the advantages of the B&C (1992) time varying inefficiency model is that the change of technical inefficiency over time can be distinguished from technical change, provided the latter is appropriately specified in the frontier function (see section 2.5). This discrimination of technical change and technical inefficiency change over time is only possible given that the technical inefficiency effects are stochastic and have the specified distributions. The next section discusses an extension of the Battese and Coelli (1992) specification, in the sense that they develop a model that allows to estimate the efficiencies and its determinants simultaneously.

2.3 Modelling Determinants of Inefficiency Effects

In this section we briefly discuss the BC(1995) specification[13] that allows for the simultaneous estimation of technical inefficiency and its determinants. A number of empirical studies (e.g. Pitt and Lee, 1981) have estimated the determinants of technical inefficiencies among firms. In a second stage analysis the estimated firm-specific technical inefficiencies are regressed on a number of characteristics, like firm

size, age and others. This two stage approach has a major drawback. In a first stage, the technical inefficiencies are assumed to be i.i.d in order to estimate them. However, in the second stage, the predicted inefficiency effects are assumed to be a function of a number of firm specific factors, which implies that they are not i.i.d unless all the coefficients of the factors are simultaneously equal to zero (Battese and Coelli, 1998). Battese and Coelli (1995) extend the approaches of Kumbhakar, Ghosh and McGuckin (1991) and Stevenson (1991) to accommodate for panel data. This allows the estimation of the parameters of the factors believed to determine the levels of technical inefficiency together with the separate components of technical inefficiency change and technical change over time. The error component accounting for technical inefficiency $u_{i,t}$ is obtained by truncation of the $N(\mu_{i,t}, \sigma^2)$ distribution where

$$\mu_{i,t} = z_{i,t}\delta \tag{4}$$

where $z_{i,t}$ is a vector of explanatory variables whose values are fixed constants and δ is a vector of unknown parameters to be estimated. Equation (4) specifies that the means of the normal distributions (which are truncated at zero to obtain the distributions of the technical inefficiency effects) are not the same but are functions of values of observable variables and of a common vector of parameters. If all δ-parameters are zero then the model is efficiently estimated by OLS and if only the constant term is different form zero we have the Aigner, Schmidt and Lovell (1977) model. The technical efficiency for the i-th firm in the t-period is now defined by $TE_{i,t} = \exp(-u_{i,t})$.

2.4 The Underlying Production Function

The model described above uses a production function $f(x;\beta)$, which can take various forms. The most known and simple is the Cobb-Douglas production function. This form has the advantage that it is easy to interpret. Its log linear transformation also immediately reveals the output elasticity's of the different inputs. The Cobb-Douglas production in its log linear form can be represented as

$$\log Y_{i,t} = \log A_{i,t} + \beta_1 \log K_{i,t} + \beta_2 \log L_{i,t} \tag{5}$$

and this representation will be used later on. Output, capital and labour of firm i at time t are denoted by $Y_{i,t}$, $K_{i,t}$ and $L_{i,t}$ respectively. The Cobb-Douglas production function, however, is a very restrictive function in that it assumes a constant elasticity of substitution between the different inputs. Therefore we need a production function that has less restrictions. A more 'unrestrictive form' is the translog production function, as it allows for non-linear impacts of the inputs and for an unrestricted elasticity of substitution and is represented in equation (6).

$$\log Y_{i,t} = \log A_{i,t} + \beta_1 \log K_{i,t} + \beta_2 \log L_{i,t} + \beta_3 (\log K_{i,t})^2 + \beta_4 (\log L_{i,t})^2$$

$$+\beta_5 \log K_{i,t} \log L_{i,t} \quad (6)$$

The output elasticity's are the result of simply deriving the translog with respect to capital and labour. The different functional forms will be explored when estimating the firm-level technical efficiencies.

2.5 Decomposition of TFP within a Stochastic Frontier Model[14]

In this section we discuss a potential methodology to disentangle total factor productivity change into its various components at the firm level using the framework suggested by Lovell and Kumbhakar (2000). To keep notation as general as possible, we assume the following representation of a stochastic frontier production function without explicitly denoting a firm index and a time index

$$y = f(x, t, \beta).\exp(v - u) \quad (7)$$

with $x = (x_1, ..., x_N)$ the input vector, β the vector of technological parameters and t a time trend. As before u captures the output oriented technical efficiency and $u \geq 0$. We want to decompose changes in firm productivity into three different components. The first is technical change (TM) which captures a shift upwards or downwards in the production frontier. The second component refers to a change in technical efficiency (TEM). This is the rate at which firms move towards or away from the frontier. A third component captures the scale elasticity of the firm. The difficulty occurs (and this is not unrealistic) when these three components change simultaneously. We define productivity[15] change as the difference between the rate of change of output and the rate of change of an input index. Total factor productivity (TFP) can formally be represented as

$$\dot{TFP} = \dot{y} - \dot{X} \quad (8)$$

where a dot above a variable indicates its rate of change ($\dot{y} = d\ln y/dt$). The change in the input is defined by $\dot{X} = \sum_n S_n \dot{x}_n$ with $S_n = \frac{w_n x_n}{E}$ and $E = \sum_n w_n x_n$. So the change in input is the sum of all different individual changes in input weighted by their respective share in total costs, where $w = (w_1, ..., w_N)$ is the vector of input prices. Taking logs of equation (7) and totally differentiating[16] results in the following expression:

$$d\ln y = \sum_{n=1}^{N} \frac{\partial \ln f(x, t, \beta)}{\partial x_n} dx_n + \frac{\partial \ln f(x, t, \beta)}{\partial t} dt - \frac{\partial u}{\partial t} dt$$

Dividing every term by dt and using the definitions of TFP, technical change and technical efficiency change we can rewrite this as

$$\dot{TFP} = TM + TEM + \sum_{n=1} \frac{\partial \ln f(x;t;\beta)}{\partial x_n}\frac{dx_n}{dt} - \dot{X}$$

Using the definition for the input change index, we can rewrite $\dot{X}$ in terms of output elasticities ($\varepsilon_n = \frac{\partial f(x;t;\beta)}{\partial x_n}\frac{x_n}{f(x;t;\beta)} = \frac{\partial \ln f(x;t;\beta)}{\partial \ln x_n}$) and elasticity of scale ($\varepsilon = \sum_n \varepsilon_n$).

$$\dot{TFP} = TM + TEM + \sum_{n=1} \frac{\partial \ln f(x;t;\beta)}{\partial x_n}\frac{dx_n}{dt} - \sum_{n=1} \frac{w_n x_n}{\sum_n w_n x_n}\frac{d\ln x_n}{dt}$$

$$\dot{TFP} = TM + TEM + (\varepsilon - 1)\sum_{n=1} \left(\frac{\varepsilon_n}{\varepsilon}\right)\dot{x}_n + \sum_{n=1} \left[\left(\frac{\varepsilon_n}{\varepsilon}\right) - S_n\right]\dot{x}_n$$

This last expression is obtained by adding and subtracting $(\varepsilon_n = \sum_n \varepsilon_n)\frac{d\ln x_n}{dt}$ and rearranging terms. It shows that the change in productivity can be decomposed into technical change, technical efficiency change and a scale component. We have an additional term that captures the change in allocative efficiency. Since we have no information on the prevalent input prices we assume away the allocative efficiency component by setting $S_n = \frac{\varepsilon_n}{\varepsilon}$.[17] This results in our final expression (9) that will be used in the empirical analysis

$$\dot{TFP} = TM + TEM + (\varepsilon - 1)\sum_{n=1} \left(\frac{\varepsilon_n}{\varepsilon}\right)\dot{x}_n \qquad (9)$$

From now on we assume a specific functional form for $f(x;t;\beta)$, that is the translog production function. If we want to estimate and calculate the three different components of TFP change, we have to introduce a time variable t into the framework. Equation (10) represents the translog production function with a time trend allowing for a non linear time trend by taking both t and t^2 into the regression.

$$\begin{aligned} \ln y_{it} = {} & \beta_0 + \sum_{n=1} \beta_n \ln x_{nit} + \beta_t . t + \frac{1}{2}\sum_{n=1}\sum_{k=1} \beta_{nk} \ln x_{nit} . \ln x_{kit} + \frac{1}{2}\beta_{tt} . t^2 \\ & + \sum_{n=1} \beta_{nt} . \ln x_{nit} . t + v_{it} - u_{it} \end{aligned} \qquad (10)$$

The technical change is easily obtained by taking the partial derivative of $\ln y_{it}$ with respect to the time trend t and the change in technical efficiency is defined as before

$$TM = \beta_t + \beta_{tt} . t + \sum_{n=1} \beta_{nt} . \ln x_{it} \text{ and } TEM = -\frac{\partial u}{\partial t} \qquad (11)$$

The third component in equation (9) depends on the elasticity of scale under which the firm is producing. If the firm exhibits constant returns to scale, a change in the input will not affect its productivity. Formally this means that the last term disappears because under constant returns to scale since $\varepsilon = 1$. In this case the change in productivity is due to the other components. However, if the firm exhibits increasing or decreasing returns to scale, we have three components of TFP change. Whether this last term is positive or negative depends on the change in the input factors. Even with increasing returns to scale ($\varepsilon > 1$) we can have an additional negative term if the rate of change of inputs is negative. The same is true under decreasing returns to scale and a positive rate of change of inputs.

