Aalto University
School of Business

Estimating technical efficiency in Finnish industry: A stochastic frontier approach

Economics

Master's thesis

Topi Hokkanen

2014

Department of Economics
Aalto University
School of Business

Acknowledgements

The topic of this thesis crystallized during my time at the Government Institute for Economic Research (VATT), where I was employed as a research assistant in the summer of 2013. I am grateful to all the VATT personnel I have had the privilege of working with – for a budding economist such as myself, your expertise and assistance has been of great value. In particular I wish to thank research director Anni Huhtala, for providing me with both guidance and allowing me to use VATT resources and databases and Elina Berghäll, whose project originally steered me towards the field of efficiency analysis. The advice of prof. Timo Kuosmanen (Aalto) and Mika Kortelainen has also been invaluable regarding the nonparametric models.

Abstract

OBJECTIVES

The objective of this thesis is to estimate the technical efficiency of Finnish industry in various sectors using stochastic frontier models. Furthermore this thesis provides a comparison between parametric and nonparametric estimation techniques utilized in the efficiency analysis literature. We estimate the technical efficiency for each sector separately, both as a cross-section and using panel data techniques. We discuss the various extensions to the models considered and identify further topics of research based on our estimation results.

METHODOLOGY AND DATA

This thesis utilizes stochastic frontier estimation techniques to estimate technical efficiency for various sectors of industry. Production frontiers are estimated both parametrically and nonparametrically. The two main approaches to frontier estimation utilized in the thesis are Stochastic Frontier Analysis (SFA) and Stochastic Nonparametric Envelopment of Data (StoNED). The dataset we utilize is a cross-sectional panel dataset of companies operating in various sectors of the Finnish industry. Technical efficiencies are estimated with both cross-sectional and panel data methods, with various specifications.

CONCLUSIONS

We find that the mean efficiency in most sectors is quite high in most models. The estimated efficiencies are sensitive to the model choice. Parametric and nonparametric estimates are found to be similar and of the same magnitudes for the majority of sectors considered. We find significant variability in technical efficiencies over time, particularly in models where firm-specific heterogeneity is controlled.

KEYWORDS: Technical efficiency, Industry, Econometrics, Stochastic Frontier Analysis, Stochastic Nonparametric Envelopment of Data, SFA, StoNED

Table of contents

1. Introduction

"Efficiency is doing better what is already being done" – Peter Drucker

Efficiency analysis has been an important topic in economic research, spanning back over 50 years. Furthermore the relevance of economic efficiency is exacerbated by the scarceness of resources, which is a basic concept in economic theory. The field of efficiency analysis is a diverse one, encompassing many academic disciplines, with applications in agriculture, industrial organization, engineering sciences, and economics. In our discipline, the utilization of efficiency analysis techniques has covered virtually every industry spanning from such diverse sectors as health care or sports to specific industries such as airlines and even broader topics, such as macroeconomic analysis (Fried, Lovell & Schmidt, 2008 p. 16). Within the field of econometrics, efficiency analysis remains a very popular topic – indeed, as noted by Kuosmanen, Johnson & Saastamoinen (2014), 13 out of the 100 most cited articles in the *Journal of Econometrics* are on topics relating to efficiency analysis.

The standard approach in microeconomics is usually to develop the theory of the firm by way of production functions and profit-maximizing behavior. The theoretical and econometric literature developed within this classic framework is vast and contains many papers which are nowadays considered as classic references in the field. However, prior to the developments in efficiency analysis the prevailing approach was to implicitly assume firms' operating at full efficiency, attributing observed variation from estimated production functions to simply statistical error[1]. In contrast to this, an interesting point to consider is the possibility of *inefficiency*. Can firms fail to obtain maximal output from their inputs, or perhaps might their costs be greater than required by their production level due to firm-specific factors? In other words, is there inefficiency in production and if so, can its determinants be identified? These are some of the basic questions that the field of efficiency analysis addresses.

[1] This is to say, the econometric approach of estimating production functions with OLS.

The econometric tools available for efficiency analysis are varied, and the field is constantly developing with regards to new techniques and methods. For the purposes of this thesis, we have selected two contrasting views to efficiency analysis, representing both parametric and nonparametric approaches. Our models include the classic, parametric Stochastic Frontier Analysis, developed by Aigner, Lovell & Schmidt (1977) and a recent nonparametric approach, called StoNED (Stochastic Nonparametric Envelopment of Data) proposed by Kuosmanen & Kortelainen (2012). These models are utilized to estimate the technical efficiency in Finnish industry for various sectors during the time period 2003-2009. The estimation results of the models are discussed, with a comparison of the parametric models both vis-à-vis each other and against the nonparametric models. The choice of stochastic frontier methods for the analysis was done so as to best capture the inefficiencies in different sectors, allowing for random disturbances in the data due to e.g. exogenous factors or measurement error.

1.1. The Finnish industrial sector

Traditionally, industry has been a significant driver of Finnish GDP, with especially the forestry and paper industries being important before the growth of the ICT sector (Kontulainen & Spolander, 2010). In recent decades though, the contribution of the industrial sector has decreased with a particularly sharp decline occurring during the financial crisis. However, current events[2] in the Finnish economy have again increased the interest in these traditional sectors as potential engines for future economic growth. The time period considered in this thesis consists of the years 2003 to 2009, containing some very adverse market conditions such as the financial crisis starting from 2008 onwards. Figure 1 below plots the output volume index for the aggregated industrial sector alongside the sector-specific indices. The aggregated volume index indicates that growth was somewhat stable up until 2008, with a sharp decline in 2008-2009. The disaggregated figures show that aside from the impact of the financial crisis which had a significant common impact, the individual sectors considered in this thesis did not develop in unison throughout the time period under study. A decreasing trend in output can be observed for metal refining, starting from 2004 onwards, while other industries seemed to have fared decently well up until the shock of 2008.

[2] i.e. the diminishing growth contribution of the ICT sector, the 'post-Nokia' –economy, etc.

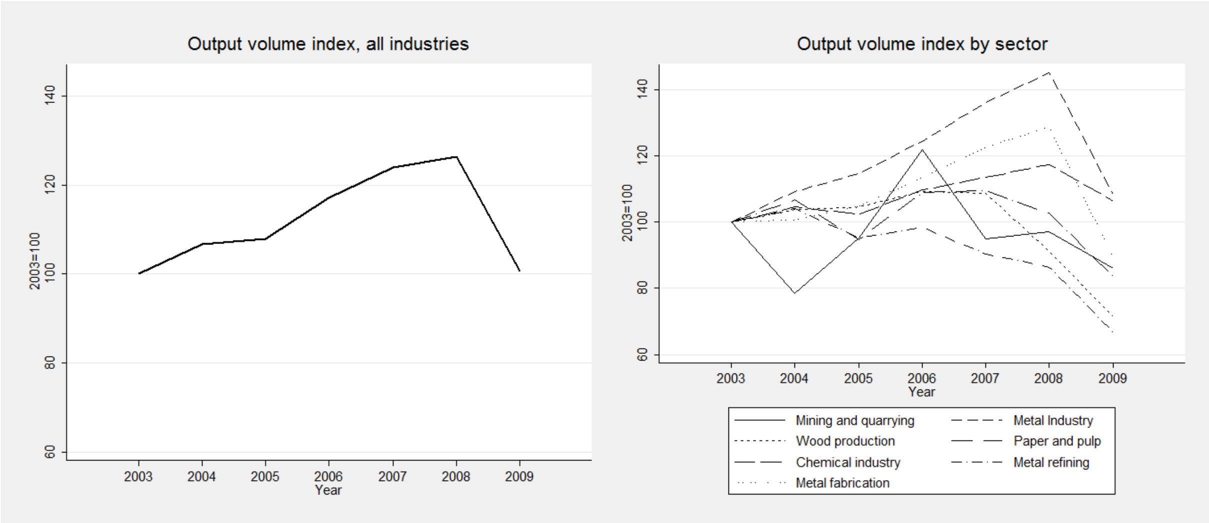


Figure 1: Aggregated and individual output volume indices, 2003-2009, 2003=100 (Statistics Finland)

However, even with a declining factor share of the economy as a whole, industry is still projected to remain quite significant sector for the Finnish economy (Honkatukia, Ahokas & Simola, 2014) in the future. For this reason it is prudent to take a closer look at the industrial sectors to ascertain whether the traditional sectors operate efficiently. The aim of this thesis is to tackle this problem by using various stochastic frontier methods. Furthermore, by estimating our models by sector, we provide a middle-ground view between the highly specific analyses of within-sector branches and country-level macroeconomic studies. Our main results from the econometric analysis are that most sectors in Finnish heavy industry are operating at relatively high mean efficiency, especially given the turbulent market conditions prevailing in our study period. We find substantial changes in technical efficiencies over time, especially when firm-specific heterogeneity is taken into account. Moreover, parametric and nonparametric efficiency estimates are found to be relatively similar for most sectors considered. The estimated technical efficiencies imply that for the chosen aggregation scale, parametric models that control for firm-specific heterogeneity seem to be best in capturing both the output elasticities and time-varying technical efficiency. Industry sectors that are estimated as the most efficient in the majority of the models are the wood and wood products sector, the quarrying sector and the metal refining and fabrication sectors.

1.2. Thesis structure

The rest of this thesis is organized as follows. The second chapter motivates the theoretical foundations for key concepts in efficiency analysis, emphasizing the link between the empirical methods and production theory. We also provide a more detailed review of the literature, going over the most relevant literature produced in the field[3] with regards to our chosen topic. Our emphasis in the third chapter is on the methods utilized in this thesis, introducing the reader to the stochastic frontier models that we utilize in the subsequent analysis. The chapter will discuss the benefits and drawbacks of each model formulation and conclude with an exposition of the possible extensions to these techniques. The fourth chapter is dedicated to summarizing the dataset used in this thesis, providing the relevant descriptive statistics and overview of our specific data with regards to the chosen variables of interest. The chapter also discusses our choice of input and output variables, with additional details provided on the calculation of certain input variables. The fifth chapter presents the estimation results for the models considered, with discussion on the estimated efficiencies in each sector and relevant parameters. We also provide a comparison of estimated technical efficiencies between the different types of models, and furthermore inspect the links between the efficiencies using Spearman rank correlations. The sixth and final chapter presents our conclusions from the estimated models and discusses some of the outstanding issues in efficiency analysis, suggesting some promising avenues for further research based on our results. Optimization algorithms, more detailed distributions on the estimated technical efficiencies and other additional visualizations are included in the appendix, where necessary.

[3] As mentioned previously, efficiency analysis literature is a vast body of work. We aim to produce a focused 'snapshot' review of the literature, given our choice of topic.

2. Analytical framework

In this chapter we provide the motivation behind efficiency analysis and introduce the relevant production theory and microeconomic theory that underpins the models used in this thesis. We will also provide the reader with a review of the efficiency analysis literature, with emphasis on studies utilizing stochastic frontier analysis. This section builds on the textbook treatises on efficiency analysis, namely Fried, Lovell & Schmidt (2008), Kumbhakar & Lovell (2000) and Coelli, Prasada Rao, O'Donnell & Battese (2005).

2.1. Production theory

In the context of efficiency analysis, production is considered as a process where producers utilize inputs (denoted by the input vector $\boldsymbol{x}$) to produce some output (denoted by $\boldsymbol{y}$). The producers transform the inputs to outputs using some production technology, which can be described either with set-theoretic concepts or the familiar approach of the production function. We begin by introducing the output and input sets alongside the technology set for some given production technology. The technology set T is defined as the set of feasible production schemes $\{\boldsymbol{x}, \boldsymbol{y}\}$, which can be produced with certain production technology specific to the production unit under observation, formally:

(1) $$T = \{(y, x): x \; can \; produce \; y\}$$

Intuitively, the boundary of this set is the *production frontier*, which relates maximal producible output for any given input vector. The input sets of the same production technology are then defined as the sets of input vectors that are feasible for each element of the output vector $\boldsymbol{y}$.

(2) $$L(y) = \{x: (y, x) \in T\}$$

Similarly, the boundary of this set forms the input isoquants for the production technology. Finally, the output set is defined as the set feasible outputs, for every possible input vector x.

(3) $$P(x) = \{y: (y, x) \in T\}$$

Likewise to the above sets, the boundary of the output set defines the output isoquants for a given output y. Figure 2 illustrates graphically the sets defined above.

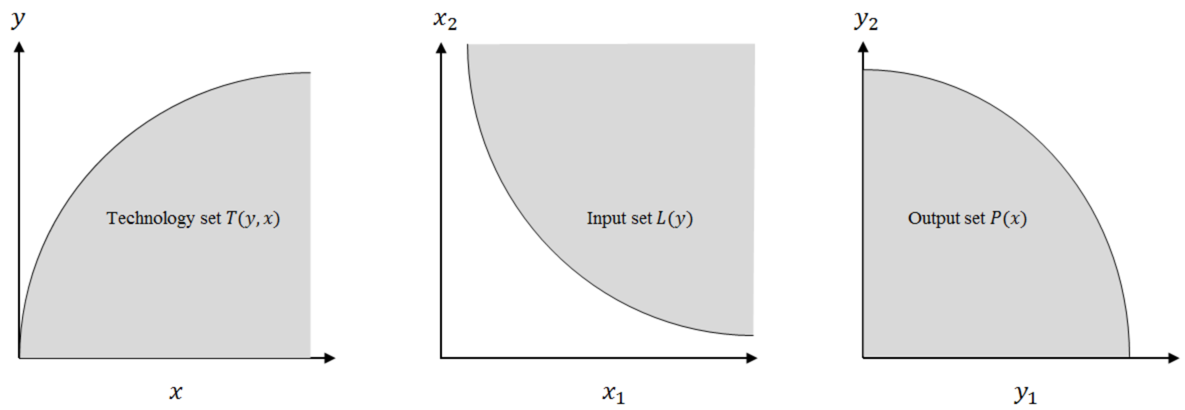


Figure 2: Illustrations of output, input and technology sets

The technology and output sets are compact, while the input set is only closed, bounded below by the input isoquant. The general properties of these sets are given in Coelli, et al. (2005), and are omitted here for sake of brevity, our purpose being only to shed some light on the forthcoming definitions of technical efficiency.

In contrast to the set representation, production can also be characterized by the familiar production function as a parametric representation of the production process for a given producer. This representation however requires that the production process is either single-output, or alternatively that the output vector can be aggregated to a composite output vector, using some optimal weights (Kumbhakar & Lovell, 2000). The function itself gives us the relation between inputs and outputs, with specific properties that depend on the chosen functional form. As described in Figure 3 below, the production function allows for consideration of multiple classes of inputs and outputs.

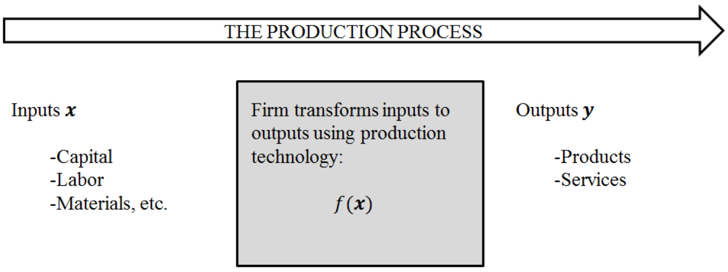


Figure 3: The production process and production function

The functional forms for production functions within the field of efficiency analysis range from the very simple to rather complex. Table 1 below summarizes some of the most utilized forms of production functions that have been utilized in applied work in the field.

Linear	$y = \beta_0 + \sum_{n=1}^{N} \beta_n x_n$
Cobb-Douglas	$y = \beta_0 \prod_{n=1}^{N} x_n^{\beta_n}$
Quadratic	$y = \beta_0 + \sum_{n=1}^{N} \beta_n x_n + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \beta_{nm} x_n x_m$
Normalised quadratic	$y = \beta_0 + \sum_{n=1}^{N-1} \beta_n \left(\frac{x_n}{x_N} \right) + \frac{1}{2} \sum_{n=1}^{N-1} \sum_{m=1}^{N-1} \beta_{nm} \left(\frac{x_n}{x_N} \right) \left(\frac{x_m}{x_N} \right)$
Translog	$y = \exp\left(\beta_0 + \sum_{n=1}^{N} \beta_n \ln x_n + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \beta_{nm} \ln x_n \ln x_m \right)$
Generalised Leontief	$y = \sum_{n=1}^{N} \sum_{m=1}^{N} \beta_{nm} \left(x_n x_m \right)^{1/2}$
Constant Elasticity of Substitution (CES)	$y = \beta_0 \left(\sum_{n=1}^{N} \beta_n x_n^{\gamma} \right)^{1/\gamma}$

Table 1: Typical production functions in efficiency analysis (Coelli, et al. 2005)

In this thesis, we utilize one of the most classic functional forms for our empirical analysis, namely the simple log-linear Cobb-Douglas production function, originally developed by Cobb and Douglas (1928). The Cobb-Douglas production function exhibits some useful properties for the purposes of our estimations, such as the handy interpretation of estimated parameters and the simple inclusion of Hicks-neutral technical progress in our estimated production frontiers.

2.2. The basic axioms of production

For production functions, we define the following four basic axioms of production. Production functions that fulfil these axioms can be said to be well-behaved in the context of efficiency analysis, i.e. if they hold, then the estimated frontier is said to be axiomatic. The axioms follow very intuitive economic logic, as will be seen below in our presentation. The basic axioms are (Kumbhakar & Lovell, 2000):

1. Nonnegativity: This axiom states that the production function transforms inputs to outputs in such a way that the resulting output is a finite, nonnegative real number. Formally, $f\colon \mathbb{R}^n \to \mathbb{R}^+$.
2. Weak essentiality: This axiom states that strictly positive production is not possible with at least one strictly positive input, i.e. if $f(\boldsymbol{x}) > 0$, then the input vector $\boldsymbol{x}$ must have at least one strictly positive element.
3. Monotonicity: This axiom states that production is nondecreasing in inputs, stated formally as follows: if $\boldsymbol{x_1} > \boldsymbol{x_2}$ then $f(\boldsymbol{x_1}) > f(\boldsymbol{x_2})$. This axiom also implies that for continuously differentiable production functions, marginal products are nonnegative for all inputs.
4. Concavity: This axiom states that the production function must be concave in inputs, i.e. for input vectors $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$, $f\big(\lambda \boldsymbol{x_i} + (1-\lambda)\boldsymbol{x_j}\big) \geq \lambda \boldsymbol{x_i} + (1-\lambda)\boldsymbol{x_j}$. This property implies the law of diminishing marginal productivity.

These axioms guarantee that production occurs in an economically feasible way, with most of the axioms having quite a colloquial logic behind them, e.g. ‘you cannot get something for nothing’ (weak essentiality) etc. Still, we must keep in mind that even these basic axioms can, and indeed may be violated in situations where more complex production processes are considered. For example, the monotonicity property may easily be violated if there is *input congestion*, which may occur in real life situations. Such violations may also occur if the definition of the input vector is extended to include so-called *bad inputs* such as emissions or other pollutants.

2.3. Defining technical efficiency

Given the production technology, production sets and input isoquants defined previously, we are now ready to define the concept of technical efficiency. Originating from the work of Farrell (1957) and Debreu (1951) the simple, informal definition of technical efficiency is that a production unit is technically efficient if for a given input-output vector either no constriction of input quantities or no expansion of output quantity is feasible[4] maintaining a given level of output or inputs, respectively. In other words, the production unit is allocating its inputs in a way that its production is *maximal*, i.e. the production coincides with the true production frontier. Formally, the definition of technical efficiency we utilize in this thesis is the following. An input-output vector $y \in P(x)$ is *technically efficient* if and only if $y^* \notin P(x)$ for $y^* \geq y$ (Coelli et al., 2005)[5]. Thus for a given input, a technically efficient producer is defined as one whose production coincides with the production frontier, with no feasible expansion in output possible with the same input level. This is the output-oriented technical efficiency, often also called the Farrell or Debreu-Farrell measure of technical efficiency, for obvious reasons. A stricter definition was given by Koopmans (1951), but as noted by Kumbhakar & Lovell (2000, p. 44) the majority of applied work utilizes the definition(s) given here.

Thus, for a production function $f(\boldsymbol{x})$, the above definitions imply an output technical efficiency measure for producer i of the form:

(5)
$$TE_i = \frac{y_i}{f(\boldsymbol{x_i})}$$

which is simply the ratio of observed output (y_i) to frontier output. This ratio lies between zero and one, with increasing values implying higher technical efficiency. Things become

[4] We consider here both input-oriented and output-oriented measures.