A variety of tests can be performed within this framework. We can check whether technical change is neutral with respect to inputs by testing the hypothesis that $\beta_{nt} = 0$ or whether there is no technical change at all by testing $\beta_t = \beta_{tt} = \beta_{nt} = 0$. The described framework is truly capturing the nature of our data. All three different components are allowed to vary over time and across firms. The difficulty lies in disentangling the technical change and the change in technical efficiency since often they both rely on the same variable t.[18]

3 Data

We use firm-level data of Belgian, Polish and Czech firms active in the manufacturing sector between 1995 and 1999. The data consist of all companies that have to report full company accounts to the national statistical offices for which at least one of the following criteria is satisfied: total turnover of at least 1 million Euro, total assets of at least 1.5 million Euro or total employment of at least 10 employees. This is a commercial database collected by "Bureau Van Dijck" which is a quoted software and consulting company on the (Euronext) stock market and the data are sold under the name of 'Amadeus'. The data include information on sales, employment, total wage bill, etc. and are representative for the manufacturing sector. Appendix 1 shows the share of employment in our sample in the total manufacturing employment, going from 90% in Belgium to 45% in the Czech Republic. Overall, our data set covers between 57% and 72% of the total manufacturing of Belgium, Poland and the Czech Republic.

Comparison of the data across countries is facilitated as Bureau Van Dijck harmonises the reported data to make comparisons of financial items across countries easier. For each firm we also have information on the sector of activity (within the manufacturing) it has its main operations in, which implies that we are able to make not only country comparisons, but also intersectoral comparisons.

We use operating revenue as a proxy for output (y) in the production function.

The number of employees and tangible fixed assets are the inputs, labour and capital respectively. We use the (nominal) book value of capital expressed in thousands of Euros and represents the true historical capital value correcting for depreciation. Our data consists of 34,026 observations over 5 years (1995-1999). We have 7,291 firms including 746 Polish firms, 1,953 Czech firms and 4,592 Belgian firms.

We drop outliers in our data since these outliers can be detrimental for the estimated production frontier and may result in over- or underestimated technical efficiencies. We drop the top and bottom one percentiles from the distribution of employment, operating revenue and capital. This guarantees us that we exclude from our sample unrealistic high values for the different variables. As a double check we look at the growth rates of the different variables and eliminate those with a year to year growth rate of over 250%. However, this means that we loose small firms, e.g. firms that had one employee in 1995 and had four in 1996. In this way, we are left with relatively big firms in terms of size and this is reflected in the average size of the firms in our sample. However, estimating a production function for small firms with e.g. only one employee is not very appealing.

3.1 The Pooled Data

The pooling of the three countries into one data set, however, introduces a few additional issues. Firstly, all Belgian firms are on average much smaller then the Polish and Czech firms. This heterogeneity issue is usually problematic when estimating production functions at the firm level. But given the framework of the Stochastic Frontier where the error term consists of a firm-specific efficiency part, the heterogeneity should be less worrying. Previous studies that looked at efficiency of firms in transition countries created a national benchmark (Piesse and Thirtle, 2000; Konings and Repkine, 1994 and others). To our knowledge there are only two papers so far where a common benchmark is created; i) Osiewalski, Koop and Steel (1997) create a world frontier to analyse output level and output growth in Poland and different Western countries, and ii) Funke and Rahn (2002) create a common benchmark for East and West Germany using stochastic frontier techniques. In this study we pool the data of the three countries enabling us to create a common benchmark for the three countries.

Secondly, a production function describes the transformation of inputs, measured in physical terms and thus we do not deflate our output since we are interested in the technical efficiency of firms.[19] These measures are not biased by inflation as we transformed all variables into Euros, provided the real exchange rate is approximately 1.

3.2 Summary Statistics

In this section we present some descriptive statistics of the data set, since an understanding of the structuring of the data is very important for the setup of the estimation of the technical efficiencies (TE).[20] Since we are dealing with an unbalanced panel and we want to compare average technical efficiencies, it is important to see the structure of the panel. Table 1 shows the number of firms that are present in the panel for 5, 4, 3, 2 and 1 consecutive year(s). We can see that 81% of the firms in our sample run through the entire time period 1995-1999. Moreover, only 5% of firms from the sample appear once or twice (consecutively) and they are all Czech firms. We end this section by reporting the average employment and output (in 1000 euro) for the three countries for every year (1995-1999). These results are reported in table 2. We can see that on average Belgian firms are much smaller and the level of employment is more stable over time as compared to Polish and the Czech firms. The average size of CEE firms is, however, decreasing over time. This is consistent with the general finding of shedding off labour in former planned economies due to massive labour hoarding during the central planned period (Roland, 2000). This time pattern is also found for other CEE countries, e.g. Slovenia (Hutchinson, 2002). Average labour productivity is higher in Belgium, although the differences are not big. This is because we use operating revenue as a measure of output. Ideally one would like to use value added to get a correct measure of labour productivity. We can thus see these differences in labour productivity as lower bound differences.[21]

4 Results

To obtain a measure for technical efficiency we use maximum likelihood estimation techniques to estimate a stochastic frontier production function. While estimating the average efficiency relative to some best-practice for a number of sectors is interesting, recently more effort has been made to estimate it at the firm level. We use the program FRONTIER 4.1 (Coelli and Battese, 1994) to estimate the maximum likelihood parameters of the production function and firm specific technical efficiency estimates.[22] We emphasize the fact that all results are based on a common best-practice frontier estimation for three countries and this is important when looking at the results.

We apply the B&C(1992 and 1995) framework to the data described in the previous section. We first look at the cross sectional case for the years 1995 and 1998. We explore the different functional forms of the underlying production function. In a second section we exploit the time dimension of the panel and allow for time varying technical inefficiency. In a third section we estimate technical efficiency at the sectoral level, leaving the implicit assumption of a common technology across

industries. Using a partiala adjustment model we check wether technical efficiency in CEEs is converging towards the EU level. Finally, we develop a methodology to estimate total factor productivity change at the firm level built on the Lovell and Kumbhakar (2000) and the B&C (1995) framework.

4.1 Cross Sectional Analysis

We use the B&C (1992) model expressed in equation (3) to estimate the technical efficiency (TE) estimates for 1995 for Belgium, Poland and the Czech Republic. We have 6,628 firms in 1995, 746 of which are Polish, 4,592 Belgian and 1,290 Czech firms all active in the manufacturing sector. We do this for both the Cobb-Douglas and Translog production functions represented in equation (5) and (6) respectively. We first report the estimates of the production functions, we denote the Cobb-Douglas as model A and the translog as model B. Table 3 represents these estimates for model A and B. We see that all coefficients are significant at the 1% level. Since model A and B are both in logs, the above estimates are elasticities (e.g. if we increase the labour force with 10% than output increases with 2.171% and 4.389% for model A and B respectively). The interpretation, however, is not so interesting since we pooled three countries and come up with constant elasticities. The log-likelihood function value of model B is higher than that of model A. With this information we can test model B against model A by means of a likelihood ratio test (provided that model A and B are nested).[23] The test-statistic is 151.9096 and thus this hypothesis has to be rejected at any appropriate significance level. Of course, this test only tells us that our model reaches a significant higher likelihood value when we use a translog production function and cannot be used on itself to judge on the underlying production functional form.

In table 4a and 4b we report the technical efficiency estimates for model A and B respectively. These tables are at the core of this paper. We see that Belgium firms are on average 3% more efficient than those in Poland and 10% more than those in the Czech Republic. There are no big differences between the TEs from model A and B. Model B results in slightly higher technical efficiencies for Poland and the most inefficient firm in Poland falls back to 16.97% which remains bigger than the one in Belgium. Of course the mean does not completely reveal the distribution of the TE in the different countries. Using the firm-specific estimates, Figure 1 represents the distribution of the TE per country, we also insert the overall average[24] (0.6218) of model A. We see that TEs are quite symmetrically distributed in all three countries. The distribution of TEs in Belgium is centered around a bigger average than the overall average. These distributions confirm our prior findings.

We can test whether these distributions are significantly different from each other by means of a Kruskal-Wallis test.[25] We have to reject this both for model A and

model B, with test-statistics of 1555.25 and 1181.28 respectively. These are to be compared with the critical values of the χ^2 distribution with 2 degrees of freedom (we have 3 countries). As a final criteria to compare the distribution of TE in the different countries, Table 5 shows the kurtosis and the skewness for the countries for model A and B. Table 5 shows that the TE in Poland is much more symmetric around its mean and with a lower peak. Both Belgium and the Czech Republic are more spread and skewed to the right around their averages. These four moments of the distribution of TEs are sensitive to the choice of the underlying production function. For the translog model (B), we see that all variances are higher and that now in all three countries the TE is skewed to the right.

We repeat the analysis that we conducted above for the year 1998.[26] We briefly report the results and focus on the comparison with the year 1995. Further on in the paper we allow for repeated observations over time and use panel data techniques. Comparing table 6 and 3 we see that he production function coefficients have not changed in sign and only marginally in elasticities, except for the elasticity of labour for model B. Table 7a and 7b represent the technical efficiency estimates for model A and B for the year 1998. The main conclusions remain the same, Belgium firms are on average more efficient than Polish firms, who are more efficient than the Czech firms. We see that the TE level is higher than in the year 1995: the overall TE average for model A and B increased by ca 5% and ca 3% respectively. The Czech firms' TE increased more than the Polish firms. We could argue that the Czech firms' efficiency increased more than that of the firms in the other two countries. Since we have an unbalanced panel we can only say something on the distribution of TEs and we explore the change of efficiency on a firm level later on.[27] We note that the overall efficiency in 1998 for model A and B is no longer only marginal different. The translog production function results in smaller estimates which makes sense since more (nonlinear) effects are considered in the production frontier than in the first order approximation (model A). Finally we report the first four moments of the three distributions of technical efficiency in Table 8 and these distributions are shown in Figure 2. Again the translog production function framework introduces a bigger variance (this can also been seen by comparing the minima and maxima of the TE under model A and B).

We use simple economic reasoning to evaluate the change over time of the technical efficiency, following Jovanovic (1982) framework, as suggested by Funke and Rahn (2002). He assumes a competitive industry with known time path of future output prices where firms differ in efficiency. Firms have a specific cost function $\phi_i c(Y)$ with $c(Y)$ a common convex cost function and ϕ_i is a firm-specific inefficiency parameter. The firm maximizes profits $\pi = p.Y - \phi_i^* c(Y)$ with $\phi_i^* = E(\phi_i | I_t)$[28]. Totally differentiating the FOC at $dP = 0$ gives us a relationship between the output

and the (in)efficiency level. Output is negatively (positively) related to inefficiency (efficiency) or formally $\frac{dY}{d\varphi_i^*} = -\frac{c'(Y)}{\varphi_i^* c''(Y)} < 0$. A firm will considers its efficiency level and will adjust later on accordingly. An efficient firm will grow because it adapts its output accordingly, whereas an inefficient firm becomes smaller. Jovanovic (1982) assumes that there is a threshold value φ^c such that all firms with $\varphi_i < \varphi^c$ will exit the market. This results in a higher mean efficiency level and a lower variance around it. Applied on our results[29], we see that indeed average efficiency increases and that the variance around it decreases over time. This is shown in Table 9. From Figure 1 and 2 we could argue that the the threshold values are different across countries, i.e. $\varphi^c_B > \varphi^c_P > \varphi^c_{CZ}$ and this results in a larger share of relative more efficient firms in Belgium.

4.2 Panel Data Analysis: Allowing Technical Efficiency to Change over Time

In this section we use the B&C (1992, 1995) specifications to estimate TE for the period 1995-1999.[30] Panel data allows us to relax the strong distributional assumptions of the cross sectional models and we can obtain estimates with more desirable statistical properties (see Kumbhakar and Lovell (2000) for a discussion). We experimented with different specifications and only report the 'preferred' specification.

The B&C (1992) specification allows for technical efficiency estimates to vary over time. However, as stated earlier this approach does not allow for a firm to become more or less efficient than another firm. The fixed ordering of firms by their TEs is not very appealing. We briefly discuss the results of the panel data model specification and check whether TE is significantly changing over time. This analysis can be done in two different ways. We can use the B&C (1992) model and estimate the parameter η together with the other parameters. A second way to perform panel data estimations - in a less structured way - is to take B&C (1995) specification and introduce determinants in the z-variable in equation (4). We follow the last approach[31] for the period 1995-1999 and use the translog production function as the underlying framework in all estimations. For these estimations we have an unbalanced panel of 7,291 firms. The model we estimate is represented as follows

$$\begin{aligned} \log Y_{i;t} &= \beta_0 + \beta_1 \log L_{i;t} + \beta_2 \log K_{i;t} + \beta_3 (\log L_{i;t})^2 + \beta_4 (\log K_{i;t})^2 \\ &\quad + \beta_5 \log K_{i;t} \log L_{i;t} + (v_{i;t} - u_{i;t}) \end{aligned} \quad (12)$$

with $u_{i;t}$ nonnegative random variables and assumed to be independently distributed as truncations at zero of the $N(\mu_{i;t}, \sigma_u^2)$ where $\mu_{i;t} = z_{i;t}\delta$. We proceed in three steps, i) we only allow for a constant term (model C.1), ii) we introduce year dummies (model C.2) and iii) we introduce subsectoral dummies (model C.3).

Firstly, we only allow for a constant term in the z-variable and this is just the Aigner, Schmidt and Lovell (1977) model specification. Table 10 shows some descriptive statistics of the TE estimates per year per country. We see that again Belgium firms are on average more efficient than their Polish and Czech counterparts, but that the Polish average is very close to the Belgian one. It is hard to compare these results with the ones obtained from the cross sectional analysis. Cross sectional TE estimates are only moment registrations of a firm's efficiency whereas the panel data approach allows to trace this more accurately because of repeated information. However, this model specification is very simple in the way the inefficiency effects u are modelled over time and it is not fully capturing the strengths of the panel data techniques.

Now we introduce more determinants of TE and estimate these simultaneously with the TE estimates. Therefore we use the B&C (1995) specification that allows for time varying inefficiencies in a more flexible way, the $z_{i,t}$ are chosen freely (see equation (4)) in contrast to the fixed function in the B&C (1992) specification (equation (3)). For now, we only use time dummies to correct for macro economic shocks and differences in inflation and producer price inflation rates between the three countries. We take the year 1995 as our reference year and thus have four year dummies for 1996, 1997, 1998 and 1999. So the vector $z = (D_1; D_2; D_3; D_4)$ plays the role of a very simple model to allow for time variant TE.[32] Maximum likelihood is used to estimate the production frontier coefficients and the technical efficiencies and the parameters $\delta_0; \delta_1; \delta_2; \delta_3$; and δ_4. We first report the estimated production function estimates in Table 11. The estimated coefficients have the same signs and magnitudes as model B in 1995 and 1998. For the same reasons as before we do not devote a lot of attention to the interpretation of the elasticities, but we do check whether these are sensible. The δ-parameters are all negative and significant from zero, this means that TE is increasing significantly over time.[33] Table 12 represents the estimated TE for the three countries for every year and the average over the five years (1995-1999). Of course this table does not exploit the fact that we have firm specific TE for 5 years (except for the firms that are not in the panel for five years). We see that Belgian firms remain more efficient on average, but the difference with Polish firms is much smaller than in the cross sectional cases. We see that this model specification delivers higher technical efficiency estimates. However, we cannot compare the TE in 1995 from Table 12 with the TE from table 4a-b since these are different model specifications. The TE of the three countries are surprisingly close to each other and this could be due to the very simple structure we choose for the $\mu_{i;t} = z_{i;t}\delta$ part of the model.

Finally, we introduce sectoral dummies ($SECT_{ij} = 1$ if firm i is in sector j) as explanatory variables for the technical efficiency in order to capture industry-specific

inefficiency effects. The system exists of equation (13), (14) and (15) and the results are presented in Table 13. We note that these estimates are averages over the period 1995-1999.

$$\log Y_{i,t} = \beta_0 + \beta_1 \log L_{i,t} + \beta_2 \log K_{i,t} + \beta_3 (\log L_{i,t})^2 + \beta_4 (\log K_{i,t})^2 + \beta_5 \log K_{i,t} \log L_{i,t} + (v_{i,t} - u_{i,t}) \quad (13)$$

$$u_{i,t} \sim i.i.d.\ N^{+}(\mu_{i,t}, \sigma_v^2) \quad (14)$$

$$\mu_{i,t} = \delta_0 + \sum_{j=1}^{22} \delta_j SECT_{i,j} + \varepsilon_{i,t} \quad (15)$$

The subsectors in bold are those with an above national average efficiency level in every country. 'Tobacco Products' is the most efficient subsector in every country and sets the benchmark.[34] We do not go further into the subsectoral analysis since up till now we have assumed identical technologies. We drop this assumption in the following section and allow for different technologies between the subsectors.[35]

4.3 Sectoral Analysis of Technical Efficiency

In this section we look for sectoral patterns within and across the three countries based on the cross sectional TE estimates of 1995 and 1998.[36] We split up the manufacturing sector in 23 subsectors following the NACE 2-digit classification (see Appendix 3). We compare the sectoral pattern of TE between 1995 and 1998. In a second stage we drop the assumption of identical production technology for every subsector and we estimate a best-practice frontier at the subsector level and suggest an interpretation for the technical (in)efficiency.

4.3.1 A Sectoral Blueprint of Technical Efficiency

We find that the subsectoral averages differ when we use model A or B, and this difference is most apparent for Poland and the Czech Republic. We demean the TEs (using model A) from their respective national unweighted average and we can identify sectors that are above the national average; these deviations are calculated in the following way: $(\overline{TE}_{j,N} - \overline{TE}_N)$.[37] In 1995 we can identify four sectors that have a positive deviation in every country, the 'Tobacco Products', 'Chemicals', 'Basic Metals' and 'Office Machinery and Computers'. There are 9 common sectors (17, 18, 20, 24, 26, 28, 33, 35 and 36) with negative deviations from their respective national average. In 1998 there are only three common sectors with a positive deviation. The sector 'Tobacco Products' appears to be the most efficient sector in each country, with an average TE in 1995 of 66.12%, 60.98% and 70.37% for Poland, the Czech

Republic and Belgium respectively. We discuss the sectoral analysis in more detail in the next section, where subsectoral frontiers are estimated at the NACE 2-digit level.

4.3.2 Subsectoral Estimation of Technical Efficiency for 1995 and 1998.

We drop the assumption that every subsector of the manufacturing sector is producing under the same technology. For this, we estimate a best-practice frontier at the 2-digit NACE level. We use the translog production function for all estimations. For some sectors we have few observations, which makes the estimation less robust. Cross sectional frontier analysis is in itself problematic when it comes to having consistent estimates, when facing few observations this problem is exacerbated.

We estimate firm level technical efficiencies for every subsector (15,..., 37) using the B&C (1992) model and keep η restricted to zero since we are dealing with a cross sectional case. We consider two specifications, a model that restricts $\mu = 0$ and another that does not have this restriction, model D and model E respectively. The former specification is the same as suggested by Pitt and Lee (1981) whereas the latter coincides with the model proposed by Stevenson (1980). These specifications are tested against each other by means of a log likelihood ratio test since these models are nested by a linear restriction.[38] The null has to be rejected for sectors 21, 23 and 28.