[5] Note here the remarkable similarity of this definition of technical efficiency to the general definition of efficient production, given e.g. in Mas-Colell, Whinston & Green (1995, p.150).

more complicated when we consider truly stochastic production frontiers, where exogenous, stochastic variation is considered alongside inefficiency, but our simple treatise here should be enough to introduce the reader to the key concept of technical efficiency. Following the simple measure of technical efficiency in (5), the production frontier for producer i can be defined as:

(6) $$y_i = f(\boldsymbol{x_i}) \times TE_i$$

Note here that the measures presented above exclude the possibility of stochastic noise in the observed outputs. When we introduce the stochastic frontier models in Chapter 3, we will see how the efficiency measures are defined in this case. Figure 4 below illustrates the concept of technical efficiency in a deterministic, single-output/single-input production process. For our analysis, the output-oriented efficiency measures are sufficient, since our main goal is the estimation of sector-specific efficiencies. However, the output-oriented measure can further be subdivided into revenue efficiency, which examines the shortfall of revenues from optimal revenues and allocative efficiency, which complements the former by considering efficiency with regard to the optimal input mix. These decompositions, while interesting, impose additional requirements for the data in the estimation process, and for our purposes will not be pursued.

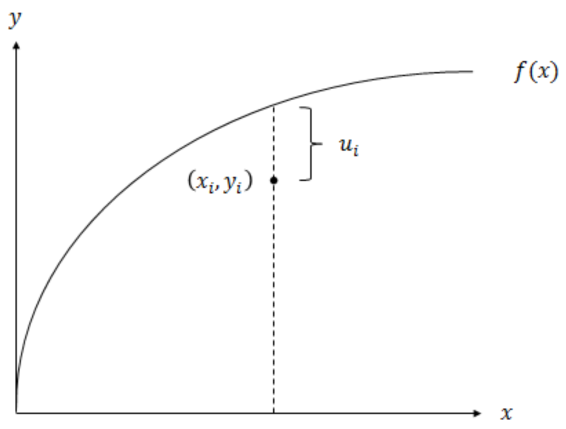


Figure 4: Technical efficiency with a deterministic production frontier

This simple figure illustrates the concept of technical efficiency in a deterministic frontier context. The Debreu-Farrell measure is obtained by projecting the observation to the production frontier and forming the ratio of observed production to maximal production. An input-oriented measure could be constructed in a similar fashion, however now the projection in this graph would be done along the input axis and the input-oriented measure would then reflect the possible constriction of inputs rather than expansion of output.

2.4. Review of current literature

Since its inception in the 1970s, stochastic frontier methods have been utilized in efficiency analysis for a variety of topics. Empirical applications have been published for virtually every sector of industry, from schooling to services and specific industries to even more entertaining topics, such as sports. For the Finnish economy, particular emphasis has been given to micro-level studies, focusing tightly on a specific branch within an industrial sector. Fields of interest have included topics such as welfare services, electricity distribution, education and agriculture and farming.

The received literature with regards to our chosen topic is scarce, with few precedents of similar types of studies using similar data. The probable reason for this is the fact that our chosen level of aggregation is quite broad, involving the estimation of sector-specific production frontiers. The clear majority of research focuses on very specific sectors of the economy, such as the aforementioned electricity distributors, or the production of specific products or crops in agriculture (see Battese, 1991). On the other end of the 'aggregation scale' these techniques have been utilized in macroeconomic contexts, using country-level data to assess e.g. the sources of productivity growth across countries (see Färe, Grosskopf, Norris & Zhang, 1994). The work contained herein falls somewhere in between these two categories, with subdivisions considered between different sectors of the economy but not as much as to zoom in on a specific branch of industry within these sectors. As of yet, we are unaware of prior studies using Finnish industrial data with similar scope.

One quite similar work to ours was presented by Fritsch & Stephan (2004) who analyze the German industry using deterministic frontier models. The data utilized consists of an exhaustive panel of German manufacturing industries for the years 1992-2002. The authors estimate mean technical efficiencies in different sectors, finding significant differences in technical efficiencies across industries. However, the authors use a restrictive, deterministic frontier model[6] which, like DEA requires by definition that a single firm is operating at full (100%) efficiency. Secondly, the authors' view of the distribution of technical efficiencies is somewhat perplexing, as they assume that the technical efficiencies themselves should display the same skewness we expect for the composed error term. As pointed out by Wang & Schmidt (2009), this is generally not the case in standard stochastic frontier models. Other cross-sectoral studies include Green & Mayes (1991), who study the technical efficiency of the manufacturing industry in the United Kingdom. The authors find that a variety of sectors are found to be technically inefficient to some degree. However, the authors note that for approximately a third of the industries under consideration, the composed residual terms were skewed in the wrong direction, thus making the estimation of technical efficiency in these sectors impossible. The authors also find significant heterogeneity within sectors, which can be considered as a relevant issue considering our topic choice, as our results also support the notion of strong within-sector heterogeneity in national industries.

Representing the macroeconomic scope of efficiency analysis literature, Koop (2001) links inefficiency to the business cycle, studying the factors influencing productivity and technical efficiency in six industries for 11 OECD countries during the years 1970-1988 with stochastic frontier methods. The results show notably uniform estimated technical efficiencies within countries, with very little variability between industrial sectors. All industrial sectors in the three Nordic countries included in the data (Sweden, Denmark and Norway, respectively) are estimated to be operating at 90% efficiency or above. By using contextual variables, the author links growth phases to the development of inefficiency, concluding that the industries under study seemed to be more efficient in times of expansion and decrease their efficiency during downturns. This result may be related to the labor hoarding phenomenon, where firms even under adverse conditions are unwilling to reduce their input levels – if we consider this

[6] indeed as observed by Cornwell & Schmidt (2008) the deterministic frontier is a tool that must be used very cautiously.

alongside the definition of technical efficiency[7], the results are somewhat intuitive if firms are 'slow' to adjust their labor inputs when faced with external shocks.

Due to their high regulation, the electricity distribution market has attracted significant interest in the context of domestic efficiency analysis studies. Kopsakangas-Savolainen & Svento (2011) analyze the cost efficiency of distributors, using many of the models also utilized in this thesis. Among the models the authors' utilize are the true random-effects and true fixed effects models by Greene (2005a, 2005b). The authors estimate mean efficiencies upwards of 80-90%, depending on the model choice. Firm-specific heterogeneity is found within the sector and is found to influence the estimated technical efficiencies, if not accounted for in the model specification. Previously, the same topic was studied by Honkatukia & Sulamaa (1999) who favor nonparametric methods over the fully stochastic frontier models. The authors reach similar conclusions, namely that distributors are estimated to be operating at a high mean efficiency for years 1996-1998.

With regards to our chosen sectors, Hyytiäinen, Viitanen and Mutanen (2010) estimate the technical efficiency of Finnish sawmills, which fall under our subsector of wood and wood products (Sector 20). The authors estimate the Pitt & Lee (1981) time-invariant panel model alongside the Battese & Coelli (1992) time-decay model, similar to our work. The sample size is decidedly small, with only 72 observations for the time period 2000-2007. In contrast to our results, the authors find little evidence for time-varying efficiency. The sawmills are estimated to be operating at a mean efficiency of 81%, which is very close to our results, where mean efficiency for the entire sector was estimated at 76% in the TRE model.

The nonparametric StoNED –model has been utilized in much the same context for Finnish industries as the parametric methods. The original application of this model was in the context of regulation of electricity distributors (Kuosmanen, Kortelainen, Kultti, Pursiainen, Saastamoinen & Sipiläinen, 2010) where the model was used to construct the benchmark cost frontier and determine cost reduction targets for the regulated firms for the Finnish energy regulation authority. Other empirical applications of this model parallel the uses of parametric

[7] i.e. producing maximal output with a given input vector.

stochastic frontier methods with similar research focus in a specific branch of industry. The published works include efficiency analysis of retail banking, electricity distribution and agriculture. Kuosmanen (2012) uses the StoNED –model to analyze the cost efficiency of Finnish electricity distributors. The model is extended to include contextual variables serving as a proxy for the operating environment. This particular subsector is estimated as highly efficient, with mean efficiency estimated as high as 92%. The operating environment is estimated to have a significant effect on total costs of a firm, with urban operators facing much steeper costs. Eskelinen & Kuosmanen (2013) utilize the model to analyze the intertemporal efficiency of retail bank branches, utilizing sales indicators as contextual variables. Large branches are found to be, on average more efficient than their smaller contemporaries. The technical efficiencies of different sales teams are found to vary significantly over the time period. Due to the novel nature of the StoNED –model, published work is yet not yet as expansive as with the earlier nonparametric and parametric frontier methods (DEA and SFA). However, even the few applications introduced here show that the model is quite flexible and particularly suitable for studying the determinants of inefficiency by way of contextual variables.

In many ways, this thesis can be considered as an intellectual descendant to the pioneering work of Aigner, Lovell & Schmidt (1977) and Meeusen & van Den Broeck. (1977) who also estimate sector-specific technical efficiencies in their originating papers on SFA. In effect, we follow in their footsteps, however now with proprietary data and modern methodology. Thus far we have introduced a simple efficiency analysis framework, within which we can analyze the efficiency of different firms by comparing their output to maximal (frontier) output. However captivating our exposition may be, two important issues still remain open. The first consideration is the question of how to estimate the theoretical frontier against which firms' performance should be measured. The second and equally important is the inclusion of stochastic noise in observed output due to measurement error or exogenous factors outside of the production units' control. These are the two questions we aim to answer in the next section.

3. Methodology

In this chapter, we will review the models utilized in this thesis. The two frontier approaches chosen for this analysis are Stochastic Frontier Analysis (SFA), and Stochastic Nonparametric Envelopment of Data (StoNED). The frontier methods are introduced and the econometric estimation techniques used for estimating the frontiers are discussed. The chapter will begin with a short overview of Data Envelopment Analysis (DEA), as this is the first widely utilized efficiency analysis technique in economic literature, after which we shift our emphasis to the stochastic frontier methods. Addressing DEA provides both an understanding of the differences between SFA and StoNED and also provides some insight to the development of efficiency analysis techniques in econometric analysis.

3.1. Data Envelopment Analysis (DEA)

Data Envelopment Analysis, originating from Farrell's seminal paper (1957) and formulated to its current form by Charnes, Cooper and Rhodes (1978) is a nonparametric efficiency analysis technique, which was developed to extend efficiency analysis from an index number-based approach to a more analytical framework. The idea behind DEA is to take a set of individual decision making units (DMUs), such as firms, and construct an efficient production frontier against which each DMU can then be compared to obtain the efficiency of the unit in question. The frontier is constructed using a mathematical programming algorithm, which assumes a general production function satisfying the axioms in chapter 2.2, without a specific functional form and approximates this nonparametrically, by way of enveloping hyperplanes. Charnes et al. (1978) describe this frontier as a production possibility surface, where produced output is maximal w.r.t. the input level.

The main desirable attributes of DEA are the basic assumptions the technique relies on, which are the following (Charnes, et al., 1978 & Banker, Charnes & Cooper, 1984); 1) convexity of input and output sets, 2) monotonicity/free disposability, 3) weak essentiality and 4) minimum extrapolation. Formally stated in our notation, these are:

1. for $\boldsymbol{x}, \boldsymbol{y} \in T$ & $\boldsymbol{x_1}, \boldsymbol{y_1} \in T$, then $[\lambda(\boldsymbol{x}, \boldsymbol{y}) + (1-\lambda)(\boldsymbol{x_1}, \boldsymbol{y_1})] \in T$ for $\lambda \in [0,1]$ meaning that convex combinations of x and y belong to the production possibility set T
2. if $\boldsymbol{x}, \boldsymbol{y} \in T$, and $\boldsymbol{x_1} \geq \boldsymbol{x}, (\boldsymbol{x_1}, \boldsymbol{y}) \in T$ and if $\boldsymbol{y_1} \leq \boldsymbol{y}, (\boldsymbol{x}, \boldsymbol{y_1}) \in T$. The logic behind this assumption is that it is always feasible to produce the same output with higher input quantity (output is non-decreasing in inputs). Similarly for output it's always feasible to produce less output using the same input quantities.
3. $\forall \boldsymbol{y} > \boldsymbol{0}, (\boldsymbol{0}, \boldsymbol{y}) \notin T$. Weak essentiality states that no strictly positive production can occur with zero inputs[8].
4. T is the intersection of all sets satisfying the above conditions. This assumption states that the estimated frontier has the property of being as close to the production set as possible, while enveloping all observations.

These assumptions mirror the basic axioms of production we defined in part 2.2. Thus a major benefit of DEA is the preservation of production axioms in estimating the efficient frontier (Thanassoulis, Portela & Despić, 2008). Compared to parametric methods, DEA is considered to be more flexible in allowing for multi-input/multi-output production processes, due to the nonparametric estimation and the exclusion of a specific production function (Bauer, 1990). DEA frontiers can be constructed in various ways, and can easily handle differing specifications of scale efficiency – the two most common models that are used in research are the Constant Returns to Scale –model and the Variable Returns to Scale –model introduced by Banker et al, (1984). The DEA model can also be specified either as input-oriented or output-oriented (cf. chapter 2). In practice, the choice of orientation is usually straightforward – e.g. in the production model it is natural to use output-oriented models and in a cost frontier model, the choice of input-orientation is usually utilized. As stated, the above characteristics make DEA a suitably flexible approach to efficiency analysis, which is one possible reason why the method is utilized to this day in modern efficiency analysis.

[8] Here we must observe that Charnes et al. (1978) state this condition with a weak inequality – however, in standard microeconomics producing nothing is always feasible, so we modify the condition slightly. The logic behind this statement still applies even when the inequality is strengthened.

However, the flexibility of DEA comes at a rather substantial cost, as the method has some significant drawbacks in the context of econometrics. The first major consideration in DEA is the fact that the method relies on identifying efficient DMUs, thus requiring that at least one DMU is operating at full efficiency (Coelli, et al., 2005). This may be an unrealistically strict assumption, and one which given data that may contain stochastic noise we should not expect. The method is also highly sensitive to outliers in the data. To see why, consider the previous fact alongside the assumption of minimal extrapolation – for an extreme outlier it is likely that the estimated frontier (and thus efficiency estimates) would be biased as the frontier must envelop all observations, including the outlier. Thus the estimated frontier may 'hang' on the outlier, with the outlying observation estimated fully efficient while other observations receive biased efficiency estimates.

The second consideration is the exclusion of noise in the data considered – however we specify the 'classic' DEA -model; we must assume that deviations from the efficient frontier are due to inefficiency alone. This assumption may be incorrect due to many reasons – for instance, if the choice of inputs and outputs omits an important variable which belongs in the process or if the data itself is subject to some random shocks or events which fall beyond the range of the DEA model. Since DEA estimates efficiency by benchmarking the DMU to the frontier (Thanassoulis et al., 2008), noisy data leads to a biased estimate of the frontier and thus also of the final inefficiencies of DMUs. Thus while the nonparametric frontier estimation can be seen as a benefit of the DEA approach (Greene, 1997), the exclusion of noise generally makes DEA ill-suited to applications where the data may contain such elements. Thorough surveys of the subsequent developments in DEA after Charnes, et al. (1978) are provided by Cook & Seiford (2009) and Thanassoulis et al. (2008).

3.2. Stochastic Frontier Analysis (SFA)

Stochastic Frontier Analysis (from here on referred to as SFA), originally developed by Aigner et al. (1977) and Meeusen & van Den Broeck (1977) serves as a counterpoint to earlier DEA methods of estimating technical efficiency. In contrast to DEA, the assumption of observed deviations from the frontier as being simply a result of inefficiency of the evaluated DMU is relaxed. The first important distinction between these two somewhat related techniques is the inclusion of statistical noise in the observed deviation from the estimated frontier. This allows for the use of efficiency analysis also in situations where we cannot with certainty assume that the 'output gap' between observed production and the optimal production is free of stochastic elements. The second distinction between these methods is the assumption of some clearly defined production technology – i.e. a parametric[9] production function. In contrast to DEA, SFA relies heavily on this assumption of the production function to be utilized in the analysis of the data, while the DEA method avoids defining an explicit production function. This leads to a different interpretation for the results from these two methods – DEA estimates the convex hull of the technology set as the minimal enveloping frontier, while SFA estimates the parameters of the production function itself.

These fundamental assumptions underpinning these two types of models make the estimation results of DEA and SFA difficult to compare. For example, in DEA, the estimated parameters for the enveloping hyperplanes serve as point estimates for the marginal productivities of the defined inputs for a given observation, but not as estimators of the parameters of the true production function. However, by construction, similar parameters in SFA serve both as estimators of the marginal productivities and the true production function parameters (Kuosmanen, et al., 2014). When these models are implemented in practice, for example by policy makers and regulators, these fundamental differences sometimes go unacknowledged (Greene, 2008). A commissioned study in energy regulation by Kuosmanen, Saastamoinen & Sipiläinen (2013) found that in many countries not only are these differences unaccounted for, they are sometimes ignored completely.

[9] a model is referred to as parametric, if all its parameters are contained in finite parameter spaces, and nonparametric if they are contained in infinite parameter spaces (Kuosmanen et al., 2014).

The stochastic frontier model retains some flexibility in allowing the actual frontier under estimation to be specified in various ways, usually according to the specific objectives of the researcher. In addition to production frontiers, the model lends itself readily to the estimation of cost frontiers (Schmidt & Lovell, 1979) and profit frontiers (Kumbhakar, 1987). These formulations differ from the simple production model by their choice of output and inputs and also the impact and interpretation of inefficiency. For the most part though, the inner workings of these extensions are almost identical to the production side models we discuss in this thesis.

3.2.1. Deterministic frontier models

Before the simultaneous innovation of Aigner, et al and Meeusen, et al (1977), earlier work on frontier estimation did not include the stochastic noise element in the model, which is why they are often referred to in literature as *deterministic frontier models*. Early efforts in frontier estimation begun with the work done by Aigner & Chu (1968), who formulated their production model in terms of a log-linear Cobb-Douglas production function with a multiplicative inefficiency term. The production function is specified as:

$$Y_i = A\prod_{i=1}^{n} X_i^{\alpha_i} \times U \tag{9}$$

Where U is defined as a random, nonnegative inefficiency term between zero and one capturing all factors that may influence inefficiency, such as the quality differences of inputs or exogenous shocks. The authors propose estimating the production frontier by way of mathematical programming, and argue that the frontier should envelop the data from above. As the frontier is constructed by way of mathematical optimization, statistical inference is dubious, as noted by the authors themselves (p. 827). Taking logarithms, the model proposed by the authors is:

$$\min_{\alpha,\beta,\varepsilon} \sum \varepsilon_i^2 \text{ , } s.t.$$

$$\ln y_i = \beta_0 + \sum \beta_i \ln x_i + \varepsilon_i$$

$$\varepsilon_i \leq 0$$

This approach has the property that the optimized frontier will bound the data from above, as the sign of the disturbances is restricted to be negative. This approach is related to DEA in the way that the shortfall of a specific observation from the frontier is interpreted entirely as inefficiency. Some other early applications of the basic ideology which would later develop into the SFA we utilize today are methods focusing on Ordinary Least Squares (OLS) estimation of the production frontiers, with different ways of 'correcting' the frontier to envelop the observed data. A benefit of OLS is its robustness to non-normality of the error term (Greene, 2003) and so it can be used to consistently estimate the model parameters, however special treatment is required for the intercept term. The two OLS approaches are Corrected OLS (COLS), developed by Winsten (1957) and Greene (1980) and Modified OLS (MOLS) by Richmond (1974). Both of these methods rely on OLS to estimate the production function parameters, but differ in their treatment of the OLS residuals ε_i. The COLS procedure shifts the frontier up by the amount of the largest residual, thus generating a frontier that truly envelops the data. As an example, using our notation, at the first stage a (log-linear) production model such as the following would be estimated by OLS:

(10) $$\ln y_i = \beta_0 + \sum \beta_i \ln x_i + \varepsilon_i$$

In the second stage the residuals would be utilized to shift the frontier to envelop the data. If we denote the maximum residual as $\varepsilon_{max} = \max(\varepsilon_i)$, the COLS intercept would be estimated as

(11) $$\beta_{COLS} = \beta_0 + \varepsilon_{max}$$

This shifts the frontier up[10] so that the observation coinciding with the largest positive residual will be on the frontier, with other observations under the frontier. Efficiency analysis in this approach can be viewed as a relative comparison[11], with the frontier observation being defined as being fully efficient and other observations receiving efficiency scores relative to the fully efficient observation. Also notable is the fact that the COLS frontier does not necessarily bound the data from above as closely as possible (Kumbhakar & Lovell, 2000) as the corrected frontier is parallel to the OLS frontier by definition. The second OLS-based approach, MOLS (Richmond, 1974) requires a distributional assumption about the

[10] The new COLS residuals are obtained by $-\varepsilon_{COLS} = \varepsilon_{max} - \varepsilon_i$

[11] Often called Comparison with the best in panel data contexts (cf. Kim & Schmidt, 2000)

inefficiency term u_i, such as exponential or half-normal. Since the mean of the OLS residuals is zero by construction, it offers little information, but the displacement is constant so that a higher order central moment of the residuals may be utilized to yield a consistent estimator of the mean of the inefficiency $E(u_i)$. After obtaining this estimate the frontier is then shifted in the same fashion as in COLS, but now using $E(u_i)$.as the shift parameter, i.e.