Before we discuss the results we have to stress the fact that the subsectoral TE estimates are not to be compared between subsectors or with the results from previous sections. A TE of e.g. 60% for a subsector j reflects the average relative technical efficiency within the subsector j, however this subsector can still be more efficient than a subsector $j+1$ with an average of 75%. The estimated TE gives an insight in the relative difference in efficiency within each subsector of the manufacturing sector, i.e a measure of firm heterogeneity. If one is interested in knowing whether a given subsector is more efficient than another, one has to follow the analysis presented in the previous section[39]. Table 14 presents the average estimated TE at the subsectoral level for the three countries and the overall average is presented as well.[40] The numbers in Table 14 should be interpreted as follows: a relative high overall TE average reflects rather homogeneous TE levels, although Belgium firms are consistently on average more efficient (with exception of subsectors 23 and 30). Subsector (35) with an overall TE of 23.29% has a wide divergence in firm TE levels within this subsector and for each country. Indeed, we see that this low average is due to the very low average of Poland and Czech Republic, 18.56% and 13.01% respectively. The same remark that was made in previous sections is still valid here, i.e. the average is very sensitive to the proportion of Belgian firms in a (sub)sector. We therefore check whether the distribution of TEs is different for

the three countries within every subsector by means of a Kruskal-Wallis (KW) test. Table 15 presents the results of this test.

We stress the fact that we performed this test within a subsector and we check whether the TE are distributed differently in Poland, Czech Republic and Belgium. We see that the distribution of TE in Belgium, Poland and Czech Republic of 6 subsectors are not statistically different from each other. The other subsectors' distributions are statistically different from each other between the three countries. However, if we want to compare subsectoral TE distributions across the three countries and perform a KW-test, we see that subsectors have significant different distributed TEs.[41] However, this test is done for the whole manufacturing sector. This result is quite intuitively, we can expect a textile producer (17) to use a different technology than a chemical firm (24).

4.4 Partial Adjustment to Analyse Convergence of CEE Countries

In the treaty of Nice 2000 a roadmap was approved that gives a potential timing of the (most early) accession of CEE countries. For firms in CEEs to be able to meet the competitive pressure in the single market, firms' efficiency level should converge towards the EU level. Given this 'road map' the European Union set up, it is interesting to see whether the efficiency in the candidate countries is converging towards the EU level. In this section we analyse whether Polish and Czech firms' TE converges to the level of Belgian firms over time. Since our time period is rather short (1995-1999), we expect to find quite low speed of adjustment parameters. We use the following Partial Adjustement (PA)-model for Poland and the Czech Republic respectively

$$TE^{j,P}_{i,t} - TE^{j,P}_{i,t-1} = \lambda_P(\overline{TE}^{j,B}_t - TE^{j,P}_{i,t-1}) \qquad (16)$$

$$TE^{j,C}_{i,t} - TE^{j,C}_{i,t-1} = \lambda_C(\overline{TE}^{j,B}_t - TE^{j,C}_{i,t-1}) \qquad (17)$$

with the following definitions: $TE^{j,P}_{i,t}$ = Technical Efficiency of a Polish firm i active in subsector j in the year t, $TE^{j,P}_{i,t-1}$ = Technical Efficiency of a Polish firm i active in subsector j in the year $t-1$, $\overline{TE}^{j,B}_t$ = Average Technical Efficiency of Belgian firms active in subsector j in the year t and λ = speed of adjustment parameter ($\lambda = 0$: no adjustment, $\lambda = 1$: fully adjustment)

For these regressions we only use the firms that are in our sample for five consecutive years, creating a balanced panel. This results in 2740 Polish and 3712 Czech observations over five years. Table 16a reports the results of the regressions for Poland and the Czech Republic. We see that the speed of adjustment parameter of Poland is higher than the one of the Czech Republic. The magnitude of λ is low as

expected and significant for both countries. This means that Polish and Czech firms are on average getting closer to the EU benchmark created by the Belgian firms.

However, the model above results in an average convergence parameter λ. We drop this assumption and introduce time dummies (for the four years 1996-1999) and time dummies interacted with the independent variable. The fact that we have no constant term in our model implies that we take all four dummies and all four interaction terms into the regression. Taking the time dummies into the regression allows us to have different year-intercepts. The PA-model with time dummies is represented in equations (18) and (19) for Poland and the Czech Republic respectively.

$$TE_{i,t}^{j,P} - TE_{i,t-1}^{j,P} = \sum_{s=1}^{4} \lambda_{Ps}(\overline{TE}_{t}^{j,B} - TE_{i,t-1}^{j,P}).D_s + \sum_{s=1}^{4} \alpha_{Ps} D_s \qquad (18)$$

$$TE_{i,t}^{j,C} - TE_{i,t-1}^{j,C} = \sum_{s=1}^{4} \lambda_{Cs}(\overline{TE}_{t}^{j,B} - TE_{i,t-1}^{j,C}).D_s + \sum_{s=1}^{4} \alpha_{Cs} D_s \qquad (19)$$

The results are shown in Table 16b. We see that for both Poland and the Czech Republic the results in Table 16a hide the differences of the convergence across years. All coefficients are significant at the 1% significance level. The convergence parameter[42] is always positive, implying that firms in Poland and the Czech Republic are getting closer to the Belgian average, which itself is changing over time. However, the rate at which they do is decreasing indicating a concave time pattern. We see that Polish firms have on average a higher speed of convergence and that the fall in the magnitude of the rate is smaller than that of the Czech firms. The extension of the PA-model indicates that there is indeed a heterogeneous process going on over time, and thus makes the simple PA-model expressed in equations (16) and (17) less appealing. Note that the speed of adjustment in the Czech Republic is 0.1091 in 1996, whereas the average over time in Table 16a suggest a parameter of only 0.0305. These findings are quite intuitive, we expect to find that the rate of convergence decreases over time since it is harder to get closer to the benchmark the smaller the distance to this benchmark becomes.

4.5 Estimating TFP Change in a System of Equations.

By combining two existing techniques (Lovell and Kumbhakar (2000) with B&C (1995)) we develop an empirical methodology to decompose productivity growth into various elements. One of these elements is technical efficiency, in this way we can see how much efficiency improvements are feeded into productivity gains. Up till we now we have assumed structured functional forms for the inefficiency effect u_{it} (e.g. BC,1992). We introduce a more flexible framework and introduce

a system of equations where determinants of technical efficiency are estimated simultaneously with the production frontier.[43] This setup allows us to capture all possible movements within the input-output space (X, Y). That is: i) keeping the frontier constant, moving closer to or further away from the frontier reflects a change in technical efficiency, ii) keeping the frontier constant, moving along the horizontal axis, reducing or increasing inputs and finally iii) a shift of the frontier itself, reflecting a change in technology or technical change. The second move (ii) depends on the economies of scale the firm is producing under. The analysis we pursue is disentangling these effects at the firm level using the methodology suggested above (section 2.3 and 2.5) as building blocks.

We use the translog production function with labour and capital as inputs.[44] We introduce a time trend and allow for a non linear trend effect. We interact the inputs with the time trend to allow for non neutral technical change. Finally we interact the time trend with country dummies to allow for country-specific technical changes[45].

We assume that the efficiency effect u_{it} is a function of subsectoral dummies, a time trend and interacted terms. We derive the different components of TFP change within this framework and give expressions for the different estimated components of the TFP change. The system of two equations (20) and (21), representing the stochastic frontier production and the technical inefficiency effect respectively, is estimated in one step using maximum likelihood and looks as follows

$$
\begin{aligned}
\ln y_{it} = {} & \beta_0 + \beta_1 \ln L_{it} + \beta_2 \ln K_{it} + \frac{1}{2}\left[\beta_3 \ln L_{it}^2 + \beta_4 \ln K_{it}^2 + \beta_5 \ln L_{it}.\ln K_{it}\right] \\
& + \beta_6.t + \frac{1}{2}\beta_7.t^2 + \beta_8 \ln L_{it}.t + \beta_9 \ln K_{it}.t + \beta_{10} BEL_i.t + \beta_{11} POL_i.t \\
& + \beta_{12} BEL_i + \beta_{13} POL_i + v_{it} - u_{it} \qquad (20)
\end{aligned}
$$

$$
u_{it} = \delta_0 + \delta_1.t + \delta_2.t^2 + \sum_{j=3}^{25} \delta_j SECT_{ij} + \sum_{j=26}^{46} \delta_j SECT_{ij}.t + \varepsilon_{it} \qquad (21)
$$

with i the firm index, t the time index and the subsectors within the manufacturing sector are denoted by j.[46] As before we have a composed error term in the production function, v_{it} is a random component that is $i.i.d.(0, \sigma_v^2)$ and independent of the inefficiency effect and the error term ε_{it} in the second equation. We allow for non linear technical efficiency changes by adding t^2 in the second equation. We include sectoral dummies in the regression ($SECT_{ij}$) and introduce an interaction term between the time trend and the subsectoral dummies to capture the subsectoral differences in technical efficiency change, since we know from the previous section that efficiencies are signifantly different across the subsectors within the manufacturing sector.

The estimated components and parameters are denoted with hats. We stress that the parameter significancy is of course an issue here, and only significant parameters are relevant throughout the whole analysis. The estimated technical change ($\widehat{TC}$), estimated technical efficiency change ($\widehat{TEC}$), estimated output elasticities ($\hat{\varepsilon}_L$ and $\hat{\varepsilon}_K$) and the scale elasticity ($\hat{\varepsilon} = \hat{\varepsilon}_L + \hat{\varepsilon}_K$) are represented in equations (22), (23), (24), (25) and (26) respectively.

$$\widehat{TC}_{it} = \hat{\beta}_6 + \hat{\beta}_7 . t + \hat{\beta}_8 \ln L_{it} + \hat{\beta}_9 \ln K_{it} \tag{22}$$

$$\widehat{TEC}_{it} = -\left(\hat{\delta}_1 + 2.\hat{\delta}_2 . t + \sum_{j=3}^{25} \hat{\delta}_j SECT_{ij} + \sum_{j=26}^{46} \hat{\delta}_j SECT_{ij} . t\right) \tag{23}$$

$$\hat{\varepsilon}_{L_{it}} = \hat{\beta}_1 + \hat{\beta}_3 \ln L_{it} + \hat{\beta}_5 \ln K_{it} + \hat{\beta}_8 . t \tag{24}$$

$$\hat{\varepsilon}_{K_{it}} = \hat{\beta}_2 + \hat{\beta}_4 \ln K_{it} + \hat{\beta}_5 \ln L_{it} + \hat{\beta}_9 . t \tag{25}$$

$$\hat{\varepsilon}_t = \hat{\phi}_0 + \hat{\phi}_1 \ln L_{it} + \hat{\phi}_2 \ln K_{it} + \hat{\phi}_3 . t \tag{26}$$

with obvious definitions for $\hat{\phi}_0$, $\hat{\phi}_1$, $\hat{\phi}_2$ and $\hat{\phi}_3$. Note that the different components are firm specific and time varying. The above expressions are in their most general form and it is easily seen how these expression simplify under different parameter restrictions like e.g. no non-neutral technical change ($\beta_8 = \beta_9 = 0$).[47] Now we have a final expression for estimating the change in total factor productivity under a simultaneous translog stochastic frontier (STSF) model. The change in TFP is defined in equation (9) and using above expressions results in equation (27).

$$\dot{\widehat{TFP}}_{it} = \widehat{TC}_{it} + \widehat{TEC}_{it} + (\hat{\varepsilon}_t - 1)\left[\frac{\hat{\varepsilon}_{L_{it}}}{\hat{\varepsilon}_t}\dot{L}_{it} + \frac{\hat{\varepsilon}_{K_{it}}}{\hat{\varepsilon}_t}\dot{K}_{it}\right] \tag{27}$$

with $\dot{L}_{it} = \ln L_{it} - \ln L_{it-1}$ and $\dot{K}_{it} = \ln K_{it} - \ln K_{it-1}$, the rate of change of labour and capital respectively. We only report the estimated components of the TFP change and the TFP change itself. The production function estimated coefficients and all the δ's are significant different from zero and are used to estimate the TFP change[48][49]. Table 17 illustrates the differences in the (average) TFP decomposition between Belgium and CEEs. TFP growth in CEEs firms is mainly due to an increase in technical change (a shift upwards of the production frontier), whereas the Belgian firms TFP growth is also driven by an increase in technical efficiency (a move towards the frontier). Belgian TFP change clearly exhibits a different pattern, having a persistence in technical efficiency change although at a decreasing rate.

The scale component is quite large for the Polish firms in comparison with that of the Czech and Belgian firms. This might be due to the size of an average Polish firm. We know from Table 2 that firm size is - on average - bigger in Poland and thus reducing labour into the production process can explain this larger share. In

order for this explanation to hold, an average Polish firm needs to produce under decreasing returns to scale. Thus, before we can interpret the scale component of the TFP change we have to make inferences about the returns to scale. Table 18 presents this average elasticity per country and per year. The increasing share of the scale component means that on average - and under decreasing returns to scale - firms in the manufacturing sector are reallocating resources, i.e. reducing labour and/or capital.[50] In the case of Poland and the Czech Republic we know from section 3 that average employment is decreasing over time. It is a general finding that firms in CEEs under the planned regime suffered from labour hoarding (Roland, 2000).

5 Interpretation of the Results and Some Policy Implications

In this section we briefly interpret the estimated technical efficiencies on a more general level. The interpretation of (in)efficiencies goes back to the concept of Leibenstein's X-inefficiency first proposed in 1966. He stated that the motivation to reduce production costs comes primarily from external pressure. This leads to assume that the reasons firms do not maximise profits is because of effort discretion. The view was later on criticised, e.g. Stigler (1976) and Demsetz (1969) to name but a few. The latter refers to the common practice of comparing the real world with an ideal but not existent world to conclude that the real world is (relatively) inefficient (Fidalgo and Garcia, 2000). It remains a fact that the very concept of technical inefficiency cannot be rationalised with the tools of neoclassical theory of the firm, since this notion violates central assumptions of economic theory. However, traditional production theory is used to develop techniques to measure such misperformances. How to go from here then?

As suggested in the sectoral analysis above, we see the technical efficiency measure rather as a measure of heterogeneity of the firms in the sample, violating the assumptions made by economic theory. It is clear that if some firms perform better than others it can be because they are different and not homogeneous as is implicitly assumed when creating a common production frontier. The issue of inefficiency than appears to be one of heterogeneity and therefore leads to the question why are firms different? Our intepretation of TE as a measure of heterogeneity coincides with the view put forward by the resource-based view of technology (Fidalgo and Garcia, 2000). This view states that technology differs across firms even in the same industries, because firms usually posses some resources and capabilities which are unique and which are ignored in the estimation of the efficiencies. This is what we find when estimating the TE at the 2-digit NACE level. It is thus very important to be aware of the common technology assumption when interpreting the results. Within

this setup, estimated TE measures the relative value of resources and capabilities not observed or not included in the model that were assumed to be homogeneous across firms. The latter can be interpreted as a measure of competitive advantage of the firm, although restricting it to the productive component. In our case this means that firms in the EU still have a competitive advantage over the firms in CEE countries and more precisely a productive advantage. However, given the results using the partial adjustment we see that this advantage is becoming smaller and thus implies that firms in CEE are getting more competitive.

6 Conclusion

We created a common production frontier for two EU-candidates (Poland and the Czech Republic) and a EU country (Belgium). As expected, Belgian firms set the benchmark and are on average more (technically) efficient than firms in Poland and the Czech Republic. However, Polish firms are very close to the Belgian level, whereas Czech firms are further away from the benchmark. Average technical efficiency (TE) is around 60% and points to the fact that firms are failing to be profit maximizers or cost minimizers as suggested by neoclassical micro economic theory.

If we estimate technical efficiency at a more disaggregated level (2 digit NACE) we find that there are big differences between subsectors of the manufacturing sector. We interpreted the technical efficiency estimates as measures of heterogeneity of TE within a subsector. We identified three common industries ('Tobacco Products', 'Chemicals', 'Basic Metals') that are relatively efficient compared to a national average. When allowing for repeated observations over time, we find evidence for an increasing TE. The panel data results confirm the results suggested in the cross sectional analysis, although the former are preferred because they are based on repeated observations.

Using a partial adjustment model we find evidence for convergence of the efficiency levels of firms in CEE countries towards an EU level. However, the convergence is at a decreasing rate, starting off higher in Poland than in the Czech Republic, suggesting a concave time pattern of convergence.

Combining two existing techniques we develop a methodology of estimating a firm and year specific decomposition of total factor productivity into several components and estimate this. In this way we can put the estimated technical efficiency in perspective. We see that is mainly technical change that drives the total factor productivity change. Firms in EU- candidate countries are reallocating resources under decreasing returns to scale and the scale component has an increasing share in TFP change. Belgian TFP change clearly exhibits a different pattern, having a persistence in technical efficiency change, although at a decreasing rate.

Notes

[1] Within this setup we allow for a CEE firm to set the benchmark as well as a Belgium firm. In this way we do not catagorise firms from CEE countries a priori as being less efficient. In a normal regression framework one could think to use country dummies, however these do catagorise producers by construction.

[2] See Fried, Lovell and Schmidt (1993) for a discussion on the different components.

[3] We note that a firm is technical inefficient if it is under the best-practice frontier and that it is technical efficient when it is on the frontier.

[4] We refer to these specifications as B&C(1992) and B&C(1995) respectively.

[5] Given that technical change is defined by an appropriate parameter model and the technical inefficiency effects in stochastic frontier model are stochastic and have the specified distribution. We incorporate this type of analysis in section 2.5.

[6] However, X is now a matrix capturing the two dimensions of the panel, i.e. time and number of producers.

[7] We stress that the vector u refers to the inefficiency effects (substracting output from the potential output). The firm specific efficiency, however, is measured in terms of technical efficiency.

[8] We present the model suggested by Lee and Schmidt (1993) as this approach shows that one can model time variant technical inefficiency effects and still use time invariant determinants (this cannot be done within the framework of traditional panel data techniques as fixed effects.) Their model is suited for applications in which the number of firms is large and the number of time series observations per firm is relatively small. Schmidt and Lee (1993) consider a linear production frontier. Let there be N firms indexed by $i = 1, \dots, N$ and T time series observations per firm, indexed by $t = 1, \dots, T$. Let $y_{i,t}$ denote the output (in logs) of firm i at time t, and $X_{i,t}$ the corresponding vector of inputs (in logs). As before $v_{i,t}$ represents statistical noise, whereas $u_{i,t} \geq 0$ represents technical inefficiency of a firm i at time t. The model in its general form can then be written as follows

$$y_{i,t} = \alpha_t + X_{i,t}'\beta + v_{i,t} - u_{i,t} = \alpha_{i,t} + X_{i,t}'\beta + v_{i,t} \qquad (28)$$

with $\alpha_{i,t} = \alpha_t - u_{i,t}$ the intercept for a firm i at time t. And α_t is the frontier intercept or the maximum possible value of $\alpha_{i,t}$. In this way they incorporated time variant technical efficiency at the firm level. Schmidt and Lee (1993) consider the model in which $\alpha_{i,t} = \theta_t \delta_i$ where the θ are parameters to be estimated. One can see that for $\theta_t = 1$ for all t, we are in the simple Panel Data model. Otherwise the temporal pattern of technical inefficiency is completely unrestricted although assumed the same for all firms. The model that these authors consider can also be compared to the two way analysis of covariance model that includes both individual and time effects. This model has $\alpha_{i,t} = \delta_i + \theta_t$.

[9] The only restriction they imposed is that all firms should have the same pattern of change.

[10] In the case of increasing technical efficiency effects, we see that firms have monotonically increasing technical inefficiency effects in the case of $\eta < 0$, the reverse is true if $\eta > 0$. However, one must bear in mind that increasing technical inefficiency effects ($u_{i,t}$) imply decreasing technical inefficiency estimates ($\exp(-u_{i,t}) = TE_{i,t}$). Thus if the estimation for parameter η is positive this implies that the efficiency of firms increase over time, at a rate determined by the exponential function depending on $(t - T)$.