(12) $$\beta_{MOLS} = \beta_0 + E(u_i)$$

In contrast to COLS, this shifted frontier may fail to envelop the data entirely, possibly leading to observations that are above the frontier. As the model is still deterministic, the interpretation of this is that the observed firm's technical efficiency is greater than unity (Kumbhakar & Lovell, 2000, Greene, 2008), a notion for which it is hard to find justification – after all, by definition technical efficiency must be restricted to lie between zero and one. In literature, the interpretation for these peculiar properties has usually been that these methods estimate the "average frontier" and not the "best-practice" –frontier, which recognizes the underlying frontier property of the model in question (Greene, 2008, p. 106). These early approaches contrast with the basic OLS –based estimation, which relates all disturbances to the frontier to purely stochastic error, while these formulations (like DEA before them) attribute residuals to the deterministic component, i.e. inefficiency. We illustrate these antecedents of SFA in Figure 5 below, where the differences in these methods are observable. The COLS frontier truly envelops the data, while some observations lie above the MOLS frontier.

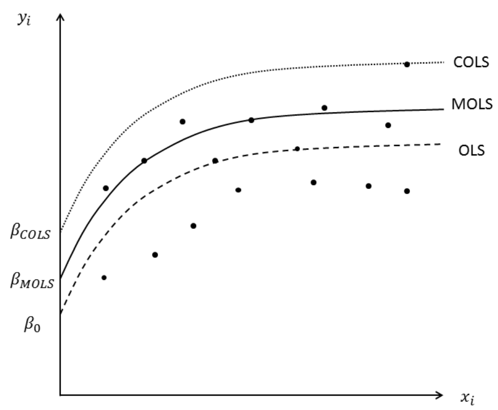


Figure 5:OLS-based production frontiers

While these models still see limited use today, they are less widespread in modern efficiency analysis, having been replaced by the fully stochastic models we consider next. In some ways, these early models can be thought to be the ‘worst of both worlds’ in layman’s terms, as they include the restrictive parametric specification of the production function, but still are unable to include stochastic error in the observations themselves. These models can thus be viewed as following the DEA lineage of assuming disturbances to be purely the result of inefficiency. A natural extension of this is, therefore, the addition of a stochastic element to the estimated frontier alongside inefficiency, addressing the issue that observed production may fluctuate due to other sources than strict inefficiency which is precisely the major innovation in the stochastic frontier model we will now introduce.

3.2.2. The cross-sectional SFA model

The stochastic frontier model (Aigner, et al, 1977 & Meeusen, et al, 1977) begins with an assumption of some production function $f(\boldsymbol{x})$, which relates the input vector to a single, non-negative output vector $\boldsymbol{y}$. In applied work, a clear majority of the research utilizes the familiar first-degree flexible[12] Cobb-Douglas or second-degree flexible transcendental logarithmic production function (abbreviated translog), as noted by Greene (2008, p.98).

Oftentimes, the choice of production technology is down to researcher preference, with studies on the correct functional forms for analysis being quite scarce. This avenue of research links closely to the broader discussion of justifiable functional forms in econometric analysis, and will not be pursued for the purposes of this thesis. A noted problem with the more flexible production functions is the fact that these specifications may yield estimates that do not satisfy the basic axioms of production mentioned earlier in chapter 2.2 (Kumbhakar & Lovell, 2000). More specifically, a more flexible functional form may create problems in the econometric estimation as well, if the function needs additional constraints on the estimated parameters to satisfy, e.g. monotonicity and concavity. A thorough survey of this particular issue was presented by Gong and Sickles (1992), who investigate the sensitivity of SFA

[12] A functional form is called first-order flexible if it provides a first-order differential approximation to an arbitrary function at a single point (Coelli, et al., 2005).

results to the choice of functional form and conclude that the choice of correct[13] production technology is imperative to obtain unbiased results.

Following Aigner, et al (1977) the frontier production function is now specified as:

$$\ln y_i = \beta_0 + \sum \beta_i \ln x_i + \underbrace{v_i - u_i}_{\varepsilon_i} \tag{13}$$

$$\varepsilon_i = v_i - u_i \tag{14}$$

where the production function is the familiar log-linear Cobb-Douglas form, y_i denotes the observed output and x_i denotes the chosen inputs for the production process. The *composed error term* (epsilon) includes both firm-specific inefficiency in u_i and stochastic noise in v_i. The inefficiency component captures the deterministic inefficiency of the unit in question, whereas the noise component represents stochastic events influencing production in a random fashion[14]. In addition, it is assumed that both error components are distributed independently of each other and the regressors. The true innovation in SFA comes from the specification of the error term in (14), which is now a convolution of two random variables, representing inefficiency and noise. When compared to standard OLS, which captures stochastic elements by a single *i.i.d.* error term, the composed error term in SFA allows us to disentangle noise from inefficiency. The distributional assumptions originally proposed by Aigner, et al. (1977) for the error terms are:

$$\begin{gathered} v_i \sim N(0, \sigma_v^2) \\ u_i \sim N^+(0, \sigma_u^2) \end{gathered} \tag{15}$$

where the $N^+()$ distribution refers to a half-normal distribution. These distributional assumptions alongside the production technology constitute the *normal-half-normal* stochastic frontier model. A parallel specification was derived by Meeusen, et al. (1977), who utilize the exponential distribution for the inefficiency term. For the normal-exponential model, the distribution of u_i is:

[13] Correct in the sense that the production technology should most accurately describe the true production function.
[14] Aigner et al. (1977) propose things such as luck, climate or machine performance.

$$u_i \sim \exp(\theta), \sigma_u = \frac{1}{\theta} \tag{16}$$

Figure 6 below plots the distributions for the composed error term epsilon, with the exponential model plotted on the right, and the half-normal model plotted on the left with varying parameter values. We see that the distributions are skewed to the left, with negative means and modes.

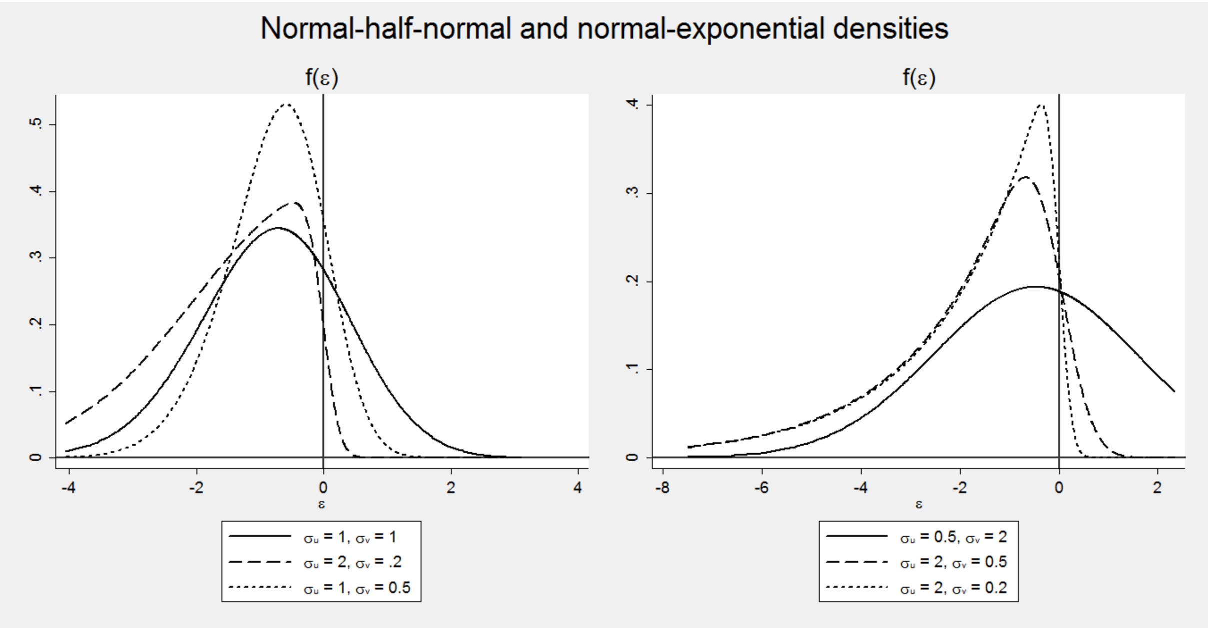


Figure 6: Density functions for ε (reproduced from Kumbhakar & Lovell, 2000)

The model is usually estimated via Maximum Likelihood (ML), as the log-likelihood of both specifications is relatively straightforward to derive and maximize (cf. Aigner et al., Meeusen et al., 1977). Furthermore, using ML in place of OLS allows unbiased estimation of all model parameters in the first stage, avoiding the pitfalls of COLS and MOLS. Although the normal-half-normal and normal-exponential variants are the most ubiquitous, various other distributions may also be considered for the inefficiency term, with the normal-truncated normal model, originally suggested by Stevenson (1980) having been used in research quite often. The difference in specification concerns again only the inefficiency terms distribution, now allowing it to follow a truncated normal distribution with a non-zero mean (equation 17 below). This variant of the standard model nests the normal-half-normal variant (when the truncation is at zero). For illustrative purposes, Figure 7 below plots some density functions for the composed error term, with different truncation means.

(17) $$u_i \sim N^{+}(\mu, \sigma_u^2)$$

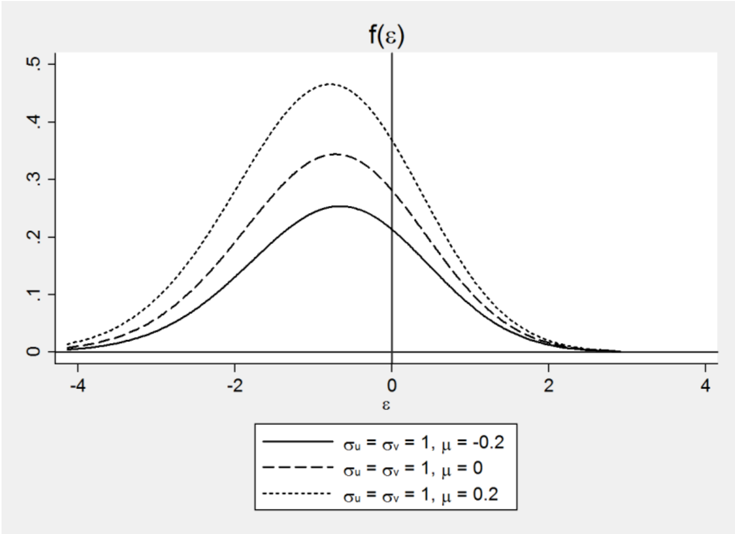


Figure 7: Truncated-normal densities for ε (reproduced from Kumbhakar & Lovell, 2000)

The truncated normal distribution has the benefit of allowing the mean of inefficiency to be non-zero, which may be preferred, for instance in situations when it is tenable that firms operate in an environment where significant inefficiency might be present. This variant also allows for additional parametrization of μ as a function of selected explanatory variables, as will be discussed later. While it is also possible to estimate the model with even more flexible distributions that nest most of the commonly utilized distributions such as the Pearson distribution or the Gamma distribution suggested by Greene (1980), the models become increasingly complex to estimate via ML. One desirable characteristic of the stochastic frontier models is, fortunately, that the model parameters and relative efficiency rankings themselves are quite robust to the choice of inefficiency distribution (Coelli et al., 2005). Still, it must be noted that the estimated inefficiencies themselves may change when the distributional assumptions are modified. Regardless of how we specify the inefficiency terms distribution, the common logic behind all these formulations is that the inefficiencies should be distributed such that the mean is either at zero or sufficiently close to zero. The reasoning here being that with firms operating in competitive markets, we should expect inefficiency to be quite low, with most of the probability mass being on the 'efficient' part of the distribution.

Mean inefficiency in the model is estimated as the first central moment[15] of the composed error term ε, as in Aigner et al. (1977):

(18) $$E(\varepsilon) = E(u) = \sqrt{\frac{2}{\pi}}\sigma_u$$

Using (18), the mean technical efficiency[16] in the sample is defined in a similar fashion than in (6) in Chapter 2.3 as a multiplicative term of the production frontier:

(19) $$TE = \exp(-E(u))$$

The formulation of the efficiency in this manner is done as to accurately describe the deterministic inefficiency– indeed, if we were to construct a simple efficiency measure using the ratio of observed production to (the now stochastic) production frontier, we would get an efficiency measure which is partially stochastic. As we are focused on productive inefficiency, such a measure is clearly not feasible (Aigner, et al. 1977, p. 25).

The returns to scale assumptions of the production function are addressed in the literature generally in one of two ways: the researcher may either require CRS to hold *a priori*, in which case an additional constraint is specified for the estimated coefficients in (13), or conversely no constraints are defined before estimation and postestimation hypothesis tests are performed to see whether CRS holds with required statistical significance. Generally postestimation procedures allow for testing the CRS hypothesis, i.e.

$$H_o: \sum_{i=1}^{n} \beta_i = 1$$

[15] For a rigorous discussion on the central moments of a random variable, we refer the reader to Johnson, Kotz & Balakrishnan (1994).

[16] To recap, we consider a producer technically efficient if they obtain maximal output with their allocated inputs – when production falls short of optimal, the producer is defined as technically inefficient, to the degree determined by their specific measure.

With flexible functional forms, such as the translog production function, obtaining this linear homogeneity however comes at the cost of requiring the specification of many additional constraints. This may affect the convergence of ML and also the estimated models. Moreover, while a flexible production function may be desirable in order to obtain a ‘best fit’ of the frontier to the data, they often have some unfortunate consequences in the econometric context. For example, many of the more flexible functional forms are no longer monotonic or globally concave (Greene, 2008) which poses a significant problem if we wish to maintain links with the standard microeconomic theory.

3.2.3. Estimating individual inefficiency

The original model by Aigner et al. (1977) provides an estimate of the mean inefficiency over all observations, but fails to estimate the firm-specific inefficiency. In applied work the aim is still often to further estimate the individual inefficiencies. In cross-sectional data however, the individual inefficiency cannot be directly estimated. This happens because the residuals from the cross-sectional model are by definition, convoluted and contain both the noise and inefficiency effects. Commenting on this dilemma, Kuosmanen et al. (2014) plainly state that it is simply impossible to estimate a firm-specific effect from a single observation subject to noise. A justifiable solution to this problem was suggested by Jondrow, Lovell, Materov & Schmidt (1982), who argue that while $E(u_i)$ remains elusive to estimate, the residuals ε_i still contain information about the inefficiency, such that a conditional estimator is feasible. To this effect they derive the conditional distribution of $u_i|\varepsilon_i$ and suggest either the mean or mode of this distribution to be utilized as the firm-specific inefficiency. The authors show that the conditional inefficiency is distributed as $u_i|\varepsilon_i \sim N^+(\mu^*, \sigma_*^2)$, truncated below at zero. Given this result, the conditional mean estimator for inefficiency is defined as:

(20)
$$E(u_i|\varepsilon_i) = \mu_* + \sigma_* \left[\frac{\phi\left(-\frac{\mu_*}{\sigma_*}\right)}{1 - \Phi\left(-\frac{\mu_*}{\sigma_*}\right)} \right]$$

where ϕ is the standard normal density function, Φ the standard cumulative normal density function and μ_* and σ_* defined as

(21)
$$\mu_* = -\varepsilon_i \times \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$$

(22)
$$\sigma_* = \frac{\sigma_u \sigma_v}{\sqrt{(\sigma_u^2 + \sigma_v^2)}}$$

with the technical efficiency estimate for each observation then obtained via

(23)
$$TE_i = \exp[-E(u_i|\varepsilon_i)]$$

which, given the support of u_i, we see is bounded above by unity and asymptotically below by zero. We will refer to this estimator as the JLMS estimator of technical efficiency, as per the authors. In cross-sectional data, the estimator is unbiased, but still inconsistent (Kumbhakar & Lovell, 2000, Greene, 2008), as the variation attributed to the distribution is independent of i. In their original work on the subject, the authors' recognize this property of the estimator, attributing this to the fact that the residuals contain only imperfect information about the inefficiency term (Jondrow et al., 1982, p. 235). The other point estimator utilized in the literature for conditional inefficiency was later derived by Battese & Coelli (1988), using a slightly differing structure, given in the equation below.

(24)
$$TE_i = E[\exp(-u_i)|\varepsilon_i] = \left[\frac{1 - \Phi\left(\frac{\sigma_* - \mu_*}{\sigma_*}\right)}{1 - \phi\left(-\frac{\mu_*}{\sigma_*}\right)}\right] \exp\left[-\mu_* + \frac{1}{2}\sigma_*^2\right]$$

These two formulations are not necessarily equivalent, due to the fact that

$$\exp[-E(u_i|\varepsilon_i)] \neq E[\exp(-u_i)|\varepsilon_i]$$

These point estimators can be viewed as representing the mean of the conditional distribution where both u_i and ε_i are drawn from and as such are inconsistent – under repeated sampling both estimators would still not equal u_i, the deep parameter value (Greene, 2008). However, this characteristic of the estimator is well-known, and as stated by Kuosmanen et al. (2014) and Cornwell & Schmidt (2008) the JLMS estimator is currently the best that can be achieved with cross-sectional data. Together these two conditional inefficiency estimators are part and parcel of modern efficiency analysis, with especially the JLMS estimator being the tool of choice for most current SFA applications. The JLMS estimator is also employed later on in the StoNED –model to estimate firm-specific inefficiency. The estimator can be applied with relative ease also in panel data models, where the error components receive an additional subscript t, indicating the time period. The JLMS estimators for the standard[17] models are presented in Greene (1997) and Kumbhakar & Lovell (2000).

The cross-sectional stochastic frontier model introduced above relies on some rather strong distributional assumptions about the error terms; however they must be maintained (in some form) so as to be able to derive both the likelihood functions and the conditional inefficiency estimators. Expanding the stochastic frontier model to a panel data context addresses some of these issues, allowing for a more accurate estimation of technical efficiency. The next chapter will discuss the stochastic frontier model in a panel data context and introduce the panel data models that will be utilized in our analysis.

3.2.4. Panel data models

As mentioned in the previous section, access to panel data is desirable in stochastic frontier analysis, as it allows for the explicit modeling of most of the major issues that arise in cross-sectional models. As outlined by Kumbhakar & Lovell (2000) and Greene (2008), the main issues to which panel data provides a solution are:

i. Relaxing the strong distributional assumptions for the convoluted error term (u and v).
ii. The allowance of correlation between the inefficiency term and inputs.
iii. The computation of conditional firm-specific inefficiency with the JLMS point estimator.

Stochastic frontier models for panel data are classified into two broad categories: *fixed-effects models* and *random-effects models*, with both allowing for estimation of either time-varying efficiency or time-invariant efficiency. We begin our exposition by assuming a panel of

[17] i.e. the normal-half-normal, normal-exponential and normal-truncated-normal models.

producers, indexed by $i = 1, \dots, I$ observed on periods $t = 1, \dots, T$. The models introduced here assume a balanced panel, where each producer is observed throughout the panel period, but our chosen data set is an unbalanced panel. Fortunately our chosen estimation methods are directly applicable even when the panel is unbalanced. For further details on the structure of our data set, we refer the reader to the next chapter where the properties of the data are elucidated in detail. We will now present the two types of panel data models we will utilize in this thesis, which differ in their handling of the temporal variation of inefficiency. We consider both a time-invariant model, where efficiency effects are assumed to be static w.r.t. time and a time-varying efficiency model, where efficiency effects have temporal variation.

3.2.4.1. Fixed or Random-effects?

Although the textbook treatises on microeconometrics (e.g. Cameron & Trivedi, 2005, Greene 2003) tend to prefer fixed-effects models over random-effects models, in efficiency analysis some of the assumptions behind fixed-effects estimation may be untenable. The first consideration (and one which we must address due to our data) is selectivity bias in unbalanced panels. In context, this property means that fixed-effects models may be used in unbalanced panels if there is no correlation between observing a unit in the panel and the error term (Hayashi, 2000). However, in SFA where the error term includes the inefficiency term it is ambivalent whether we can make such an assumption – after all, should we not observe firms with higher inefficiency to 'drop out' of our panels due to competition. This problem is obviously exacerbated in long panels (Cornwell & Schmidt, 2008), where the selection bias may be considerable. According to Koski & Pajarinen (2013), the exit rates of companies are significantly related to their productivity, with declining productivity increasing the probability of exit. While productivity in itself is not a complete proxy for technical efficiency, this result provides some basis for our preference of random-effects estimation. Nurmi (2004) estimates the survival rates for Finnish manufacturing plants for years 1989-1997, including the sectors analyzed in this thesis. We find that for our sectors the survival rates are for the most part quite high (p. 38). Thus this problem is unlikely to cause significant bias in our estimations, particularly since our panel is relatively short (cf. Chapter 4).