[11] This is because the technical inefficiency effects of different firms at any given time period t are equal to the same exponential function ($[\exp - \eta(t - T)]\ \eta\ \exp(T - t)$) of the corresponding firm specific inefficiency effects at the last period of the panel (u_i).

[12] Since we have a rather small time dimension in our data, we could argue that the ordering of firms' efficiency is not changed. However this can be questioned because we are dealing with firms

in transition countries and our panel starts at the year 1995. After 1995, the transition process really took off and it can be expected that this process can alter the ordering of firms according to their technical inefficiency.

[13] This approach follows on the previous sections although there is a fundamental difference in the way time variant technical inefficiencies are modelled.

[14] This section relies mostly on Lovell and Kumbhakar (2000) and references herein.

[15] This definition is a conventional Divisia index of productivity change provided we are in a scalar output case.

[16] We assume that the technical efficiency component is a function of time, $u(t)$. This general representation allows us to incorporate different specifications (e.g. Battese and Coelli (1992) and Kumbhakar (1990)) into our analysis.

[17] This assumption implies that we assume perfect competition in the product market since prices are equal to marginal costs. When assuming the latter, the estimated total factor productivity estimates are biased if product markets are imperfectly competitive (Hall, 1988). This assumption of perfect factor markets is debatable, however, it is just a working assumption often used in empirical work (see e.g. Levinsohn and Petrin, 2001).

[18] For instance Battese and Coelli (1992) specification $u_{it} = [\exp(-\eta(t-T))]u_i$ is an example how the inefficiency variable is depending on t. However, we will partially overcome this problem by using a system of equations where technical efficiency itself is a function of different variables that change over time, but no longer directly depending on t.

[19] Note though that as we scale everything in logs this price difference is captured by the constant term. However, the producer price index may also capture monopoly power in the output market.

[20] The programm FRONTIER has a very stringent setup of the data set in order for the programm to converge.

[21] To fully compare labour productivity between the countries one needs to correct for quality differences as well. Since we construct a different measure to compare these countries, we do not go into these issues.

[22] We use estimate here and not measure - as often used throughout the literature - since these are actual estimates with standard errors around them. This allows us to make inferences about the TEs. Whereas deterministic approaches like DEA produce efficiency measures, that are just point estimates (Horrace and Schmidt, 1996).

[23] This test-statistic is obtained by taking twice the difference of the log likelihood values, and under the null hypotheses ($\beta_3 = \beta_4 = \beta_5 = 0$) has a χ^2 distribution with 3 degrees of freedom.

[24] This average is sensitive to the proportion of Belgian firms in the sample, which in our case is more than one third. However, we also encounter a lot of very inefficient Belgian firms.

[25] The null hypotheses is that the distribution of technical efficiency is the same in the three countries.

[26] We would like to compare the last period of our sample (1999) with the beginning of the sample (1995). However, the estimation of TE in the last year of the sample did not converge. This can be due to the lack of variation in this year.

[27] We did not create a balanced panel for the years 1995-1998 because this would introduce a sample selection bias, that is we would only compare the efficiency of firms that were active both in 1995 and in 1998.

[28] I_t is the information set at time t and as $t \to \infty$ $\omega_i^* = \omega_i$.

[29]We do not have the full population of …rms to adress the entry-exit issue in full.

[30]Note that here we are able to produce estimates for the year 1999. This is because the panel data approach uses both the variation in cross section and in time to estimate the technical e¢ciency.

[31]The …rst approach results in very low TE estimates due to various reasons. Even if we supply the program with starting values as to guide the maximum likelihood estimation iterative procedure, using di¤erent starting values, the results from the cross sectional estimations and various combinations of them, the results are not robust. This is why we prefer the B&C (1995) model.

[32]Alternatively one can include a time variable in stead of dummies. We use this approach when estimating the total factor productivity change later on.

[33]To see this, note that $\exp(x) > 0$ if $x < 0$ and that $TE_{i;t} = \exp(-u_{i;t})$.

[34]The tobacco products (16) industry has a high and stable demand for its products. The international competition coupled with a rather inelastic demand could explain this …nding.

[35]One could argue that introducing country dummies is exactly doing this at a more aggregated level, but then it is no longer informative to estimate technical e¢ciency. If one is interested in the frontier country di¤erences, an OLS regression with country dummies will do. By estimating e¢ciency at the 2 digit NACE level, we can still compare CEE …rms relatively to a EU benchmark (only within that subsector).

[36]The results of the panel data are less robust to functional form than the cross sectional results, so we base our sectoral analysis on the cross sectional TE estimates.

[37]Using following de…nitions that $\overline{TE}_{j;N}$ is the average TE of sector j from country N and $\overline{TE}_N$ is the national average TE for country N, with $j = (15; :::; 37)$ and $N = $ (Poland, Czech Rep., Belgium).

[38]The null hypothesis of having a correct restriction is tested against having an incorrect restriction, the test-statistic has to be compared with 3.841 (χ^2 critical value with one degree of freedom at the 5% signi…cance level).

[39]Although one could argue that assuming the same technology over all subsectors and over three countries is not realistic.

[40]The missing sectors 15, 16, 18, 32 and 36 are better estimated with OLS or need di¤erent starting values which results in very high TE estimates of ca. 0.99, meaning that there are no ine¢ciency e¤ects. The TE estimates of the three sectors where the restriction $\mu = 0$ was rejected, are based on model E.

[41]The null hypotheses states that TEs are distributed the same over every subsector. But just as in a test of overidentifying restrictions, we can not identify which sector(s) is (are) causing this hypothesis to be rejected. Think of the classical example of Verbeek (2000), if a pub allows you to get three beers and only pay two, could you tell which one is for free?

[42]We only report the adjustment parameters in tables 16a and 16b.

[43]This refers to Battese and Coelli (1995) representation. The ine¢ciency e¤ect u_{it} is assumed to be a normal-truncated variable with mean μ_{it} and variance σ_u^2. The mean μ_{it} is itself a function of vector of variables z_{it}. This is exactly the second equation in the system of equations.

[44]We use a di¤erent variant of the translog production function here. The second order terms are weighted by $1/2$, this simpli…es the further analysis.

[45]This implies that we have to include the country dummies as well to control for country di¤erences.

[46]There are 23 different subsectors in the manufacturing sector (according to the NACE rev.1 classification). Leaving the first subsector (15) out to avoid the dummy-trap, this implies having 22 dummies (from 3 to 25) and having 22 interaction terms (from 26 to 46).

[47]Except for the fact that we assume away the allocative efficiency component because we do not have any information on the prevailing input prices.

[48]There are 45 coefficients (for every subsector of the manufacturing (22) and interacted with the time trend (22) and the constant term) in the inefficiency model. Negative coefficients imply that the sector is more efficient than the reference sector (15, Food Products) and if the interacted term is significant, this means that technical efficiency is significantly changing over time in that given subsector. We do not go further into this, since our main interest lies in the different components of the TFP decomposition.

[49]The production function coefficients (β) and the coefficients from the inefficiency model (δ) are used to calculate the TFP decomposition. The table below shows the coefficients of the production function, where Y: significant at any appropriate significance level and N: not significant.

β	coefficient	significant
β_1	0.870	Y
β_2	-0.400	Y
β_3	-0.048	Y
β_4	0.029	Y
β_5	0.015	Y
β_6	0.190	Y
β_7	-0.019	Y
β_8	-0.005	Y
β_9	-0.001	N
β_{10}	-0.024	Y
β_{11}	-0.010	N
β_{12}	2.06	Y
β_{13}	4.24	Y

The coefficients for labour and capital (including the higher order) are as before. Coefficients β_6 and β_7 suggest a shift upwards of the production function, however at decreasing rate over time. The interpretation of the coefficients β_8 and β_9 (both inputs interacted with the time variable) confirm that labour is a flow variable and is significantly changing over time, whereas capital is a stock variable and is reflected in an insignificant coefficient. From the above table it is also clear that the Czech Republic has a significant higher shift upwards in the production frontier, i.e. both coefficients β_{10} and β_{11} are negative. This result is also found when we decompose the total factor productivity into its various components (see table 17). The last two coefficients imply that Poland and Belgium have a higher intercept relative to the Czech Republic.

[50]These are the summary statistics of the estimated scale elasticities. To make inferences about them (CRS, DRS or IRS) we need to test these various hypothesis.

7 References

1. Aigner, D., Lovell, K. and Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models, Journal of Econometrics, 6, pp 21-37.

2. Battese, G.E. and Coelli, T.J. (1988). Prediction of Firm Level Technical Efficiencies with a Generalised Frontier Production Funtion and Panel Data, Journal of Econometrics, 38, pp. 387-399.

3. Battese, G. E., Coelli, T.J. and Colby T.C. (1989) Estimation of Frontier Production Functions and the Efficiencies of Indian Farms Using Panel Data from ICRISAT's Village Level Studies, Journal of Quantitative Economics, 5(2), pp. 327-48.

4. Battese, G.E. and Coelli, T.J. (1992). Frontier Production Functions, Technical Efficiency and Panel Data: with Application to Paddy Farmers in India, Journal of Productivity Analysis, 3, pp. 153-169.

5. Battese, G.E. and Coelli T.J. (1995). A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data, Empirical Economics, 20, pp. 325-332.

6. Coelli, T. (1996). A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, Draft, pp. 33.

7. Coelli, T., Rao, D.S.P. and Battese, G.E. (1998). An Introduction to Efficiency and Production Analysis, Kluwer Academic Publishers, pp. 275.

8. Cornwell, C., Schmidt, P. and Sickles, R.C. (1990). Production Frontiers with Cross Sectional and Time Series Variation in Efficiency Levels, Journal of Econometrics, 46, pp. 185-200.