The second and more significant point to consider is the fact that fixed-effects models may capture other time-invariant effects which are not included in the model and by construction attribute this to inefficiency (Greene, 2005b), again diluting the precision of efficiency estimation giving support to using a random-effects model in this context. Greene (2008) observes the differences in the models poignantly, stating that fixed-effects models in essence revert back to the deterministic frontier specification, while having the benefit of robustness and the random-effects models carry with them the assumption of inefficiency being uncorrelated to the inputs themselves. However, the assumption is likely to be a benign one (p.229) and a much more important factor is the temporal modeling of inefficiency. In contrast to Greene, Cornwell and Schmidt (2008) assert that the primary consideration is whether the individual effects are allowed to correlate with the inputs and choose the estimation method accordingly.

Kumbhakar & Lovell (2000, p.106) state that for relatively short panels with a significant number of observations a random-effects based approach is preferable to the fixed-effects treatment, with ML methods being an efficient choice if the distributional assumptions are satisfied. Kim & Schmidt (2000) compare various specifications of panel methods in the stochastic frontier context, and observe that fixed-effects estimation performs generally poorly (p.92), with ML with distributional assumptions for the error components being generally well-suited for the task of efficiency estimation. Given the composition of our data, we feel that the choice of random-effects models over fixed-effects is justifiable, based on the arguments presented in the surveys and literature referenced above. We consider two 'classic' random-effects models and an extended random-effects model which given its structure (introduced shortly) offers a more powerful version of the classic models.

3.2.4.2. Time-invariant technical efficiency

We consider the following random-effects model, with time-invariant technical efficiency. Following Pitt & Lee (1981) the stochastic production function in equation (13) is now rewritten as

(25) $$\ln y_{it} = \beta_0 + \sum_{n=1}^{N} \beta_n \ln x_{nit} + v_{it} - u_i$$

where the distributional assumptions for the error terms follow the standard-normal-half-normal approach, i.e.

(26) $$v_{it} \sim N(0, \sigma_v^2)$$
$$u_i \sim N^+(0, \sigma_u^2)$$

and furthermore it is assumed that u_i and v_{it} are distributed independently of each other and the inputs. The log-likelihood function can then be derived and maximized with respect to the parameters to obtain estimates for the betas and the error components. This model carries with it the (likely stringent) assumption of time-invariant inefficiency as the inefficiency term is not allowed to change over time. After obtaining estimates for all parameters, the JLMS estimator can be utilized to obtain estimates of producer-specific efficiencies with the adjustment of replacing ε_{it} with $\bar{\varepsilon}_i$ and σ^2 with σ^2/T (Schmidt & Kim, 2000).

3.2.4.3. Time-varying technical efficiency

We consider two specifications of time-varying efficiency models, the time-varying decay model proposed by Battese & Coelli (1992) and the true random effects model by Greene (2005a, 2005b). The time-decaying efficiency model has the benefit of relative simplicity, when considered to alternative formulations proposed by Lee and Schmidt (1993) and Cornwell, Schmidt and Sickles (1990), as it only requires the estimation of a single additional parameter, which characterizes the temporal pattern of inefficiency. For the time-decay model of Battese & Coelli (1992), the production function is rewritten as:

(27) $$\ln y_{it} = \beta_0 + \sum_{n=1}^{N} \beta_n \ln x_{nit} + v_{it} - u_{it}$$

(28) $$v_{it} \sim N(0, \sigma_v^2)$$

and, as before the error components are assumed to be distributed independently of each other and the regressors. The model variant proposed by the authors is the normal-truncated-normal model, however, the time-varying inefficiency term is now parametrized with a time trend[18], defined as:

(29) $$\begin{gathered} u_{it} = \beta(t) \times u_i \\ u_i \sim N^+(\mu, \sigma_u^2) \end{gathered}$$

(30) $$\beta(t) = \exp\big(-\eta(t-T)\big)$$

To see the temporal variability of efficiency allowed in this model, consider the behavior of the trending function. Technical efficiency is defined as before in equation (23), with the slight modification that u_{it} is utilized in place of u_i. We see now that

(31) $$\dot{\beta}(t) = -\eta \underbrace{\exp\big(-\eta(t-T)\big)}_{>0}$$

(32) $$\dot{TE_i} = -\dot{\beta}(t) \underbrace{\exp(-\beta(t)u_i)}_{>0}$$

Thus the estimated parameter η determines the pattern of 'time-decay' in the model. If $\eta > 0$ then technical efficiency is increasing over time (at an increasing rate) and conversely if $\eta < 0$ then the technical efficiency is decreasing over time. The model nests the Pitt & Lee (1981) time-invariant model, as when $\eta = \mu = 0$, there is no temporal variation in efficiency, and the inefficiency component follows a half-normal distribution. Thus a suitable test for the time-invariance assumption is the value of parameter η – if the estimated parameter is both significant and non-zero, then technical efficiency can be said to vary over time, with the above interpretation.

The second time-varying model we consider is the True Random Effects (TRE) model, proposed by Greene (2005a, 2005b). This model was developed to extend the functionality of

[18] Another formulation following the same logic is Kumbhakar (1990), who suggests a trend function of the type $\beta(t) = [1 + \exp(\gamma t + \delta t^2)]^{-1}$

the classical approaches and to address the common problems associated with these models. The first issue tackled by this model is the (still) rigid structure for inefficiency effects over time employed by the previous model. It is unclear whether a simple trend, which is estimated to be the same for all firms in the sample, provides an accurate description of time-varying efficiency. The second issue scrutinized are the effects of unobservable heterogeneity across the observed firms, which in other models are by construction included in the inefficiency term – thus the observation that previous models may, in fact, be "… *picking up heterogeneity in addition to or even instead of inefficiency"* (p. 7). Greene's formulation of the TRE model offers a solution to both of these issues, and provides a valuable contribution to the econometrics literature on efficiency analysis. The production function in the model is written as:

(33) $$\ln y_{it} = \alpha + w_i + \sum_{n=1}^{N} \beta_n \ln x_{nit} + v_{it} - u_{it}$$

where we utilize the distributional assumptions of the normal-exponential model for the composed error term:

(34) $$v_{it} \sim N(0, \sigma_u^2)$$ $$u_{it} \sim \exp(\theta), \sigma_u = \frac{1}{\theta}$$

The inefficiency component is assumed to vary randomly in time, with the firm-specific, time-invariant term w_i capturing any cross-firm unobserved heterogeneity. The distributional assumption for w_i is purposefully loose, with the only requirements that the variable has a mean of zero and a finite variance (normality is not required). Because of the addition of this term, however, the estimation procedure is slightly more complicated than with the previous models. Greene (2005a) derives the likelihood function and suggests estimating the parameters by way of simulated maximum likelihood. The individual technical efficiencies can be estimated by way of modifying the JLMS estimator slightly, as in Greene (2005b), where now the conditioning must be done w.r.t. both the composed error and w_i [19]. This model allows for the inefficiency to vary over time in a more flexible way than the time-decay

[19] i.e. we must obtain $E[u_{it}|\varepsilon_{it}(w_i)]$ in place of the usual $E(u_{it}|\varepsilon_{it})$.

model introduced earlier, and the inclusion of a firm-specific effects allows untangling unobserved factors from inefficiency, which otherwise would bias our estimates.

Thus far we have considered various different specifications of the stochastic frontier production model, with each having a distinctly different flavor to them. The classic cross-sectional model provides us with an overview of a specific snapshot in time with regards to the sectors considered, while the panel data models give us a sense of how the technical efficiency varies throughout our panel period. Still, these models all come with the fundamental assumption of the parametric production function, which we now relax as we introduce the fully nonparametric StoNED –model.

3.3. Stochastic Nonparametric Envelopment of Data (StoNED)

The second method used in this thesis is a rather new, nonparametric estimation technique developed by Kuosmanen (2006) and Kuosmanen & Kortelainen (2012), called Stochastic Nonparametric Envelopment of Data (StoNED). This technique can be viewed as a synthesis of the earlier DEA and SFA techniques, including both the nonparametric frontier estimation inherent in DEA and the stochastic element included in SFA.

A notable amount of work has been done in the field of efficiency analysis to hybridize the two rather different approaches of DEA and SFA. The main objective has been to produce a model that would incorporate both the nonparametric, axiomatic frontier estimation from DEA alongside the stochastic component from SFA. Fan, Li and Weersink (1996) were among the first to propose a semiparametric kernel regression model in which the benefits of DEA and SFA would be utilized. Other contributions towards this unified framework in efficiency analysis include the models proposed by Kumbhakar, Park, Simar & Tsionas (2007) who follow Fan et al. (1996) in proposing a semiparametric local polynomial regression model and Banker & Maindiratta (1992) who extend the functionality of the basic SFA model by the nonparametric techniques employed in DEA.

However, most of these models[20] have proven to be either computationally very difficult or otherwise cumbersome to work with. As also noted by Kuosmanen (2006) aside from computational issues, none of the proposed models produce results that satisfy the basic production axioms. StoNED is a framework that unifies these two models, while still maintaining a statistical pedigree – that is to say, the method allows for statistical inference and is based on justifiable assumptions. StoNED retains the nonparametric estimation inherent to DEA alongside a stochastic, composed error component as in the SFA approach (Kuosmanen, 2006). While still a new model, StoNED has already received a rather substantial vote of confidence, having been adopted as the efficiency analysis method of choice by the Finnish Energy Authority (EMVI, Energiamarkkinavirasto) for the 2012-2015 regulatory period (Kuosmanen, Kortelainen, Kultti, Pursiainen, Saastamoinen & Sipiläinen, 2010).

The presentation of this model follows the approach by Kuosmanen et al. (2014) and the seminal papers by Kuosmanen (2006) and Kuosmanen & Kortelainen (2012). To begin and to give some insight into where this method fits in modern efficiency analysis, Table 2 below classifies the typical methods in efficiency analysis by both model type (parametric vs. nonparametric) and their handling of inefficiency. The top row contains methods which attribute stochastic disturbances to simple statistical error, the middle row attributes them fully to inefficiency and the last row contains the composed error models which are utilized in this thesis, where the disturbances contain both inefficiency and stochastic elements.

	Parametric	**Nonparametric**
Central tendency	OLS	CNLS
Deterministic frontier models	COLS, MOLS	DEA
Stochastic frontier models	SFA	StoNED

Table 2: Methods used in efficiency analysis (Kuosmanen et al. 2014)

[20] For a rigorous comparison of some earlier approaches of these types of models, we refer the reader to Badunenko, Henderson, & Kumbhakar (2011).

A further benefit of the StoNED -method to take into account is the fact that most of the earlier models used in efficiency analysis can be viewed as being nested in the StoNED –model (Kuosmanen & Kortelainen, 2012). Specifying the model with certain parameter restrictions or a parametric functional form leads to the StoNED –frontier producing the classic DEA and SFA –models, respectively[21]. As such the StoNED –model is directly applicable in the same contexts as the previous classic techniques, and furthermore perhaps also in contexts where both the classic models would be deemed unsuitable. A thorough comparison of StoNED against the other classic techniques was presented by Andor & Hesse (2014), who conclude that the model performs remarkably well under various types of Monte Carlo simulations, particularly excelling in situations where the data in question is subject to substantial noise. Methodologically, the model builds on a regression technique called Convex Nonparametric Least Squares (CNLS), originally introduced by Hildreth (1954) and extended by Hanson & Pledger (1976) in the first stage, applying the Method of Moments technique to the residuals in the second stage to both correct the frontier and obtain estimates of individual inefficiencies using the JLMS estimator. Alternatively, in the second stage, untangling the composite error term into its components can be achieved using a technique called quasi-likelihood estimation introduced by Fan et al. (1996), and extended to the StoNED model by Kuosmanen et al. (2014).

We begin in a similar fashion as in SFA by assuming a frontier production function $f(\boldsymbol{x})$, transforming inputs to outputs such that it represents the maximal output for a given input level. However, we now relax the parametric functional form assumption of SFA and assume that the function f belongs to a class of continuous, monotonic increasing and globally concave functions, denoted F_2, which can be non-differentiable. Again, observed output may fluctuate from the frontier output due to inefficiency and noise, which are captured in the composite error term epsilon, analogous to the SFA formulation.

(35)
$$\begin{gathered} y_i = f(\boldsymbol{x_i}) + \varepsilon_i \\ \varepsilon_i = v_i - u_i \end{gathered}$$

[21] A rigorous proof for DEA appears in Kuosmanen & Johnson (2010)

Due to the exclusion of an explicit production function, the model must now be estimated using nonparametric methods. The general form of the CNLS problem is very difficult to solve, as the family F_2 contains infinitely many functions from which the optimal f can be chosen from. This issue was tackled by Kuosmanen (2008) who shows that the optimal solution can in fact be reduced to a much simpler quadratic programming (QP) problem that retains the qualities of the original problem. The frontier estimator for the StoNED model is the CNLS regression, which is formulated in the following way.

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} \varepsilon_i^2 \,, \quad \text{s.t.} \tag{36}$$

$$y_i = \alpha_i + \boldsymbol{\beta}_{\boldsymbol{i}}' \boldsymbol{x}_{\boldsymbol{i}} + \varepsilon_i \quad \forall i \tag{37}$$

$$\alpha_i + \boldsymbol{\beta}_{\boldsymbol{i}}' \boldsymbol{x}_{\boldsymbol{i}} \le \alpha_h + \boldsymbol{\beta}_{\boldsymbol{h}}' \boldsymbol{x}_{\boldsymbol{i}} \quad \forall i, h \tag{38}$$

$$\boldsymbol{\beta}_{\boldsymbol{i}}' \ge 0 \quad \forall i \tag{39}$$

Unlike the more general problem, this problem can now be solved using modern algebraic solvers (such as MINOS, CPLEX, etc.). This algorithm constitutes the standard CNLS regression with an additive error term. The interpretation of the model is intuitive when we go over the algorithm one equation at a time. The objective (36) is to minimize the sum of squared residuals, mirroring closely the standard OLS approach. The first constraint (37) defines the hyperplane to which every observation is projected. The estimated coefficient $\boldsymbol{\beta}_{\boldsymbol{i}}'$ represents economically the vector of marginal products for $\boldsymbol{x}_{\boldsymbol{i}}$ (Kuosmanen & Kortelainen, 2012). Note that each observation receives a specific marginal productivity vector or perhaps the same vector if they are projected to the same hyperplane – however, the marginal productivities are not required to be equal for all observations as with SFA. The second constraint (38) is at the heart of the CNLS –model, as it specifies concavity of the hyperplanes between all pairs of observations. This constraint is often called the ‘Afriat inequality’, after Afriat (1972)[22]. Figure 8 provides a graphical illustration of a CNLS production frontier alongside the Afriat inequality.

[22] Afriat’s theorem (1972) states that for n observations and m inputs the following statements hold:

I. A globally concave function $f: R^m \to R$, satisfying $y_i = f(\boldsymbol{x}_{\boldsymbol{i}})$ exists
II. $\exists(\alpha_i, \beta_i) = (\beta_{i1}, \dots, \beta_{im})^T, y_i = \alpha_i + \boldsymbol{\beta}_{\boldsymbol{i}}^{\boldsymbol{T}} x_i \ \forall i$ satisfying $\alpha_i + \boldsymbol{\beta}_{\boldsymbol{i}}^{\boldsymbol{T}} x_i \le \alpha_j + \boldsymbol{\beta}_{\boldsymbol{j}}^{\boldsymbol{T}} x_i \quad \forall\, i, j\ i \neq j$

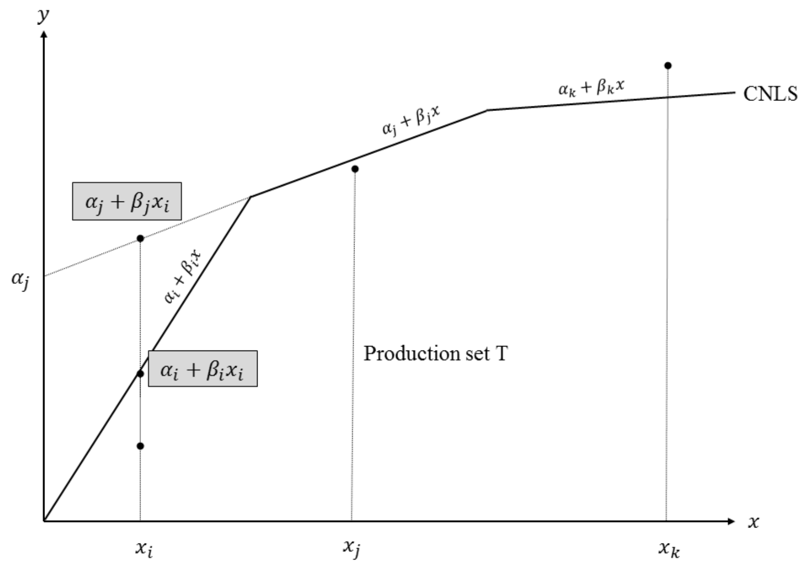


Figure 8: CNLS production frontier

In the figure above, we see the Afriat inequality in action. In this simple visualization, the dotted line represents the hyperplane of observation x_j when projected onto the region of x_i's hyperplane. When evaluated at x_i, the concavity constraint ensures that the dotted line segment lies always above the original hyperplane. This provides the necessary concavity to the production frontier, maintaining the CNLS frontier as the convex hull of the production set. For a more thorough discussion on Afriat's theorem and CNLS, we refer the reader to Lee, Johnson, Moreno-Centero & Kuosmanen (2013). The third constraint (40) states that the estimated frontier is monotonic, i.e. production is non-decreasing in input quantity. The CNLS frontier is thus the piecewise linear approximation of the true production function. By way of the Afriat inequalities, the estimated frontier now also preserves the fundamental production axioms, which is a major benefit to nonparametric estimation.

However, due to the specifics of our particular data set, we must adjust the model slightly. As our data justifiably includes firm-specific heterogeneity, a log-transformed model is preferable to the additive error model above (Kuosmanen et al, 2014). For the multiplicative error model, consider a production function of the type:

(40) $$y_i = f(\boldsymbol{x_i}) \exp(\varepsilon_i)$$

Where the only modification to the previous model is the now multiplicative error term. Taking the logarithm, we get the regression equation for the multiplicative error model:

(41) $$\ln y_i = \ln f(\boldsymbol{x}_i) + \varepsilon_i$$

And finally the complete model is stated as:

(42) $$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\phi},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} \varepsilon_i^2 \text{, s.t.}$$

(43) $$\ln y_i = \ln(\phi_i + 1) + \varepsilon_i$$

(44) $$\phi_i = \alpha_i + \boldsymbol{\beta}_i'\boldsymbol{x}_i - 1$$

(45) $$\alpha_i + \boldsymbol{\beta}_i'\boldsymbol{x}_i \leq \alpha_h + \boldsymbol{\beta}_h'\boldsymbol{x}_i$$

(46) $$\boldsymbol{\beta}_i \geq 0$$

Unlike in SFA, where we can apply the transformation to the variables themselves due to the specific functional form assumption, now we must apply the logarithm to the approximating function φ. If we were to take a log-transformation of the variables themselves, the actual estimated model would be a piecewise log-linear frontier[23] – i.e. the hyperplanes would themselves be essentially Cobb-Douglas production functions! This obviously would defeat the entire purpose of CNLS as we would end up right back with a type of parametric frontier we wish to abstract away from. To deal with the computational issue of not taking logarithms of zero, we must add one to the logarithmic regression equation (43), compensating for this with subtracting one from the supporting hyperplane equation in (44). The interpretation of the constraints remains, with the difference that the marginal productivity vector is now inside the logarithmic transformation. This modification also allows us to utilize a Farrell measure of technical efficiency, which can now be estimated and interpreted in the same manner as in SFA.

Note also the deep connection that the StoNED –model has with DEA. As the production frontier is estimated using tangential hyperplanes, this allows us to consider production

[23] For further information on this type of frontier estimation, we refer the reader to Banker & Maindiratta (1986).

models with multiple outputs and inputs with relative ease by using a distance function –based approach, when compared to the parametric estimation in SFA. This flexibility is another notable benefit of the StoNED –model, when compared to parametric methods.

Similarly to DEA, both variants of the model can accommodate different types of scale efficiencies. The scale efficiency in the model is determined by the value of the estimated intercept term α_i:

- for $\alpha_i = 0 \; \forall \; i$, the estimated model satisfies constant returns to scale
- for $\alpha_i \geq 0 \; \forall \; i$, the frontier exhibits non-increasing returns to scale
- for $\alpha_i \leq 0 \; \forall \; i$, the frontier exhibits non-decreasing returns to scale

Kuosmanen & Kortelainen (2012) state that for the additive error model, these specifications influence the estimated inefficiencies in the 2nd stage, recommending the use of the multiplicative error model if these assumptions are to be imposed.

After the production frontier has been consistently estimated with CNLS in the first stage, we can proceed with the estimation of technical efficiency. To this end, we utilize the CNLS residuals, and the information they contain on the inefficiency term. We must note here that the following method assumes explicitly the standard normal-half-normal model variant. At present, alternative distributional assumptions have been considered, but no published results have yet surfaced. The MoM estimation utilizes the CNLS residuals, denoted here as ε_i^{CNLS} and closely resembles the MOLS approach discussed earlier. As the residuals sum to zero (Kuosmanen et al., 2014), we can utilize the central moments of the residuals to obtain estimates of σ_u and σ_v. For the estimated residuals the second and third sample moments are:

(47) $$\widehat{M}_2 = \sum_{i=i}^{n} \left(\varepsilon_i^{CNLS} - \bar{\varepsilon}^{CNLS}\right)^2 \times \frac{1}{n}$$

(48) $$\widehat{M}_3 = \sum_{i=i}^{n} \left(\varepsilon_i^{CNLS} - \bar{\varepsilon}^{CNLS}\right)^3 \times \frac{1}{n}$$

where we see that the second moment is simply the sample variance and the third moment is the nominator of the skewness measure of the joint density function for epsilon. The derived theoretical equivalents to these are (cf. Kumbhakar & Lovell, 2000):

(49) $$M_2 = \left(\frac{\pi - 2}{\pi}\right)\sigma_u^2 + \sigma_v^2$$

(50) $$M_3 = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(1 - \frac{4}{\pi}\right)\sigma_u^3$$

inserting the estimated moments to the above equations and solving for σ_u and σ_v we get the Method of Moments estimators for our parameter values:

(51) $$\hat{\sigma}_u = \left[\frac{\hat{M}_3}{\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(1 - \frac{4}{\pi}\right)}\right]^{\frac{1}{3}}$$

(52) $$\hat{\sigma}_v = \left[\hat{M}_2 - \left(\frac{\pi - 2}{\pi}\right)\hat{\sigma}_u^2\right]^{\frac{1}{2}}$$

After the estimates have been obtained, the JLMS estimator of $E(u_i|\varepsilon_i^{CNLS})$ can be utilized directly to obtain firm-specific estimates of the inefficiency term, but with the slight modification that the conditioning is now on the CNLS residuals. As with classic SFA (Aigner et al., 1977), the expected value of inefficiency μ, is obtained from

(53) $$\mu = \sqrt{\frac{2}{\pi}} \times \hat{\sigma}_u$$

The estimated frontier is then corrected in (41) by using the mean value of inefficiency as

$$f(\boldsymbol{x}_i)\exp(\mu)$$

This is the technique utilized for estimating the technical efficiency in most applications of the StoNED –method. Published applications of this model include Kuosmanen (2012) who estimates the efficiency of Finnish electricity distributors, Eskelinen & Kuosmanen (2013), who focus on analyzing the branch efficiency in Finnish retail banking and Mekaroonreung & Johnson (2012) who estimate shadow prices for greenhouse gases for U.S. coal power plants. In our analysis we apply the process outlined by Kuosmanen et al (2014) for the StoNED – estimation for our purposes, noting here that our main goal is to estimate the technical efficiencies, and the frontier itself is of less interest to us. Our analysis proceeds with the following steps: in the 1st stage we utilize the multiplicative error CNLS –regression (equations 42 to 46) to obtain both the production frontier and the CNLS residuals, and in the 2nd stage we use the Method of Moments approach to estimate the firm-specific inefficiencies by way of the JLMS estimator. The estimated technical efficiencies and marginal productivities are then compared to their parametric counterparts.

A noted problem with this model appears, however, if the estimated third moment $\left(\widehat{M}_3\right)$ is positive as it is part of the skewness measure of the composed error term epsilon, which should be left-skewed (cf. Figures 6 & 7 in chapter 3.2.). If we estimate a positive third moment, this implies that epsilon is not negatively skewed, and this is problematic. Greene (2008) posits that this skewness of the residuals serves as an 'internal diagnostic' of the frontier model, suggesting that a positive estimate implies either a misspecified model or an unsuitable application – however according to Kuosmanen & Kortelainen (2012) such a blunt diagnosis might be too strict, as the 'wrong skewness' problem arises sometimes even with correctly specified frontier models. The SF literature provides alternative ways of dealing with this issue, with some authors proposing a substitution of an arbitrarily small value for σ_u – Coelli (1995) suggests using $\sigma_u = 0.05$, while Kuosmanen (2012) substitutes $\sigma_u = 0.00001$. Andor & Hesse (2014) favor the original suggestion by Kuosmanen & Kortelainen (2012), setting $\widehat{M}_3 = -0.0001$. From the applied econometricians' point of view, these suggestions are still somewhat ambiguous as to whether the problem lies in the data or the model itself, and if under such circumstances valid inferences can be made on the estimated technical efficiencies. Furthermore, since most of these recommendations are provided in the context of testing the frontier model using simulated data (where, of course the true DGP[24] is known), we are unsure of how tenable these ad hoc -methods are in applied analysis.

[24] Data Generating Process

While this model generalizes both the DEA and SFA approaches, its one drawback lies in the formulation. The StoNED –model is computationally quite demanding, with computation times increasing markedly when the sample size increases (Lee et al., 2013). This results from the Afriat inequality (38), which involves the enveloping hyperplanes. As we require the constraint to hold for each pair of observations, we thus have for n observations a system of $n(n-1)$ inequalities. As such, sample sizes of several thousand observations[25], which are often the norm in SFA or DEA can easily begin to test the hardware limits in terms of memory capacity. As acknowledged by Kuosmanen et al. (2014) this is indeed one of the outstanding issues with StoNED –models. Approaches to solve this bottleneck have been suggested (cf. the models proposed by Lee et al., 2013) but for the purposes of this thesis, the author was unable to procure access to more sophisticated algorithms other than the basic model. This purely practical problem of computational difficulty necessitates that we select a suitable subsample from our data for the sectors considered. This point is discussed in more detail in the next chapter which deals with our data set.

The StoNED –model addresses a central theme in econometrics, which is the juxtaposition of structure and flexibility. As noted by Bauer (1990), the more structure imposed on the models, the more precise the estimates become, but now with the important caveat that the structural assumptions must be correct in the first place. Coupling together nonparametric estimation with the composed error term is in our view a good solution to this problem, allowing a good mix of precision and flexibility.

[25] Indeed, for a sample of a thousand observations we would have 999 000 linear convexity constraints, a considerable amount for a computer to hold in working memory!

3.4. Other considerations in efficiency analysis

3.4.1. Contextual Variables

While stochastic frontier methods provide a way of examining the inefficiency of various economic decision-makers, the classic methods still do not provide a clear-cut way of identifying the particular determinants of inefficiency. In this section we will briefly discuss the ways in which stochastic frontier methods have been extended to estimate the impact of outside factors to inefficiency. Contextual variables are defined as exogenous factors to the classic model that may help explain firm-level inefficiency (Cornwell & Schmidt, 2008). E.g. for production models these can include such measures as the quality of management or ownership structure (state vs. privately owned), and for cost frontier models variables which characterize the operating environment or market properties have often been utilized (cf. Kuosmanen, 2012, Kopsakangas-Savolainen & Svento, 2011). Early approaches to introduce exogenous determinants of inefficiency to the stochastic frontier context utilized a two-stage estimation method derived from DEA, where in the first stage, SFA is utilized to obtain the inefficiencies of individual firms, and in the second stage the predicted inefficiency scores are utilized as the dependent variable in (typically) truncated regression or tobit regression with the exogenous variables. However, as noted by Kumbhakar & Lovell (2000) this approach is highly inconsistent. This is due to the fact that the first-stage assumption in SFA is to assume the inefficiency terms to be distributed according to the various specifications discussed earlier, yet in the second stage the terms are assumed to be functions of the contextual variables[26]. This problem necessitates joint estimation of both the frontier and the contextual effects in the first stage. The usual method for this in SFA is to parametrize the error term, or its distribution with a function that relates either the mean or variance of inefficiency to the contextual variables. With this procedure, the frontier model can then be estimated with ML consistently in the first stage. The inclusion of contextual variables in SFA is closely related to the allowance of heteroscedasticity in either error component, where the inefficiency and error terms are also handled in a notably similar fashion.

[26] One could even argue here that if such factors are said to influence inefficiency, then they should already belong in the frontier production function itself.

In the nonparametric StoNED –model, the effects of contextual variables can be handled in a more flexible way than in SFA, with fewer restrictions on the variables themselves. Within the model, the contextual variables are introduced into the CNLS regression itself, maintaining the flexibility of StoNED when compared to SFA. This method was proposed by Johnson & Kuosmanen (2011) who refer to the model as StoNEZD (StoNED with Z-variables). A notable benefit of this method is the joint estimation of the effects of contextual variables and the frontier itself, neither requiring a parametric functional form or the parametrization of the inefficiency terms' distributional properties. Recent applications of the StoNEZD method include Kuosmanen (2012), Kuosmanen et al. (2010), Kuosmanen, Saastamoinen and Sipiläinen (2013) and Eskelinen & Kuosmanen (2013). Another desirable trait of the StoNEZD –model are the lesser restrictions imposed on the contextual variables. Whereas in parametric estimation, the contextual variables are required to be uncorrelated to the inputs, which is a difficult assumption to justify intuitively, in StoNEZD the effects can be estimated even when the contextual variables are correlated with input factors. While these methods extend the reach of basic frontier estimation methods, they also clearly require good quality data regarding the contextual variables themselves.

At this stage, having motivated both the theoretical basis and the methods for efficiency analysis we are ready to move onto application. We now bring these models to bear on our data to ascertain the efficiency of Finnish heavy industry.

4. Data description

The dataset utilized in this thesis consists of a panel of companies operating in the Finnish industry during the years 2003 to 2009. The data was extracted from the YRTTI –database, maintained at the Government Institute for Economic Research (VATT). The database contains the complete yearly financial information reported by Finnish companies to the taxation authority. The data consists of 75385 observations, with yearly observations and firm-year counts[27] illustrated in Figure 9 below. As we see from the leftmost graph, the number of observations stays relatively stable throughout our panel period, with about 10000 observations per year. The panel year histogram indicates that our data is an unbalanced panel, but we have a very respectable proportion of companies that we observe throughout the entire panel period.

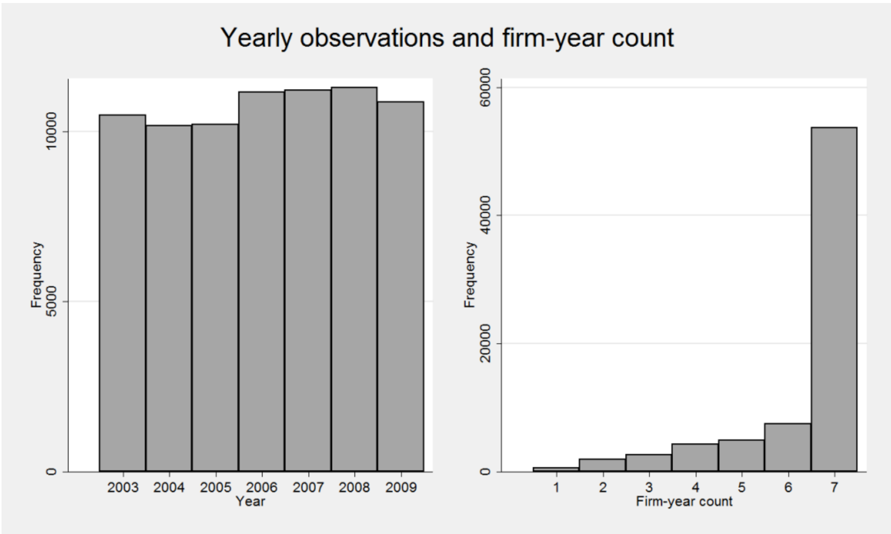


Figure 9: Panel characteristics

We restrict our analysis to the relatively heavy industries, specifically to the following sectors (classification as per the 2-digit level Standard Industrial Classification 2002, by Statistics Finland). Measuring by turnover, together these sectors comprise 44% of the industrial sector[28]. The specific industries analyzed in this thesis are the following:

[27] i.e. a single firm observed during a single year denotes a single *firm-year*
[28] Source: Statistical Yearbook of Finland 2013, Statistics Finland (author’s own calculation)

Table 3: Industry classification by sector (SIC2002)

Sector code	Industry	N obs.
10	Mining of coal and lignite	3837
13	Mining of metal ores	248
14	Other mining and quarrying	2673
20	Manufacture of wood and wood products	21335
21	Manufacture of pulp and paper products	1499
23	Manufacture of coke, refined petroleum, oil refining	132
24	Manufacture of chemicals and chemical products	2532
25	Manufacture of rubber and plastic products	4752
27	Manufacture of basic metals	258
28	Manufacture of fabricated metal products	35797

We see that some sectors are somewhat underrepresented in our data, while other sectors have an abundance of observations. The sectors with a low number of observations are primarily ones where we observe a low number of relatively large firms, with high market capitalizations. Measuring by turnover, the largest sectors in our data are oil refining (23) and the paper and pulp sector (21). As we find later on when we estimate our models, some sectors did in fact suffer heavily from a low sample size, with particularly the simple cross-sectional model being affected by this.

4.1. Descriptive statistics

Table 4 below presents the means for our main variables of interest over the panel period. We find that the firms in our data are somewhat heterogeneous, with large variability in company turnover and the number of persons employed.

Table 4: Descriptive statistics (pooled)

Sector	10	13	14	20	21	23	24	25	27	28
Personnel	5.253	15.146	10.264	10.150	102.926	197.000	56.921	31.580	85.598	11.739
Turnover*	730.0	3679.0	2030.8	1724.7	63635.3	421013.0	18249.2	4569.2	28752.4	1303.2
Capital*	474.0	1385.5	976.8	437.7	33627.0	33017.3	5136.2	1514.3	7394.4	328.3
Materials*	144.6	500.6	727.7	999.9	35473.4	361547.9	9901.8	2310.3	18241.6	549.8
SME (%)	94.7%	94%	97%	93.5%	87.5%	78%	89.3%	96.4%	91.8%	96.9%

* = 1000 EUR

The overall proportion of SMEs[29] (Small and Medium-sized Enterprises) in our data is 95.18%, which is typical of the Finnish economy as a whole. However, our proportion of SMEs is still slightly lower than for the whole economy, where the proportion of SMEs is 97-99%[30] with slight differences on whether classified by the personnel or turnover criteria. In this aspect our data contains a higher proportion of larger companies – perhaps this is not very surprising, given that we restrict our attention to the industrial sector, which is known for its behemoth companies in Finnish parlance. We see that the mean number of employees is highest in oil refining (23), paper and pulp industry (21) and in metal refining (27). The lowest mean numbers of employees for the average firm are found in coal and lignite mining (10), wood production (20) and the manufacturing of metal products (28). Mean turnover is highest in oil refining (23), followed by the paper and pulp industry (21), with physical capital being also highest in these two sectors. We compute some standard aggregate measures of productivity for our selected sectors, with Figure 10 providing a per worker comparison of turnover and physical capital, and Figure 11 plotting the development of mean turnover alongside the input usage for materials and physical capital over the panel period.

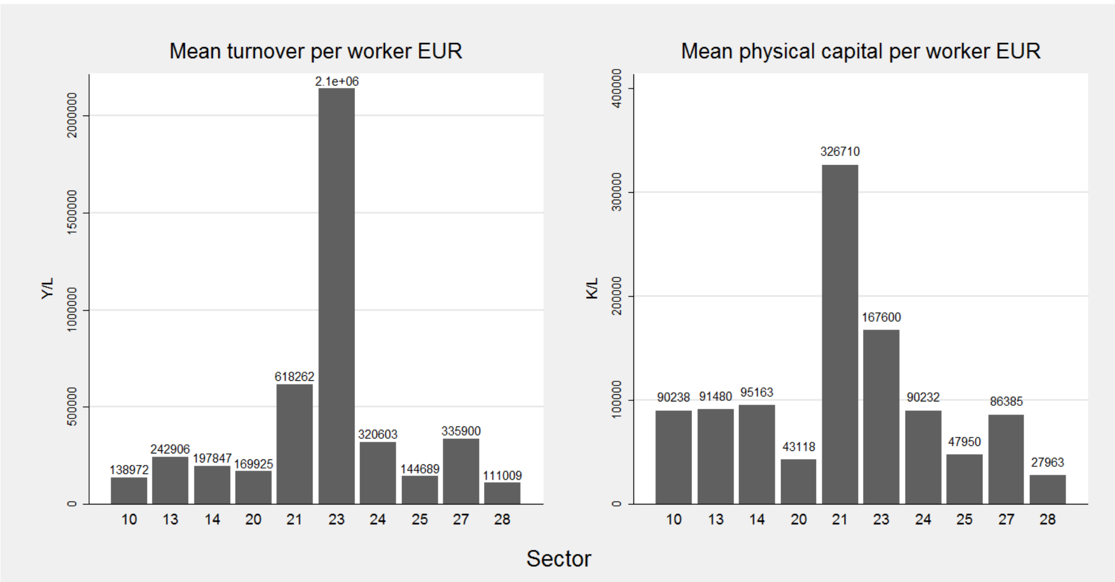


Figure 10: Per worker measures of turnover and capital

[29] The definition of an SME provided by Statistics Finland is one in which the number of employees is under 250 and the yearly turnover not exceeding 50 million €.
[30] This data was obtained from the Business Registry overview 2012, provided by Statistics Finland.

We see from the per worker comparison graphs that the oil and paper industries are heavily capitalized and have a significantly higher mean turnover than the rest of our chosen sectors.

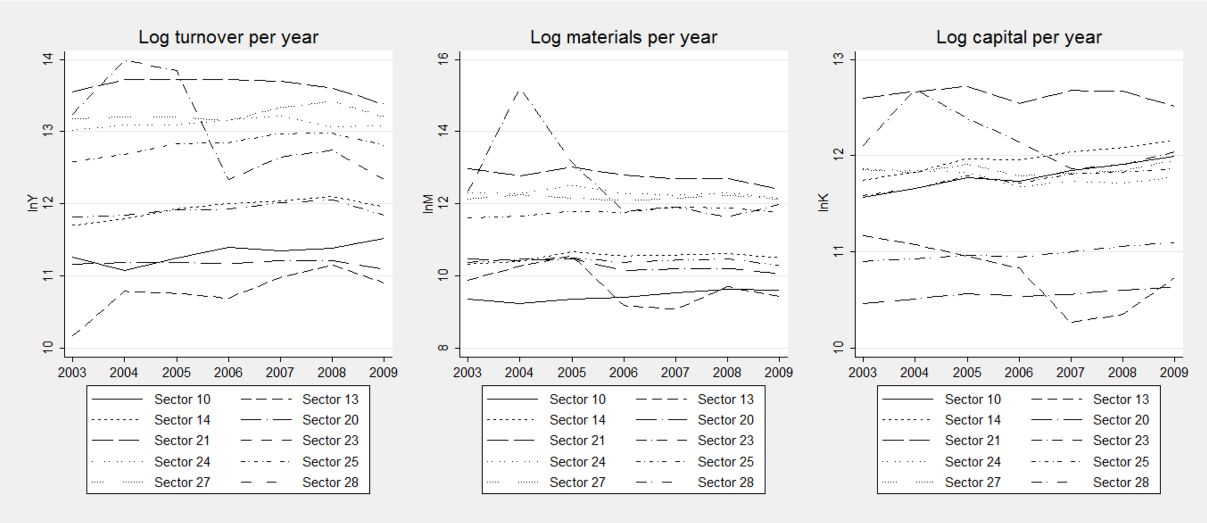


Figure 11: Turnover and input usage during 2003-2009

We find that, overall the majority of sectors have increased their turnover during the period, with some sectors experiencing rather significant fluctuations and some remaining relatively stable. Of the sectors whose turnover has declined over the period, the wood production industry (20) seems to have been hit the hardest, with a rather significant drop in turnover occurring during 2005-2006. The input usage for materials seems to be rather uniform for the majority of the sectors, with metal ore mining (13) and oil refining (23) having decreased material usage during the period. For physical capital, for the former sectors the capital stock declines during the observation period, while for other sectors we see an upward trend of capital accumulation.

4.2. Model variables

The output variable is defined for our production models as the deflated company turnover. To maintain comparability across time periods, we deflate the variables with the aggregated producer price indices[31] obtained from Statistics Finland, defining our first panel year (2003) as the base year. The reason for this is to more accurately capture the inefficiency effects and to smooth out estimation error resulting from input price changes under the timeframe studied (especially poignant when we consider frontier estimation for panel data).