9. Demsetz, H. (1969). Information and Efficiency: Another Viewpoint, Journal of Law and Economics, 12, pp. 1-12.

10. European Commission (2001). Regular Report on the Czech Republic's Progress Towards Accession.

11. European Commission (2001). Regular Report on Poland's Progress Towards Accession.

12. Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (1993). The Measurement of Productive Efficiency, Technique and Applications, Oxford University Press, pp 119.

13. Funke, M. and Rahn, J. (2002). How Efficient is the East German Economy? An Exploration with Microdata, Economics of Transition, volume (10), pp 201-223.

14. González, E. and A. Cárcaba (2000), A Resource-Based Interpretation of Technical Efficiency Indexes, WP, Universidad de Oviedo, pp 17.

15. Green, W.H. (1990). A Gamma Distributed Stochastic Frontier Model, Journal of Econometrics, 46, pp. 141-164.

16. Hall, R.E. (1988). The Relation between Price and Marginal Cost in U.S. Industry, Journal of Political Economy, 96(5), pp. 921-947.

17. Horrace, W.C. and Schmidt, P. (1996). Confidence Statements for Efficiency Estimates from Stochastic Frontier Models, Journal of Productivity Analysis, 7(2-3), pp 257-282.

18. Hutchinson, J. (2002). Firm Size Distribution: From Gibrat to Sutton. Comparing Belgium and Slovenia using Firm Level Data, LICOS mimeo, pp. 33.

19. Jovanovic, B. (1982) Selection and the Evolution of Industry, Econometrica, 50, pp.649-670.

20. Kodde, D.A. and Palm, F.C. (1986). Wald Criteria for Jointly Testing Equality and Inequality Restrictions, Econometrica, 54, pp. 1243-1248.

21. Konings, J. and Repkine, A. (1997). How Efficient are Firms in Transition countries? Firm level evidence from Bulgaria and Romania, DP no 66/1997 LICOS, pp. 34.

22. Kumbhakar, S.C. (1990). Production Frontiers, Panel Data and Time Varying Technical Inefficiency, Journal of Econometrics, 46, 201-211.

23. Kumbhakar, S., Ghosh, S. and J. McGuckin, J.(1991). A generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms, Journal of Business and Economic Statistics, 9, pp.279-286

24. Kumbhakar, S.C and Lovell, C.A.K. (2000). Stochastic Frontier Analysis, Cambridge University Press.

25. Lee, Y.H. and Schmidt, P. (1993). A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency, in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (1993). The Measurement of Productive Efficiency, Technique and Applications, Oxford University Press, pp 119.

26. Leibenstein, H. (1966). Allocative Eﬃciency vs "X-ineﬃciency", American Economic Review, 56, pp. 392-415.

27. Levinsohn, J. and Petrin, A. (2001) Estimating Production Functions Using Inputs to Control for Undobservables, WP.

28. Meeusen, W. and van den Broek, J. (1977). Eﬃciency Estimation from Cobb-Douglas Production Functions wit Composed Error, International Economic Review, 18, pp. 435-444.

29. Osiewalski, J., Koop, G. and Steel M.F.J. (1997). A Stochastic Frontier Analysis of Output Level and Growth in Poland and Western Economies, TUCER, DP nr 9785, pp 21.

30. Piesse, J. and Thirtle, C. (2000). A Stochastic Frontier Approach to Firm Level Eﬃciency, Technological Change, and Productivity during the Early Transition in Hungary, Journal of Comparative Economics, 28, pp. 473-501.

31. Pitt, M.M. and Lee, M.F. (1981). The Measurement and Sources of Technical Ineﬃciency in the Indonesian Weaving Industry, Journal of Development Economics, 9:43-64.

32. Reifschneider, D. Stevenson, R. (1991). Systematic departures from the frontier: A framework for the analysis of ﬁrm ineﬃciency, International Economic Review, 32 (3), pp.715-723.

33. Roland, G. (2000). Transition and Economics: Politics, Markets and Firms, MIT Press, The MIT Press,

34. Schmidt, P. and Sickles, R.C. (1984). Production Functions and Panel Data, Journal of Business and Economic Statistics, 2, pp. 367-374.

35. Stevenson, R.E. (1980). Likelihood Functions for Generalised Stochastic Frontier Estimation, Journal of Econometrics, 13, pp. 57-66

36. Stigler, G. (1976). The Xistence of X-ineﬃciency, American Economic Review, 66, pp. 213-216

Table 1: The Structure of the Panel (1995-1999)

Number of years in panel	Number of firms				
	Poland	Czech Rep.	Belgium	Total	%
5	685	928	4315	5928	81
4	57	415	247	719	10
3	4	266	30	300	4
2	0	266	0	266	4
1	0	78	0	78	1
Total	746	1953	4592	7291	100

Table 2: Summary Statistics (x1000 euro)

Year	Poland			Czech Rep.			Belgium		
	output	empl	LP	output	empl	LP	output	empl	LP
1995	417	896	0.4651	149	471	0.3161	57	97	0.5911
1996	421	886	0.4746	148	419	0.3528	58	94	0.6219
1997	455	836	0.5448	154	375	0.4097	63	94	0.6695
1998	441	774	0.5697	147	334	0.4410	65	94	0.6862
1999	404	682	0.5930	157	341	0.4595	64	93	0.6906

LP: labour productivity (output/emp); Exchange rate used: EC, 2001

Table 3: Estimating model A and B for 1995 (N = 6628)

Dependent Variable	Model A		Model B	
$\log Y_i$	coefficient	standard error	coefficient	standard error
$\log K$	0.5511	0.0092	-0.6646	0.0621
$\log L$	0.2171	0.0098	0.4389	0.0531
$(\log K)^2$	-	-	0.0408	0.0036
$(\log L)^2$	-	-	-0.1441	0.0051
$\log K \log L$	-	-	0.0845	0.0070
constant	6.31	0.0987	12.58	0.2801
log-likelihood	-9884.5011		-9808.5463	

*: significant at 1%

Table 4a: Technical Efficiency Model A (1995)

TI (A) 1995	Poland	Czech Republic	Belgium	Overall
average	0.5937	0.5503	0.6453	0.6210
median	0.5904	0.5498	0.6521	0.6262
min	0.2662	0.0147	0.1053	0.0147
max	0.844	0.822	0.8405	0.844
Observations	746	1290	4592	6628

Table 4b: Technical Efficiency Model B (1995)

TI (B) 1995	Poland	Czech Republic	Belgium	Overall
average	0.6180	0.5433	0.6445	0.6218
median	0.6191	0.5479	0.6462	0.6821
min	0.1697	0.0054	0.1146	0.0054
max	0.8823	0.8327	0.8543	0.8823
Observations	746	1290	4592	6628

Table 5: Moments of Distribution of TE for Model A and B (1995)

Moments	Poland		Czech Republic		Belgium	
	A	B	A	B	A	B
mean	0.5937	0.61798	0.5503	0.5433	0.6453	0.6445
variance	0.0039	0.0059	0.0058	0.0095	0.0051	0.0054
kurtosis	4.1984	4.8277	6.6529	4.6408	6.4181	5.4824
skewness	0.1450	-0.4361	-0.6267	-0.5621	-0.9075	-0.5834

Table 6: Estimating model A and B for 1998 (N = 7068)

Dependent Variable	Model A		Model B	
$\log Y_i$	coefficient	standard error	coefficient	standard error
$\log K$	0.5497	0.0086	-0.6032	0.0513
$\log L$	0.2208	0.0092	0.2534	0.0484
$(\log K)^2$	-	-	0.0424	0.0032
$(\log L)^2$	-	-	-0.1093	0.0047
$\log K \log L$	-	-	0.0704	0.0065
constant	6.2758	0.1091	12.4303	0.2374
log likelihood	-10492.620		-9883.676	

*: significant at 1%

Table 7a: Technical Efficiency Model A (1998)

TI (A) 1998	Poland	Czech Republic	Belgium	Overall
average	0.6556	0.6254	0.6935	0.6722
median	0.6521	0.6262	0.6983	0.6746
min	0.3591	0.1245	0.2182	0.1245
max	0.8319	0.8262	0.8463	0.8463
Observations	735	1801	4532	7068

Table 7b: Technical Efficiency Model B (1998)

TI (B) 1998	Poland	Czech Republic	Belgium	Overall
average	0.6505	0.5958	0.6759	0.6528
median	0.6481	0.6026	0.6764	0.6570
min	0.2422	0.0628	0.2167	0.0628
max	0.8748	0.8278	0.8581	0.8748
Observations	735	1801	4532	7068

Table 8: Moments of Distribution of TE for Model A and B (1998)

Moments	Poland		Czech Republic		Belgium	
	A	B	A	B	A	B
mean	0.6556	0.6505	0.6254	0.5958	0.6935	0.6758
variance	0.0020	0.0038	0.0040	0.0063	0.0026	0.0036
kurtosis	7.3027	6.6178	12.1067	8.6479	5.3983	4.5218
skewness	-0.0891	-0.4917	-1.4570	-1.3133	-0.6529	-0.3858

Table 9: Comparing Technical Efficiency over Time (1995, 1998)

	Moments	Poland		Czech Republic		Belgium	
		A	B	A	B	A	B
1995	mean	0.5937	0.6798	0.5503	0.5433	0.6453	0.6445
	variance	0.0039	0.0059	0.0058	0.0095	0.0051	0.0054
1998	mean	0.6556	0.6505	0.6254	0.5958	0.6935	0.6758
	variance	0.0020	0.0038	0.0040	0.0063	0.0026	0.0036

Table 10: Results TE Model C.1: Only Constant Term ($\mu_{it} = \delta_0$)