In our models, we utilize three input variables, which are:

- Labor (L): defined as the number of employees a company has on staff for a given year. The variable is an aggregate indicator of labor input, defined as the number of full-time and full-time equivalent persons employed by the company.
- Intermediate materials (M), defined as the deflated value of the procured materials and supplies for a firm for a given year. Following Coelli et al. (2005) we consider a suitable deflator index for the variable to obtain the real value of material inputs for a given year.
- Capital Stock (K), defined as the value of the physical capital stock the company operates. Details on the computation of the capital stock are found under the next subsection.

Obtaining the labor and material input variables is straightforward from the data, as we have the nominal values for these inputs directly. For the chosen inputs, our estimations would benefit from more detailed data, such as subdivision between skilled and unskilled labor, but unfortunately such detailed measures could not be procured from the data. These inputs categorize broadly the industry production process we consider in our models.

4.2.1. Computation of capital stock

In the computation of the company-specific capital stock variable, we utilize a technique called Perpetual Inventory Method (PIM), which is a way of computing physical capital stocks by utilizing the flows of investments (time series data on investment) for a certain asset type. The investment series was obtained by utilizing the depreciation database for the

[31] The selected index is the producer price index for commodities.

companies in our panel data set, included in YRTTI. For the purpose of accuracy, we compiled the depreciation data from the start of the YRTTI –database (year 1994), and compute the capital stocks for our companies for the period 2003-2009. The yearly depreciation rates for capital stock were obtained from Statistics Finland ASTIKA time series database, computing the sector specific rate for sector i as

(54) $$\delta_i = \sum_{t=1975}^{T=2012} \left(\frac{D_{it}}{K_{it}}\right) \times \frac{1}{(T-t)}$$

where D denotes the gross depreciation of the capital stock for sector i at time t, and K denotes the gross capital stock for the same sector for the same period. We compute the depreciation rate as the average rate over the entire time period. After the depreciation rates are obtained, the individual capital stocks are calculated by utilizing the reported value of undepreciated moveable assets as the base value (denoted as b_{it}) and using the yearly addition reported for the asset class as a proxy variable for yearly investment in physical capital (denoted as k_{it}). The asset class under consideration is defined as machines and equipment. The capital depreciation profile considered is geometric depreciation, i.e. we assume that capital depreciates at a constant rate, defined by δ_i. The real-valued capital stock for firm *i* at time *T* is computed recursively as:

(55) $$K_{iT} = (1-\delta_i)^{T-t_0} b_{it_0} + \sum_{t=t_0}^{T-1} (1-\delta_i)^{T-t_0} k_{it}$$

This represents an investment series stretching out to the first period (t_0) the firm is observed, with summation to a desired end period (T). As observed by Coelli et al., (2005) having an accurate measure for the capital stock variable is important for the purposes of estimating the frontier models. Many empirical applications of SFA have utilized the simple book values for capital inputs, however this measure is subject to both accounting practices and tax planning, and as such the PIM measure is usually preferred if it can be reliably calculated.

4.3. Subsample selection for StoNED

As mentioned in chapter 3.3, the StoNED –model is computationally intensive, and given the size of our data, we must restrict our analysis to a subsample of the firms observed during a single panel year. In our analysis what this means is that we must construct representative subsamples of 100-150 observations for a given sector to be able to estimate the StoNED –model. The subsample is selected using random sampling, with sample size restricted to 100 observations for a given sector. We use a pseudorandom number generator to generate scores for each observation, sort the observations and select the first 100 observations for the subsample by sector. This process should give us a relatively unbiased subsample for a given sector for our data set, while maintaining a computationally feasible sample size. We subsample our data for the nonparametric model by utilizing our last panel year (2009) as the base population and generate the random subsample by the above manner. We find that for some sectors, the number of available observations falls below acceptable standards, so we do not estimate the nonparametric models for these sectors. Our chosen subsample contains the following sectors, with the number of observations presented in the table below:

Sector code	Industry	N obs.
10	Mining of coal and lignite	99
14	Other mining and quarrying	99
20	Manufacture of wood and wood products	99
21	Manufacture of pulp and paper products	99
24	Manufacture of chemicals and chemical products	99
25	Manufacture of rubber and plastic products	99
27	Manufacture of basic metals	99
28	Manufacture of fabricated metal products	99

The sample sizes for both the metal ore mining sector (13) and oil refining sector (23) were notably low at 54 and 22 observations, respectively. For this reason, we exclude these sectors from our nonparametric models, as the estimated models would likely yield highly biased results due to the small sample size.

5. Estimation results

This chapter presents our estimation results for the parametric and nonparametric models, respectively. Outputs and inputs are selected according to the previous chapter. In panel models, we allow for Hicks-neutral technical change[32] with the inclusion of a time trend variable T. Additional estimated parameters are explained for each model where necessary. All parametric models are estimated without imposing CRS, with hypothesis tests performed during postestimation. The p-values for these tests are provided in the estimation tables, for the sake of brevity. For some models and sectors, the frontier model did not achieve convergence due to either a low sample size or model misspecification. These sectors are omitted from the results, with accompanying discussion on the possible reasons for non-convergence.

5.1. SFA models

5.1.1. Cross-sectional normal-exponential model

For our cross-section, we estimate the following standard normal-exponential model:

$$\ln y_i = \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \beta_3 \ln M_i + v_i - u_i$$
$$v_i \sim N(0, \sigma_v^2)$$
$$u_i \sim \exp\left(\frac{1}{\sigma_u}\right)$$

With the final panel year (2009) selected as the period for our cross-sectional analysis. Table 5 presents our estimation results for the cross-sectional model for all the sectors considered.

[32] i.e. we allow the frontier output to vary according to the time trend, while maintaining the same marginal rates of technical substitution between inputs (Romer, 1996).

Table 5: Cross-sectional normal-exponential SFA results

Sector	**10**	**14**	**20**	**21**	**24**	**25**	**27**	**28**
	$\ln y_i$	$\ln y_i$	$\ln y_i$	$\ln y_i$	$\ln y_i$	$\ln y_i$	$\ln y_i$	$\ln y_i$
$\ln K$	0.272^{***}	0.263^{***}	0.138^{***}	0.196^{***}	0.139^{***}	0.0674^{***}	0.105^{***}	0.0960^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln L$	0.502^{***}	0.496^{***}	0.481^{***}	0.448^{***}	0.550^{***}	0.456^{***}	0.511^{***}	0.568^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln M$	0.154^{***}	0.253^{***}	0.411^{***}	0.378^{***}	0.387^{***}	0.437^{***}	0.406^{***}	0.330^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_0	6.822^{***}	6.387^{***}	5.806^{***}	5.787^{***}	6.292^{***}	6.638^{***}	6.402^{***}	7.291^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma_u^2)$	-3.273	-1.597^{***}	-3.049^{***}	-2.338^{***}	-2.971^{***}	-2.330^{***}	-2.975^{***}	-2.436^{***}
	(0.064)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma_v^2)$	-0.872^{***}	-1.775^{***}	-1.612^{***}	-1.780^{***}	-1.445^{***}	-1.963^{***}	-1.685^{***}	-1.696^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_u$	0.195	0.450	0.218	0.311	0.226	0.312	0.226	0.296
$\hat{\sigma}_v$	0.647	0.412	0.447	0.411	0.486	0.375	0.431	0.428
λ	0.301	1.093	0.487	0.756	0.466	0.832	0.525	0.691
p_{CRS}	0.092	0.698	0.004	0.377	0.000	0.007	0.212	0.416
LogL	-257.9	-165.2	-974.1	-118.5	-180.6	-326.8	-190.1	-2303.3
N	251	186	1348	158	226	473	271	3020

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

For some of the sectors included in the analysis, the sample size is very low, thus affecting the convergence of the models and likewise the reliability of the results. The sectors which did not converge are metal ore mining and oil refining sectors (13 and 23) and for this reason are omitted from the table. The sample size could have been augmented by considering a pooled model, but this would obviously affect the comparability of the cross-sectional models when we interpret the output elasticity and technical efficiencies. The estimated output elasticities of inputs vary considerably between sectors, with highest elasticity of capital in both of the mining sectors (10 and 14), followed closely by the paper and pulp sector (21). The output elasticity of labor seems to be relatively stable across all sectors considered, with remarkably similar parameter estimates in all sectors. The output elasticity for intermediate products and materials exhibits higher variability across sectors, with the rubber and plastic manufacturing industry (25) and the wood production industry (20) receiving the highest estimates and conversely both the mining sectors (10, 14) receiving a low estimate for this parameter. Furthermore all our specified inputs are estimated as statistically significant, with all receiving positive coefficients. The signal-to-noise ratio[33] varies from 0.3 to slightly above unity, indicating that some inefficiency is found in all valid sectors. The null hypothesis of constant returns to scale in production cannot be rejected for 5 of the valid sectors (10, 14, 21, 27 & 28) while for the remaining sectors the CRS hypothesis is rejected at the 1% significance level.

We estimate the technical efficiency of each sector with the JLMS estimator. Table 6 below presents the relevant summary statistics for the estimated technical efficiencies for each sector.

Table 6: Estimated technical efficiency, cross-sectional model

	Mean	Median	S.D.	Min	Max	N
Sector 10	.8245032	.8302928	.0493239	.4149393	.8960347	250
Sector 14	.6783292	.6982899	.1550031	.00373	.9067243	186
Sector 20	.8098265	.8246998	.0778864	.0618402	.9396582	1348
Sector 21	.7505505	.7682191	.1226527	.1315873	.9239034	158
Sector 24	.803422	.8187942	.0791801	.2349097	.9631763	226
Sector 25	.7544979	.7786691	.1199259	.008998	.9447996	473
Sector 27	.8049767	.817481	.0853935	.2132607	.99	271
Sector 28	.7588835	.7832275	.1111705	.0089885	.9446272	3020
N						5954

[33] i.e. the estimated parameter lambda, defined as $\lambda = \frac{\hat{\sigma}_u}{\hat{\sigma}_v}$

The mean efficiency across sectors seems notably uniform, with most sectors estimated to be in the range of 75%-83% efficient, except for the quarrying industry (14). The most efficient sectors in this model are estimated to be the coal mining sector (10) with a mean efficiency of 82%, followed closely by wood production (20), metal refining (27) and chemical manufacturing (24), with efficiencies of 81% and 80%, respectively. The lowest estimated mean efficiency is estimated for the quarrying sector (14), at 68% average efficiency. Figure 12 below provides a visualization of the mean efficiencies for each sector under analysis.

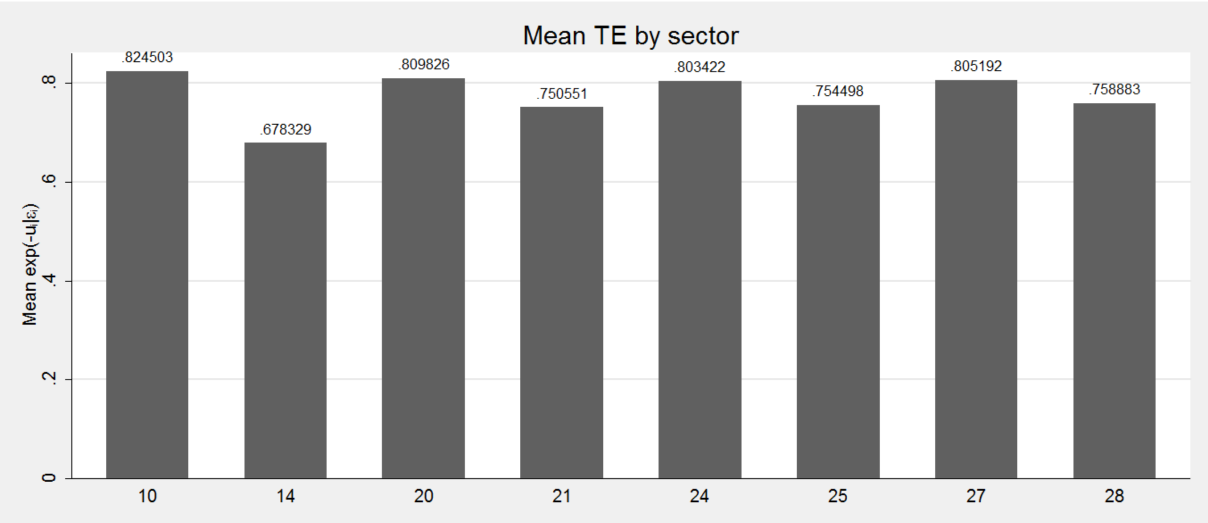


Figure 12: Mean technical efficiency by sector

5.1.2. Panel data models

5.1.2.1. Time-invariant model

We estimate the Pitt & Lee (1981) time-invariant random-effects model as:

$$\ln y_{it} = \beta_0 + \beta_1 \ln K_{it} + \beta_2 \ln L_{it} + \beta_3 \ln M_{it} + \delta T + v_{it} - u_i$$

$$v_{it} \sim N(0, \sigma_v^2)$$
$$u_i \sim N^+(0, \sigma_u^2)$$

Table 7 presents the estimation results for our time-invariant model.

Table 7: Time-invariant panel SFA results

Sector	**10**	**13**	**14**	**20**	**21**	**23**	**24**	**25**	**27**	**28**
	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$
$\ln K_{it}$	0.188^{***}	0.0755	0.243^{***}	0.201^{***}	0.274^{***}	-0.0842	0.194^{***}	0.104^{***}	0.229^{***}	0.103^{***}
	(0.000)	(0.352)	(0.000)	(0.000)	(0.000)	(0.143)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln L_{it}$	0.496^{***}	0.635^{***}	0.495^{***}	0.425^{***}	0.279^{***}	0.552^{***}	0.459^{***}	0.422^{***}	0.462^{***}	0.486^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln M_{it}$	0.186^{***}	0.426^{***}	0.226^{***}	0.351^{***}	0.404^{***}	0.650^{***}	0.338^{***}	0.410^{***}	0.313^{***}	0.310^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
T	0.0195^{**}	0.00716	0.0183^{**}	0.00163	-0.0156^{**}	0.0435	0.0127^{*}	0.0115^{***}	0.0175^{***}	0.0224^{***}
	(0.004)	(0.873)	(0.002)	(0.435)	(0.007)	(0.141)	(0.019)	(0.001)	(0.000)	(0.000)
β_0	8.101^{***}	6.370^{***}	7.142^{***}	6.330^{***}	5.419^{***}	5.966^{***}	7.034^{***}	6.846^{***}	6.699^{***}	7.883^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma_u^2)$	1.406^{***}	0.377	0.915^{***}	0.868^{***}	0.842^{***}	0.995^{*}	1.339^{***}	0.711^{***}	0.943^{***}	0.821^{***}
	(0.000)	(0.123)	(0.000)	(0.000)	(0.000)	(0.041)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma_v^2)$	0.210^{***}	0.459^{***}	0.139^{***}	0.127^{***}	0.122^{***}	0.0940^{***}	0.149^{***}	0.125^{***}	0.116^{***}	0.114^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_u$	1.186	0.614	0.957	0.932	0.917	0.998	1.157	0.843	0.971	0.906
$\hat{\sigma}_v$	0.458	0.678	0.373	0.357	0.350	0.307	0.386	0.354	0.341	0.338
λ	2.589	0.906	2.563	2.613	2.622	3.253	2.996	2.384	2.848	2.679
p_{crs}	0.000	0.041	0.031	0.000	0.005	0.148	0.599	0.000	0.820	0.000
LogL	-1467.7	-79.60	-899.2	-5998.8	-657.6	-28.45	-1171.0	-1976.6	-1073.6	-11772.3
N	1663	71	1365	9651	1125	48	1641	3381	1862	20704

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The output elasticity estimates for the time-invariant model show a little more variability than the cross-sectional model. The estimates for elasticities of the inputs are statistically significant and positive for most sectors, with few exceptions. The output elasticity of capital is statistically insignificant even at the 10% level for metal ore mining and oil refining (sectors 13 and 23). However, the sample size for these sectors is quite low, especially given the length of our panel which can influence the results. As with the cross-sectional model, we find high output elasticity of capital for the mining sectors (10, 14) and also the wood production and metal refining sectors (20, 27). Particularly in metal refining (27), the output elasticity of capital is much higher when compared to the cross-sectional model.

For the labor input, the estimated output elasticity is again relatively stable across sectors, with only the metal ore mining and paper and pulp sectors (13, 21) receiving estimates that differ somewhat substantially from the other sectors. We find that for paper production (21), the estimated output elasticity of labor is in this model now much lower than with the cross-sectional model. Intermediate products and materials is the input with highest variability in output elasticity, still we find that the estimates for this model are pretty much in line with the previous cross-sectional SFA model. The highest elasticities of the material input are found in oil refining (23), metal ore mining (13) and rubber and plastic manufacture (25). The time trend variable is estimated to have a statistically significant effect in 6 of the 10 sectors considered, with estimated technological progress of 1-2% per year. For 4 sectors, the trend variable is estimated to have a statistically insignificant effect at the 1% significance level. The null of CRS is rejected for half of the sectors, with the mining sectors (10, 13), oil refining (23) and the chemical and plastic industries (24, 27) estimated to maintain CRS in production.

We estimate the technical efficiency of each sector with the JLMS estimator. The summary statistics of estimated technical efficiencies are presented below in Table 8.

Table 8: Estimated technical efficiency, time-invariant model

	Mean	Median	S.D.	Min	Max	N
Sector 10	.6692877	.755369	.2200278	.0779095	.9872186	3837
Sector 13	.781378	.8281071	.1120344	.2716586	.860218	248
Sector 14	.6930793	.7577068	.2106908	.1319497	.9623473	2673
Sector 20	.7184975	.8164089	.2035004	.0240199	.9848098	21335
Sector 21	.5759172	.5134617	.217789	.0958686	.9793308	1499
Sector 23	.7722552	.8684067	.19728	.301295	.9288165	132
Sector 24	.5914191	.5535239	.2402759	.0540433	.9752855	2532
Sector 25	.6417187	.6102683	.1924603	.0013292	.9869947	4752
Sector 27	.5849272	.5383598	.2314885	.0562682	.9816268	2580
Sector 28	.676789	.6995192	.2125119	.0482854	.989049	35797
N						75385

For the estimated technical efficiencies, most sectors fare much worse in the time-invariant model than in the cross-section, with estimated mean efficiencies in the range of 57-77%. The largest differences between efficiency estimates when compared to the cross-sectional model are found in metal refining (27), chemical manufacture (24) and coal and lignite mining (10). Intuitively, the estimated mean efficiencies seem to be on the low side, with efficiencies of 57% and 58% estimated for the paper industry (21) and metal refining (27), respectively. However, this aspect of our results is not too surprising when we consider the very rigid assumptions this model relies on. The time-invariance assumption of inefficiency (and as a result also technical efficiency) can be viewed as especially restrictive, particularly if the data contains firm-specific heterogeneity, as noted by Greene (2005a). Figure 13 below summarizes the estimated mean efficiencies by sector.

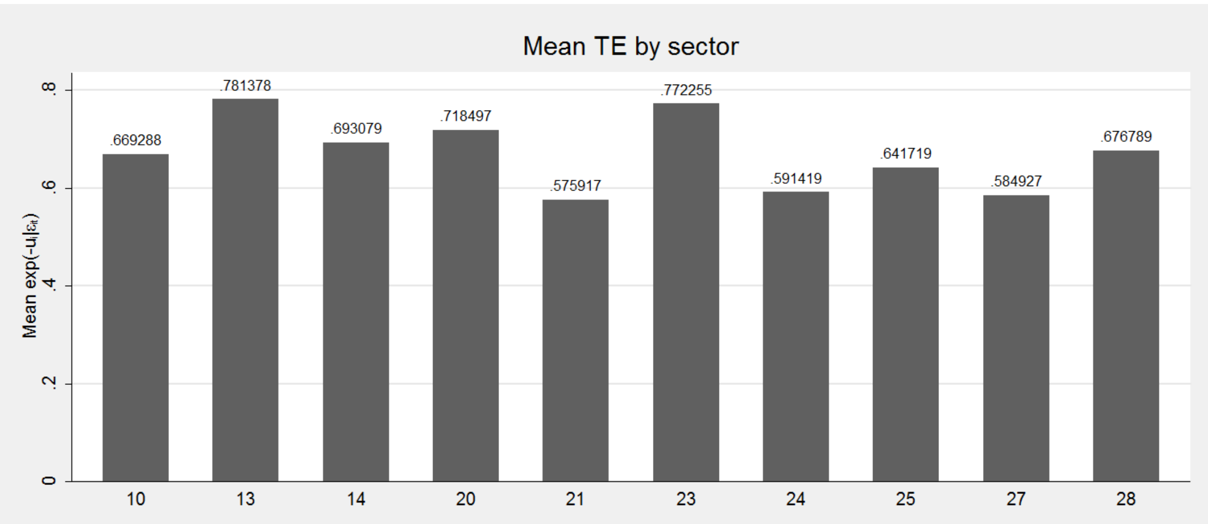


Figure 13: Mean technical efficiency, time-invariant model

5.1.2.2. Time-varying models

We estimate the Battese & Coelli (1992) time-decay model as:

$$\ln y_{it} = \beta_0 + \beta_1 \ln K_{it} + \beta_2 \ln L_{it} + \beta_3 \ln M_{it} + v_{it} - u_{it} + \delta T$$

with the following assumptions for the distributions of the error components and the time trend function:

$$\begin{aligned} v_{it} &\sim N(0, \sigma_v^2) \\ u_{it} &= \beta(t) u_i \\ u_i &\sim N^+(\mu, \sigma_u^2) \\ \beta(t) &= \exp\big(-\eta(t - T)\big) \end{aligned}$$

Table 9 presents the estimation results for this model for all the sectors considered. Graphs for the estimated trend functions are relegated to the appendix.

Table 9: Time-decaying panel SFA results

Sector	**10**	**13**	**14**	**20**	**21**	**23**	**24**	**25**	**28**
	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$
$\ln K_{it}$	0.282***	0.0814	0.230***	0.164***	0.219***	-0.0359	0.164***	0.107***	0.111***
	(0.000)	(0.322)	(0.000)	(0.000)	(0.000)	(0.621)	(0.000)	(0.000)	(0.000)
$\ln L_{it}$	0.368***	0.610***	0.469***	0.396***	0.325***	0.396***	0.436***	0.419***	0.444***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln M_{it}$	0.133***	0.430***	0.225***	0.403***	0.423***	0.705***	0.398***	0.441***	0.351***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
T	0.0736***	0.0497	0.0853***	0.0565***	0.0117	0.0999	0.0760***	0.0498***	0.0608***
	(0.000)	(0.384)	(0.000)	(0.000)	(0.481)	(0.124)	(0.000)	(0.000)	(0.000)
β_0	10.78***	6.052***	7.846***	7.513***	7.343***	5.870***	8.242***	7.905***	8.749***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma^2)$	-0.617***	5.403	-0.902***	-1.046***	-1.188***	-0.798*	-0.736***	-1.111***	-1.041***
	(0.000)	(0.541)	(0.000)	(0.000)	(0.000)	(0.037)	(0.000)	(0.000)	(0.000)
γ	0.928***	6.242	0.765***	0.766***	0.687***	1.610**	0.951***	0.693***	0.938***
	(0.000)	(0.481)	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.000)	(0.000)
μ	4.402***	-558.7	1.634***	2.361***	2.530***	1.857*	2.863***	2.364***	2.287***
	(0.000)	(0.910)	(0.000)	(0.000)	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)
η	-0.0148***	-0.103	-0.0445***	-0.0227***	-0.00900	-0.0387	-0.0222***	-0.0165***	-0.0181***
	(0.000)	(0.222)	(0.000)	(0.000)	(0.175)	(0.291)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_u$	0.622	14.89	0.526	0.490	0.450	0.612	0.588	0.468	0.504
$\hat{\sigma}_v$	0.391	0.657	0.359	0.334	0.320	0.274	0.365	0.331	0.315
p_{crs}	0.000	0.058	0.000	0.000	0.032	0.361	0.900	0.002	0.000
LogL	-1261.2	-79.07	-851.7	-5441.0	-556.5	-26.73	-1078.1	-1803.3	-10689.4
N	1663	71	1365	9651	1125	48	1641	3381	20704

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The output elasticities from the time-decay model differ somewhat from the earlier cross-sectional and time-invariant models. For the elasticity of capital, particular differences are found in metal ore mining (13) and oil refining (23) where the elasticity of capital is statistically insignificant. For the most part, the estimated parameters are much in line with earlier models w.r.t. capital elasticities. For labor input, the highest elasticity is estimated in metal ore mining (13) but for other sectors we find remarkably similar results across the sectors when compared to the earlier models. For material inputs, the estimated elasticities are almost one-to-one comparable to the time-invariant model. For the metal refining sector (27), this model failed to converge. Thus we omit the estimation results for this sector, being unable to make any trustworthy inference based on the model estimates.

A notable issue in our results is the estimate of the time-decay parameter eta (η). For the majority of sectors, we estimate a statistically significant and negative coefficient. As explained in subsection 3.2.4. this now implies that counterintuitively, technical efficiency decreases over time. We also find that in cases where eta is statistically significant, the time trend is estimated also to have a significant and particularly high positive effect, while the trend variable is statistically insignificant in models where eta is also statistically insignificant[34]. The range of parameter estimates for eta suggests that for our sectors, technical efficiency is decreasing at a rate of 1.5-4.5% per year. We estimate the technical efficiencies for all sectors with the JLMS estimator, the summary statistics are provided below in Table 10.

Table 10: Estimated technical efficiency, time-decay model

	Mean	Median	S.D.	Min	Max	n
Sector 10	.3180353	.3679933	.1380169	.0015823	.7730885	1663
Sector 13	.7788491	.8114699	.0812613	.4655049	.852858	71
Sector 14	.6437501	.6770985	.1328407	.0568507	.8334417	1365
Sector 20	.5370632	.5877486	.1403843	.0067725	.909677	9651
Sector 21	.548958	.5822697	.108591	.0388075	.8979213	1125
Sector 23	.5439452	.6088429	.1834756	.0581147	.7844315	48
Sector 24	.5107344	.5468742	.125607	.0144755	.9434608	1641
Sector 25	.5605792	.5943768	.11236	.0010015	.8910145	3381
Sector 28	.5717149	.6177716	.1348086	.0003036	.9242431	20704
N						41511

[34] This might indicate some form of collinearity issues, however in both the original application of the model in Battese & Coelli (1992) and subsequent work, a time-trend variable is generally included in the model.

For the viable sectors, mean efficiencies are estimated as notably low in comparison to the time-invariant and cross-sectional models. The mean efficiency across all sectors varies between 30-78%, with metal ore mining (13) estimated as the most efficient sector by far. In contrast to earlier models, the coal mining industry (10) is estimated to be the least efficient of the sectors. To see the estimated time-varying effects we present the mean estimated efficiencies by sector alongside the estimated means over the panel period in Figure 14 below.

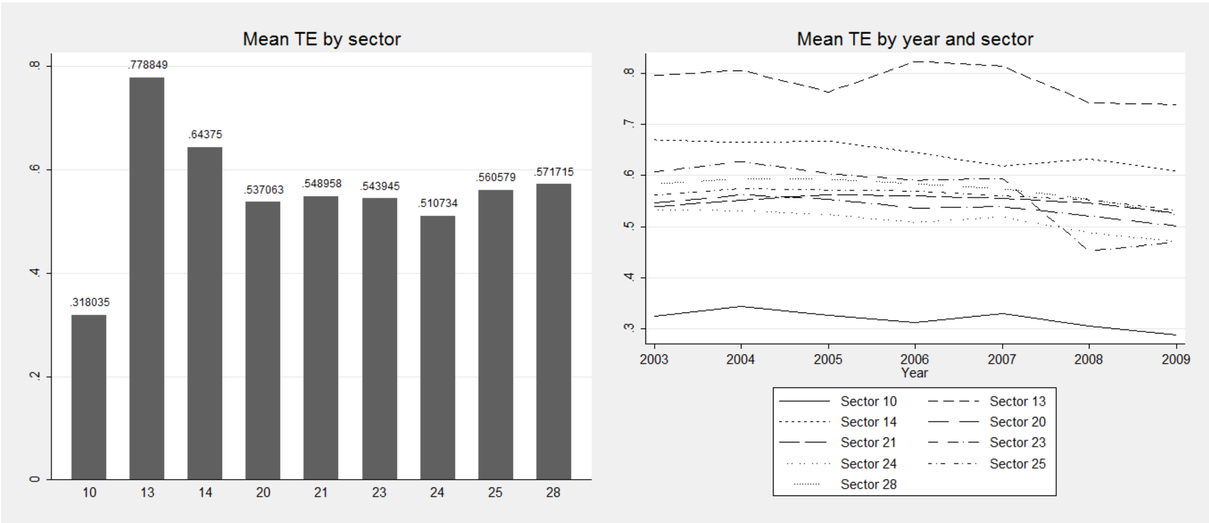


Figure 14: Mean technical efficiency by year and sector, time-decaying model

We find that the estimated technical efficiency varies across the panel period, decreasing over time. The 'bumps' in the line graphs are due to the unbalanced nature of our panel data. Looking at the overall mean across the time period, we find that the efficiencies are very low. Even when considering the peculiarities of the time period under analysis (e.g. the financial crisis), a mean efficiency of 50% can be considered exceptionally low. Thus with the exception of the output elasticity parameters, we find that the results of this model stand in stark contrast to the models estimated previously. Greene (2005b) notes that this specific model may be biased in short panels due to the parametrization of efficiency effects over time. This becomes especially relevant if we find heterogeneity within the sectors themselves, because in this case we are essentially forcing all firms in a specific sector to have the same monotonic pattern of time-varying efficiency (η). This assumption may prove to be too strict, which is why we aim to address this issue with our next model.

Our final parametric model is the following true random-effects model, by Greene (2005a, 2005b):

$$\ln y_{it} = \alpha + w_i + \beta_1 \ln K_{it} + \beta_2 \ln L_{it} + \beta_3 \ln M_{it} + v_{it} - u_{it} + \delta T$$

where w_i is a time-invariant term capturing firm-specific heterogeneity and α is an intercept parameter common to all firms in the sector, resembling the fixed-effects model. As before, T is the time trend variable, allowing for Hicks-neutral technological progress.

The distributional assumptions for the error components are:

$$v_{it} \sim N(0, \sigma_v^2)$$
$$u_{it} \sim \exp\left(\frac{1}{\sigma_u}\right)$$

For the firm-specific factors, no distributional assumptions are made beforehand other than the assumption that they are distributed with a finite variance. Table 11 presents the estimation results for the true random-effects model.

Table 11: True random-effects panel SFA results

Sector	**10**	**13**	**14**	**20**	**21**	**24**	**25**	**27**	**28**
	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$	$\ln y_{it}$
$\ln K_{it}$	0.215***	0.160*	0.294***	0.176***	0.185***	0.0986***	0.163***	0.182***	0.127***
	(0.000)	(0.047)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln L_{it}$	0.332***	0.544**	0.383***	0.336***	0.219***	0.298***	0.282***	0.500***	0.410***
	(0.000)	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln M_{it}$	0.146***	0.343***	0.174***	0.354***	0.375***	0.380***	0.369***	0.290***	0.304***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
T	0.0208***	0.0475	0.0227***	0.00667***	-0.00114	0.0195***	0.0134***	0.0217***	0.0259***
	(0.000)	(0.235)	(0.000)	(0.000)	(0.679)	(0.000)	(0.000)	(0.000)	(0.000)
α	7.758***	6.362***	6.672***	6.247***	6.814***	7.337***	6.719***	6.962***	7.350***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma_u^2)$	-2.457***	-1.071***	-2.536***	-2.574***	-2.619***	-2.358***	-2.656***	-2.639***	-2.764***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(\sigma_v^2)$	-2.692***	-4.035**	-3.791***	-3.971***	-5.592***	-5.802***	-4.899***	-3.905***	-3.796***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\theta}$	0.968***	0.507**	0.639***	0.681***	1.043***	1.724***	0.833***	0.649***	0.620***
	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_u$	0.293	0.586	0.281	0.276	0.270	0.308	0.265	0.267	0.251
$\hat{\sigma}_v$	0.260	0.133	0.150	0.137	0.0611	0.0550	0.0864	0.142	0.150
λ	1.125	4.403	1.872	2.012	4.421	5.595	3.070	1.883	1.676
p_{crs}	0.000	0.683	0.000	0.000	0.000	0.000	0.000	0.035	0.000
LogL	-1285.9	-65.97	-660.5	-4487.6	-373.9	-837.5	-1149.0	-788.5	-8562.2
N	1663	71	1365	9651	1125	1641	3381	1862	20704

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In this model, the oil refining sector (23) suffered a lack of convergence, which may be caused by a very low sample size (cf. Tables 7 & 9), and for this reason we omit this sector from the tabled results. The output elasticities in the TRE model differ somewhat from our earlier SFA models. For capital, highest elasticity is estimated for quarrying (14). For other sectors, we find that the model estimates fall somewhere between the earlier models, and exhibit greater variability between sectors than in the previous models. For the labor input, highest elasticities are estimated for mining sectors (10, 13, and 14) and metal refining (27), which falls in line with earlier models. In contrast to earlier models, the elasticities for the material input seem to be somewhat uniform, bar sectors 10 (coal and lignite mining) and 14 (quarrying). The time trend variable is estimated to have a statistically significant effect for the majority of the sectors. The impact of this variable is now more in line with the time-invariant model than the time-decay model, where these parameter estimates were found to be somewhat inflated.

Significant firm-specific heterogeneity is found in all sectors, indicated by the parameter estimate of theta ($\hat{\theta}$), which characterizes the simulated standard distribution of the firm-specific intercept term (w_i). In this regard, our results for Finnish industrial production are in line with Green & Mayes (1991), who likewise infer significant within-sector heterogeneity to be present in the UK industries, although using different models. The signal-to-noise ratio varies substantially across the sectors, with significant inefficiency estimated to be present in all sectors. Also in contrast to the earlier models, the null hypothesis of CRS is maintained only in metal ore mining (13) and chemical manufacturing (24) at the 1% significance level, with rejection of CRS in all other sectors considered.

We estimate the technical efficiencies of the firms by way of the modified JLMS estimator, as defined by Greene (2005b). Table 12 presents the summary statistics of the estimated efficiencies. To illustrate the estimated time-varying efficiency effects, Figure 15 provides both the estimated mean efficiencies by sector over the panel time period, alongside a plot for the mean efficiencies in each sector for each year.

Table 12: Technical efficiency by sector, TRE -model

	Mean	Median	S.D.	Min	Max	N
Sector 10	.7470894	.7871113	.1404787	.0231889	.9393414	1663
Sector 13	.6177982	.6687851	.2250328	.0510581	.9070556	71
Sector 14	.7539788	.7980068	.1600758	6.61e-06	.9491786	1365
Sector 20	.7585419	.8150398	.1758432	0	.9560826	9649
Sector 21	.7341525	.7991204	.2101491	0	.9563442	1125
Sector 24	.6639304	.7300599	.224807	0	.9470574	1641
Sector 25	.7579121	.8216035	.1924036	0	.9618658	3381
Sector 27	.7689176	.8296016	.1770161	0	.9564107	1862
Sector 28	.7814448	.831233	.1571242	0	.9595957	20700
N						41457

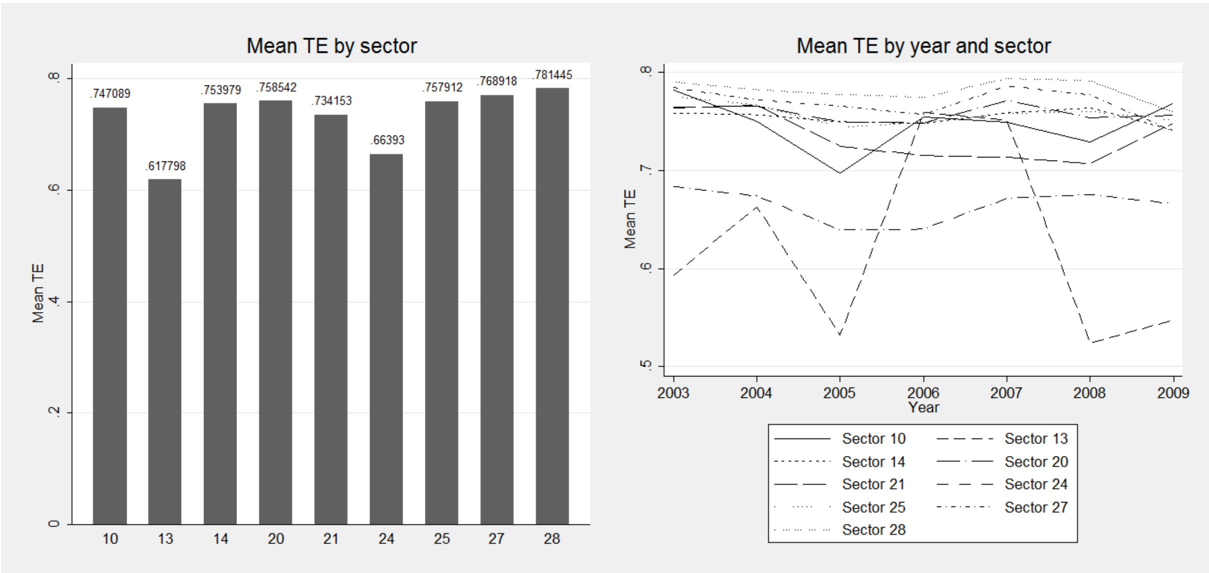


Figure 15: Estimated technical efficiency by sector and year, TRE model

Looking at the development of efficiency over our panel period, the true random-effects model gives us a fairly comprehensive view of the efficiency changes in the sectors. We find that for several sectors, the estimated efficiency is almost uniform, with very small changes over the period, but several sectors seem to have rather significant changes in their efficiency over the period. The diverse range of effects, combined with the significant firm-specific heterogeneity may be the reason why we see such low efficiency estimates in the time-invariant and the time-decaying SFA model. The mean efficiency is significantly higher for all sectors in the TRE model, and the sector plot reveals some interesting developments during our panel period. Both mining sectors (10, 13) seem to have suffered some misfortunes during the years 2004-2005, with their efficiency showing a clear drop for these periods. Also these sectors seem to have been impacted heavily from the demand slump beginning in 2008 – of the two, metal ore mining (13) seems to have fared much worse than its fellow sector, with mean efficiency fluctuating more heavily and increasing only to 55% by the end of our panel period. Most all sectors show a drop during the 2004-2005 periods, however the majority seems to have recovered from this efficiency drop by year 2009, with some sectors estimated to have increased their efficiency in comparison to 2003. For paper and pulp production (21) we clearly see a downward trend in efficiency, beginning in 2005 with the trend ending somewhere in 2008 and by 2009 the sector is estimated to have reached the pre-drop efficiency. The impact of the global financial crisis is the probable reason for a clear drop in estimated efficiency for 2008 for all sectors, however as also noted by Koop (2001) clear procyclicality of the technical efficiencies is hard to pin down.

An interesting point to consider here is the fact that during 2005, the paper and forestry industry in Finland underwent a heavily publicized and costly labor battle, which our estimations show had a significant effect on the sectors technical efficiency. It is well within the realm of possibility that that this event had significant knock-on effects on most of the sectors we consider in this thesis as well. Given that we find a more or less consistent downward trend in the estimated efficiencies of most sectors starting approximately around the time of this event, it offers one compelling explanation to our estimated temporal variability in the efficiencies.

Also notable is the result that for some sectors, the estimated technical efficiencies are very close to being time-invariant. In quarrying (14), wood manufacturing (20) and manufacturing of metal products (28) the technical efficiency is estimated to be rather uniform throughout our panel period, with the graphs being essentially flat. Comparing the efficiencies of these sectors to the time-invariant model, we find that the time-invariant efficiency estimates are still substantially lower for these sectors. Given the significant estimates of firm-specific heterogeneity, this seems to indicate that the earlier models are indeed picking up some of this heterogeneity in the composed error term, thus leading to biased efficiency estimates. The overall results from the TRE model seem quite encouraging as the decomposition of firm-specific heterogeneity allows us to analyze efficiency in the sectors in a more precise manner.

Moreover, we find results that are in line with major events in the Finnish industry as a whole during the time period.

5.2. StoNED

We estimate the following StoNED –model, now imposing CRS by restricting $\alpha_i = 0$[35].

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\phi},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} \varepsilon_i^2 \text{, s.t.}$$

$$\ln y_i = \ln(\phi_i + 1) + \varepsilon_i$$

$$\phi_i = \beta_{ki}K_i + \beta_{li}L_i + \beta_{mi}M_i - 1$$

$$\beta_{ki}K_i + \beta_{li}L_i + \beta_{mi}M_i \leq \beta_{kh}K_i + \beta_{lh}L_i + \beta_{mh}M_i \forall\, i, h$$

$$\beta_{ij} \geq 0$$

As mentioned previously under chapter 3.3., the definition of the model (non-parametric estimation) makes tabulating the estimated parameters somewhat cumbersome. As each observation is projected onto a hyperplane, we have at the most n different hyperplanes with a parameter vector $\boldsymbol{\beta_i}$ which is $[n \times 3]$, we present our results as in Kuosmanen (2012). The estimated parameter vector is multiplied by the expected value of inefficiency (μ) to obtain the marginal productivities of the inputs in question. We note here that for certain sectors, we encountered the wrong skewness problem, and as a result, are unable to provide estimation results for these sectors (this issue is discussed further when we estimate the technical efficiencies). Table 13 presents the summary statistics of the estimated marginal productivities for valid sectors. The actual number of hyperplane segments forming the CNLS frontier ranged from 13-20 in our analysis depending on the sector. This compares favorably with Kuosmanen (2008), who finds similar results in Monte Carlo simulations for a Cobb-Douglas DGP for a similar sample size and signal-to-noise ratio. For interested readers, the optimization algorithm used for this model is provided in Appendix A.