Year	Statistic	Poland	Czech Rep.	Belgium
	Obs.	746	1291	4592
1995	Average	0.7334	0.6668	0.7473
	Minimum	0.1057	0.0000	0.0292
	Maximum	1	0.8783	0.8856
	Obs.	728	1445	4493
1996	Average	0.7322	0.6672	0.7522
	Minimum	0.1637	0.0002	0.2649
	Maximum	1	0.8735	0.8913
	Obs.	722	1640	4509
1997	Average	0.7355	0.6693	0.7533
	Minimum	0.3877	0.0019	0.0978
	Maximum	0.8767	0.8584	0.8891
	Obs.	735	1801	4532
1998	Average	0.7316	0.6681	0.7547
	Minimum	0.1058	0.0019	0.1004
	Maximum	1	0.8767	0.8928
	Obs.	734	1531	4527
1999	Average	0.7231	0.6692	0.7534
	Minimum	0.1431	0.0134	0.094
	Maximum	0.8827	0.8705	0.8937

Table 11: The Production Function Coefficients of model C.2

Dependent Variable	Model C	
$\log Y_i$	coefficient	standard error
$\log K$	-0.6115	0.0244
$\log L$	0.2934	0.0230
$(\log K)^2$	0.0412	0.0015
$(\log L)^2$	-0.1214	0.0021
$\log K \log L$	0.0765	0.0030
constant	12.32	0.1076
log likelihood	-46808.932	

*: significant at 1%

Tabel 12: Results TE Model C.2: Year Dummies ($\mu_{it} = \delta_0 + \sum \delta_l D_t$)

Year	Poland		Czech Rep.		Belgium		Overall	
	TI	Obs.	TI	Obs.	TI	Obs.	TI	Obs.
1995	0.7127	746	0.6400	1291	0.7295	4592	0.7100	6629
1996	0.7379	728	0.6759	1445	0.7581	4493	0.7380	6666
1997	0.7478	722	0.6870	1640	0.7649	4509	0.7445	6871
1998	0.7352	735	0.6738	1801	0.7584	4532	0.7344	7068
1999	0.7406	734	0.6923	1531	0.7684	4527	0.7483	6792
1995-1999	0.7347	3665	0.6750	7708	0.7557	22653	0.7352	34026

Table 13: Results TE Model C.3: Sectoral Dummies ($\theta_{i,t} = \delta_0 + \sum_{j=16}^{37} \delta_j SECT_{i,j} + \varepsilon_{i,t}$)

Sector	Poland	Czech Rep.	Belgium	Overall
Food Products	0.7265	0.6743	0.7372	0.7216
Tobacco Products	0.8908	0.8752	0.8966*	0.8930
Textiles	0.6891	0.6457	0.7276	0.7076
Wearing Apparel	0.7475	0.6678	0.7510	0.7375
Leather	0.7998	0.7829	0.8180	0.8001
Wood	0.6747	0.5790	0.6950	0.6722
Pulp & Paper	0.6684	0.6242	0.7145	0.6945
Publishing & Printing	0.7383	0.6735	0.7580	0.7475
Coke & Petroleum	0.6948	0.6480	0.6480	0.6460
Chemicals	0.7212	0.6631	0.7613	0.7383
Rubber & Plastic	0.7062	0.6326	0.7400	0.7110
Other Non-Metallic Mineral	0.6081	0.5214	0.6580	0.6208
Basic Metals	0.7095	0.6793	0.7428	0.7173
Fabricated Metal	0.6838	0.5900	0.6915	0.6701
Machinery Equipment	0.6566	0.6055	0.7166	0.6670
Office, Machinery & Computers	0.8870	0.8745	0.8920	0.8880
Electrical Machinery	0.6752	0.6181	0.7230	0.6806
RTv & Communication	0.6700	0.6564	0.7518	0.7066
Medical	0.5666	0.4841	0.6020	0.5628
Motor Vehicles	0.7281	0.6556	0.7492	0.7181
Other Transport Equipment	0.6687	0.5561	0.6884	0.6394
Furniture	0.7262	0.6538	0.7393	0.7235
Recycling	0.8603	0.8145	0.8618	0.8608
Average	0.6978	0.6249	0.7248	0.6993

*: Tobacco Products (16) sets the benchmark (TE = 1)

Table 15: Kruskal-Wallis Test for Distribution of TE between Countries (1995)

Distributions of TE ...	Subsectors (NACE 2-digit level)
differ significantly	17, 20, 21, 22, 24, 25, 26, 27, 28, 29, 31, 34, 35
do not differ significantly	19, 23, 30, 31, 33, 37

Table 14: Subsectoral Technical Efficiency Model D, E

Subsector	Poland		Czech Rep.		Belgium		Overall	
	1995	1998	1995	1998	1995	1998	1995	1998
17	0.6089	-	0.5316	-	0.6383	-	0.6163	-
19*	0.4998	0.2434	0.3444	0.2227	0.5238	0.5393	0.4549	0.3479
20	0.6336	0.5891	0.4905	0.3893	0.6480	0.6246	0.6241	0.5696
21	0.5291	0.5441	0.3597	0.4053	0.5896	0.6169	0.5474	0.5701
22	0.6808	0.6093	0.6071	0.4994	0.7182	0.6999	0.7060	0.6682
23*	0.6655	0.5966	0.1587	0.3742	0.3913	0.5123	0.4325	0.5139
24	0.6365	0.5816	0.5763	0.4856	0.6826	0.6647	0.6599	0.6182
25	0.5826	-	0.3892	-	0.6445	-	0.4918	-
26	0.6384	-	0.5427	-	0.6636	-	0.6358	-
27	0.5861	0.6586	0.5338	0.6240	0.6321	0.7161	0.5951	0.6753
28	0.6688	0.6593	0.5382	0.5334	0.7059	0.6831	0.6786	0.6445
29	0.5972	-	0.5490	-	0.6432	-	0.6068	-
30*	0.5996	0.5576	0.2408	0.0992	0.5063	0.5426	0.4906	0.4798
31	0.4965	0.5827	0.4502	0.5069	0.6039	0.6589	0.5423	0.5912
33*	0.5574	-	NO	-	0.7237	-	0.6165	-
34	0.4933	0.6130	0.2847	0.5038	0.5069	0.6564	0.4446	0.6014
35	0.1856	0.6704	0.1301	0.5872	0.3207	0.7087	0.2329	0.6559
37	0.5120	-	NO	-	0.5729	-	0.5692	-

(*) Observations less than 50, - : OLS is better or no TI effects, NO: no observations, See Appendix 3 for NACE classification.

Table 16a: Results of the Simple PA-Model

	Poland	Czech Republic
β	0.1020	0.0305
Standard Error	0.0093	0.0044
R^2	0.0417	0.0128

Table 16b: Results of the Extended PA-model

	β_1	β_2	β_3	β_4	R^2
Poland	0.1863	0.1898	0.0788	0.0890	0.0996
	(0.0193)	(0.0203)	(0.0210)	(0.0196)	
Czech Republic	0.1091	0.0680	0.0428	0.0268	0.0487
	(0.0124)	(0.0131)	(0.0130)	(0.01263)	

standard errors are in parantheses

Table 17: Average TFP Decomposition

Country	TFP4	T4	TE4	scale
Poland	9.66	7.68	0.92	1.03
	100	79.49	9.48	10.70
Czech Rep.	10.15	8.58	1.07	0.46
	100	84.60	10.56	4.58
Belgium	10.79	6.20	3.94	0.64
	100	57.49	36.51	5.95

Averaged over 1996-1999, shares in italic

Table 18: Average-, Maximum- and Minimum Elasticity of Scale

Country		1995	1996	1997	1998	1999
Poland	average	0.7833	0.7845	0.7851	0.7839	0.7824
	max	0.9264	0.9257	0.9331	0.9490	0.9457
	min	0.5091	0.4995	0.5054	0.4511	0.5197
Czech Rep.	average	0.7734	0.7681	0.7651	0.7610	0.7619
	max	0.9750	0.9731	0.9718	0.9534	0.9517
	min	0.5067	0.3887	0.4105	0.4600	0.3887
Belgium	average	0.8355	0.8365	0.8340	0.8309	0.8286
	max	1.1068	1.1021	1.043	1.0976	1.1054
	min	0.5601	0.5459	0.5300	0.5355	0.5171

Figure 1: Distribution of TE per country in 1995

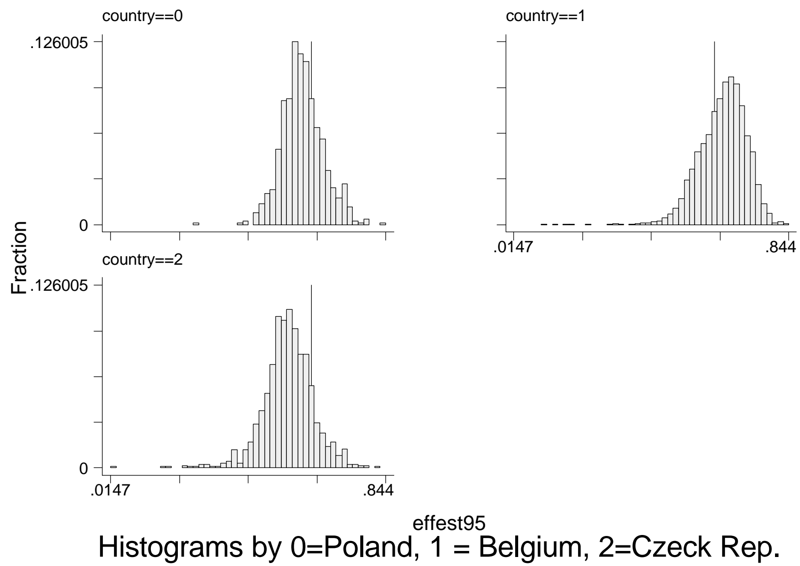

Figure 2: Distribution of TE per country in 1998