[35] The author refers to advice received from prof. Kuosmanen in the model specification. Another way of justifying this restriction is to state that we expect firms to be operating at the *technically optimal production scale* (Coelli, et al, 2005, p. 59)

Table 13: Summary statistics of marginal productivities in the StoNED -model

β_{ij}	Mean	S.D.	Min	Max	n
β_{ki}					
Sector 14	.3986314	1.187267	9.31e-10	11.63156	99
Sector 20	.5647135	1.134969	0	9.639045	99
Sector 21	.1760216	.1624187	0	.3607891	99
Sector 24	.5673164	.7110923	0	1.484563	99
Sector 25	.2451989	.7199075	0	6.427761	99
Sector 28	.240064	.9155385	0	8.419203	99
β_{li}					
Sector 14	31971.14	29992.78	9.31e-10	184610.4	99
Sector 20	26977.3	18342.71	0	60492.8	99
Sector 21	10278.98	10918.24	0	29223.55	99
Sector 24	12216.77	49892.09	0	499277.5	99
Sector 25	7710.87	12127.41	0	107963.2	99
Sector 28	6380.964	3490.127	0	14257.6	99
β_{mi}					
Sector 14	8.425549	45.40911	.5423432	376.4385	99
Sector 20	1.375527	1.73844	.4886344	9.615304	99
Sector 21	.5149583	.3603314	.3161999	2.431649	99
Sector 24	.6148749	.3074948	.2134038	2.231255	99
Sector 25	.4827802	.6391428	.0116071	5.710814	99
Sector 28	.7317588	1.038242	.0347925	6.960261	99
N					594

Due to nonparametric estimation, the marginal productivities are much more varied than in the parametric models. Since the marginal productivities represented here are summary statistics over the entire in-sample sector, they are also harder to interpret in intuitive economic terms than in parametric SFA. Unfortunately, this is the cost for relaxing the production function assumption, as the parameters no longer have a neat interpretation as output elasticities. Nonetheless, we estimate high marginal productivities for capital in the wood production and chemical manufacturing industries (20, 24). For labor, distinctively high estimates are found for wood production and quarrying sectors (20, 14) and the same sectors are estimated to have the highest productivities for material inputs as well.

Following Kuosmanen et al. (2014) and Kuosmanen & Kortelainen (2012), the technical efficiency is obtained from the CNLS residuals by utilizing the Method of Moments approach, with the conditional efficiency via the JLMS point estimator (cf. Chapter 3.2.3.). Equations for the moment conditions that are used to decompose the residuals are found in chapter 3.3. We must note here that for coal and lignite mining (10) and metal manufacturing

(27), we find that the third central moment $\widehat{M_3}$ of the residuals was estimated to be positive. As discussed previously, this is not a specifically unusual problem in the context of efficiency analysis, and there has been discussion on how to deal with this 'wrong skewness' of the residuals in both parametric SFA and nonparametric estimation. Given the relative validity of our results in the previous section, we might suggest that this effect might be due to some peculiarity in our data rather than misspecification of our models. Still, we do not posit that we have found an optimal solution to this problem, and for this reason we report the (full) estimation results for the sectors in which the regular moment conditions apply.

Table 14 below presents the summary statistics for the JLMS estimators of technical efficiency, computed as $TE_i = \exp\left(-E\left(u_i \middle| \varepsilon_i^{CNLS}\right)\right)$ alongside the usual error component estimates ($\hat{\sigma}_u, \hat{\sigma}_v$ and the signal-to noise ratio λ).

Table 14: Technical efficiency estimates for the StoNED model

Sector	**14**	**20**	**21**	**24**	**25**	**28**
TE						
Mean	.4937208	.5444598	.813469	.7927677	.8291774	.8659342
S.D.	.0994944	.0954808	.1249652	.1319052	.1221598	.1057647
Min	.0748019	.0898081	.2936328	.2382404	.4745087	.5119669
Max	.7015002	.7388261	1	.9858699	.9956139	.9865161
$\hat{\sigma}_u$	1.166955	1.055883	.3852133	.4540763	.3269823	.3255435
$\hat{\sigma}_v$	.8955932	.7799989	.3074054	.4045265	.2366103	.3688377
λ	1.302996	1.353699	1.253112	1.122488	1.381945	.8826198
N	99	99	99	99	99	99

We compute the coefficient of determination R^2_{StoNED} for our StoNED –models. As the method relies on the CNLS regression, the computation is done analogously to standard regression methods. However, since the model is specified as a log-linear relation, the coefficient now measures the proportion of explained variance in the logarithm of the output. The R^2_{StoNED} is estimated as:

$$R^2_{StoNED} = 1 - \frac{Var(\varepsilon_{CNLS})}{Var(\ln y)}$$

As a robustness check, we also inspect the CNLS residuals with the standard Shapiro-Wilk test for normality. If we can reject the null hypothesis of normality with a high degree of confidence, this indicates that our model does not suffer from misspecification. The estimated StoNED –model is tested for heteroskedasticity with the standard White (1980) test, which is directly applicable in the context of this model as the residuals are obtained from the CNLS regression. The test is implemented as an artificial regression of the squared CNLS residuals against the regressors, the squared regressors and their interactions. The test score for N observations is then obtained as $\widehat{W} = NR^2$, and is distributed as $\chi^2(n^2)$ with the lowercase n denoting the number of inputs used. The p-value for the null of homoscedasticity is then defined as $P(\widehat{W}) = 1 - \chi^2(n^2, \widehat{W})$. Table 15 below presents the R-squared alongside the results of the Shapiro-Wilk –test and the White test, respectively.[36]

Table 15: Coefficients of determination and robustness test scores for StoNED

Sector	14	20	21	24	25	28
R^2	.8651	.9319	.9784	.9612	.9781	.9171
S-W p-value	0.000	0.000	0.000	0.000	0.000	0.003
$\widehat{W}$	11.08	4.237	5.043	1.367	6.966	52.73
p-value	.263	.895	.831	.998	.641	0.000

We find that the StoNED models have good explanatory power, as evidenced by the high coefficients of determination, ranging from 0.87 to 0.98. For the robustness tests, we find that we can reject the null hypothesis of normality for the residuals of each of our models. In addition to this, the White tests imply that the majority of our estimated models do not suffer from heteroskedasticity, with only sector 28 (production of plastic and rubber products) identified as heteroskedastic. As stated by Kuosmanen et al. (2014, p. 36) this test does not explicitly identify either error component (or both) as heteroskedastic, as it only indicates that the composed residuals are heteroskedastic. This makes the heteroskedasticity correction for the model somewhat cumbersome, as it involves estimating a doubly heteroskedastic model, where both error components (v_i, u_i) are treated as heteroskedastic. Given the relative similarities of the estimated technical efficiencies (cf. next subsection), we propose that even for this sector our model is not significantly biased.

[36] The full White test regression results are suppressed for brevity, but we note that for all our models, none of the regressors received a statistically significant regression coefficient.

5.3. Comparison of estimated output elasticities

The parameter estimates of different inputs in our SFA models can be interpreted directly as output elasticities, or alternatively representing a part of the marginal productivity of the input in question, with certain caveats. Given this interpretation, we compare our estimated parameters to the input-output tables provided by Statistics Finland. Although this comparison remains rough, as the tables represent the contributions of different input elements to output prices, we still propose that this gives us a rough guideline on whether our estimated elasticities are of the right magnitude. Table 16 below presents the average cost contributions of different classes of inputs for years 2008-2011, obtained from Statistics Finland with corresponding inputs selected for physical capital, labor and material inputs.

Table 16: Output cost components (source: Statistics Finland, author's calculations)

Sector code	Industry	Capital	Labor	Materials
10-14	Mining and quarrying	0.161	0.307	0.246
20	Wood and products of wood	0.135	0.351	0.209
21	Paper and paper products	0.165	0.278	0.335
23	Coke and refined petroleum products	0.037	0.061	0.796
24	Chemicals and chemical products	0.095	0.216	0.454
25	Rubber and plastics products	0.103	0.295	0.371
27	Basic metals	0.082	0.185	0.605
28	Fabricated metal products	0.088	0.329	0.352

Comparing the above table to our estimated output elasticities[37] in Tables 5, 7, 9 and 11, we find that for most inputs, our estimations are surprisingly close to the values obtained from the national accounts tables. The most differences we observe is for the labor input, where our estimates are systematically higher – this can possibly be due to the fact that our input data only contains the aggregate measure of FTE personnel, where a more accurate measure of the labor input would be the actual labor hours utilized. The TRE –model estimates seem to be very close to the tabled values, along with the cross-sectional model. The time-invariant model compares a little less favorably to the tabled cost components. For the time-decay model, even though the estimates for technical efficiency are suspect, many parameter estimates are still notably similar to the cost contributions above. We find that for models and

[37] Note that the input-output tables contain only a single aggregate measure for our three mining and quarrying sectors (10, 13, 14), thus for these sectors we have compared the averaged output elasticities over the mining sectors to the table values.

sectors where CRS is found to hold with statistical significance, the estimates are notably close to the tabled values.

5.4. Comparison of estimated technical efficiency

For estimated technical efficiencies, we compare the mean efficiency in each sector produced with different models. Figure 16 below presents the estimated means of the parametric cross-sectional, time-invariant and true random-effects model and the nonparametric StoNED – model. We see that for certain sectors the nonparametric estimates are much in line with our parametric estimations, but differ quite markedly for the quarrying sector (14) and wood manufacturing sector (20). Given the relative robustness of both our model specifications, we propose that where these estimated efficiencies are sufficiently close, either approach to efficiency estimation probably works relatively well, and where the efficiencies differ there is some evidence that one model provides a better fit to our data. Still, even when significant differences are present it is difficult to provide a conclusive recommendation of one model type over the other.

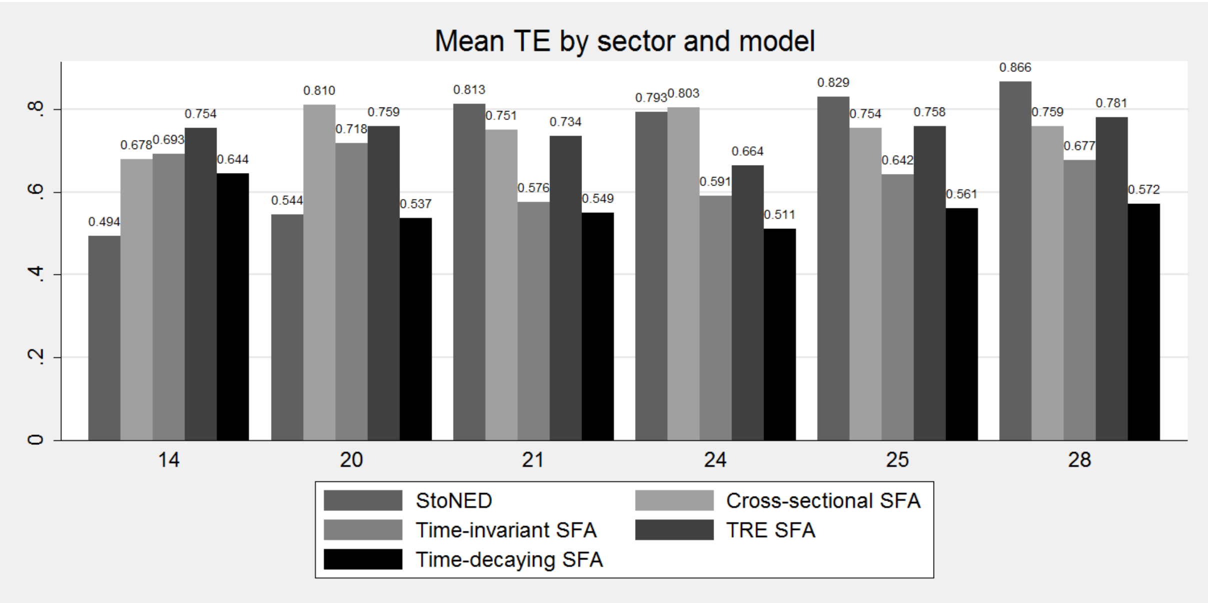


Figure 16: Technical efficiency comparison across all applicable models

We note that intuitively, the StoNED efficiencies for the quarrying (14) and the wood production sector (20) seem artificially low, a peculiar finding when we consider the conclusions of Andor & Hesse (2014), who state that the StoNED –model has a tendency to overestimate the technical efficiency. However, the authors also conclude that StoNED –models have a distinct advantage over parametric ones when the noise-to-signal ratio is high. In our models, for all viable sectors, the estimates of lambda are quite low and this may be one reason for the divergence in estimated efficiencies. One also cannot rule out the effects of the (necessary) random sampling – while in theory, our StoNED sample data should be unbiased, the low efficiencies can also arise as an artefact of the specific data subsampling process. Figure 17 below presents a similar comparison across the parametric SFA models.

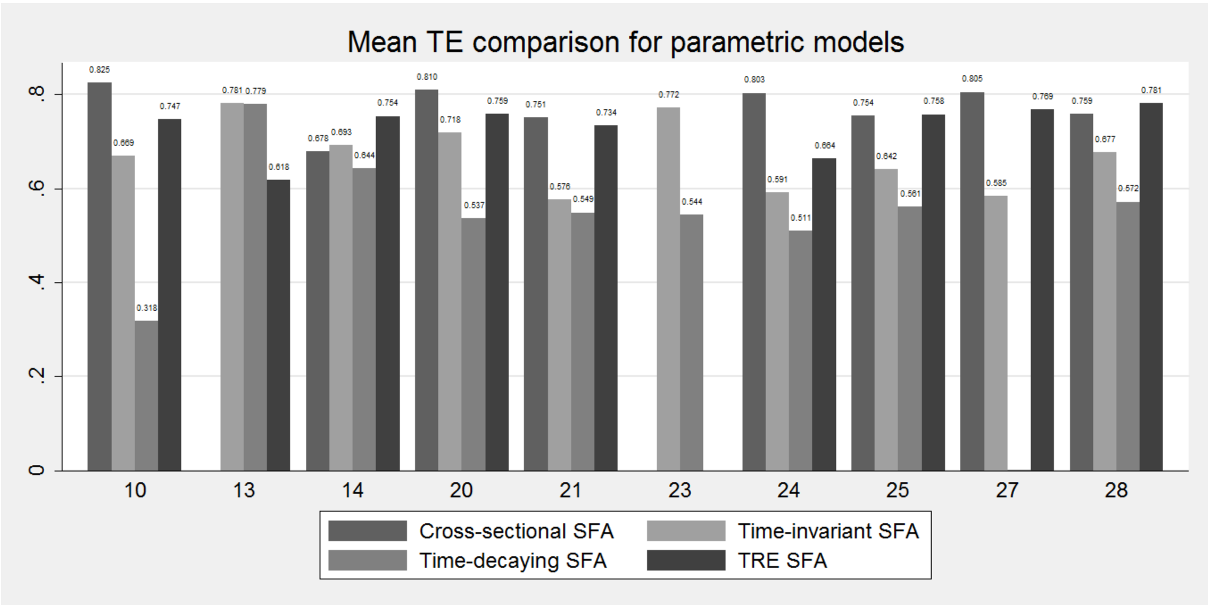


Figure 17: Technical efficiency comparison for parametric models

Among the parametric models, we find that the estimated technical efficiency varies substantially across the model type. We see that the time-decaying model estimates a consistently lower efficiency for all sectors, while the TRE and cross-sectional models exhibit notable similarity between their efficiency estimates. Given the time-decay models somewhat peculiar results of decreasing technical efficiency over time, this is hardly surprising. Another somewhat surprising finding is the similarity between the time-decaying and time-invariant model even in sectors where the time trend (η) is estimated to be statistically significant (i.e. in sector 14 and 24). Considering that these two models rely on quite different assumptions, we would not expect this to happen[38]. Given the specifics of our data and the time period, which includes the financial crisis, a flexible model that allows for efficiency to vary throughout time is most likely to give us unbiased results. To illustrate these differences, we provide a figure of the overall mean estimated efficiencies across all sectors for our three panel models – the bias of the time-decay model can be clearly seen as the overall mean efficiency is estimated to be close to 50%, which even under difficult market conditions is a very low estimate.

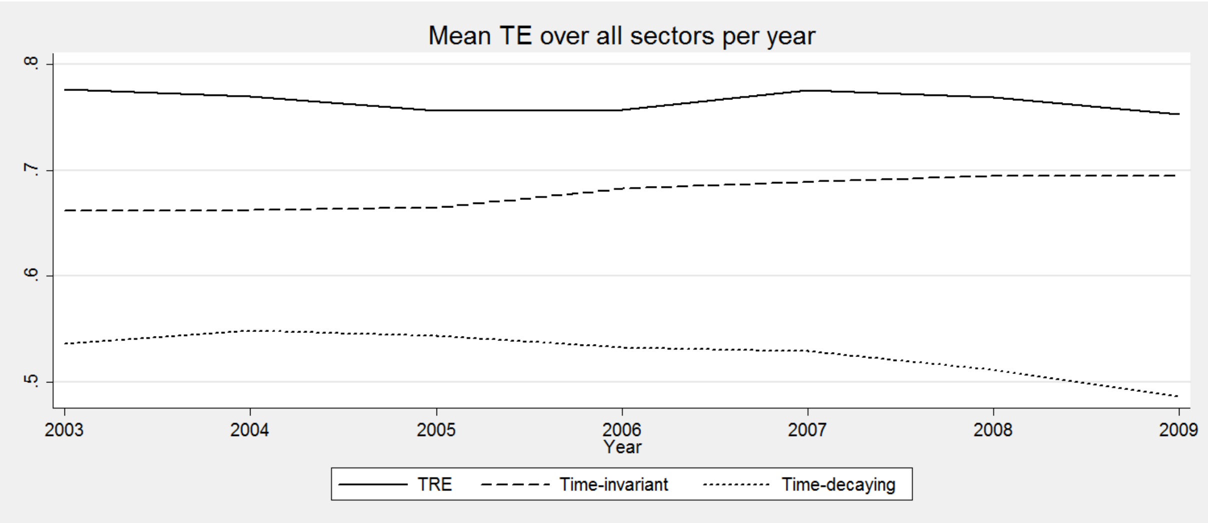


Figure 18: Mean technical efficiency comparison by year and model type

It is likely that the time-invariant model imposes too restrictive assumptions on the technical efficiency, which can be seen in the downward bias of the efficiency estimates. Conversely, when we relax this assumption with the time-decaying model we find that rather than providing us a clearer picture of technical efficiency over time, the estimation results are counterintuitive and still exhibit a rather substantial downward bias. This may be due to the fact that in relaxing the time-invariance assumption we in effect replace it with another assumption of the common efficiency trend within sectors. In comparison to these models, the flexible TRE model seems to provide us with both a good fit to the data and reasonable technical efficiency estimates. Especially telling is the observed effects over time in the TRE

[38] That is to say, similar efficiency measures for these models could be expected when the time trend is estimated as statistically insignificant, as in that case the time-decaying model collapses back into a time-invariant model.

model – we see clearly the impacts of both the 2005 strike and the financial crisis reflected in our industry efficiency estimates.

To further see how our estimated models relate to one another, we estimate Spearman rank correlations between the estimated technical efficiencies both across the models and also specifically for each sector. A particular detail for the TRE model is worth noting here: due to the model specification, we expect the TRE efficiencies to correlate somewhat less with all other models, as it differs markedly from our other models (cf. Greene, 2005b). We compute the correlations between the StoNED efficiency and the cross-sectional SFA model (1), the time-invariant panel SFA model (2), the time-decay SFA model (3) and the true random-effects SFA model (4). Table 17 presents the rank correlations over all sectors, with table 18 providing the sector-specific correlations between our estimations. Stars denote significant correlations at the 1% level.

Table 17: Spearman rank correlations across all sectors

	StoNED	SFA (1)	SFA (2)	SFA (3)	SFA (4)
StoNED	1				
SFA (1)	.3715*	1			
SFA (2)	.0794	.5040*	1		
SFA (3)	-.3260*	.0830	-.0575	1	
SFA (4)	.1387*	.2994*	-.1494*	.2835*	1

We find that the StoNED efficiencies correlate quite weakly with their parametric counterparts when we consider correlation across all sectors. The highest correlation between the nonparametric estimate and the parametric estimate is found between the StoNED model and the cross-sectional SFA model. The nonparametric model also correlates negatively with the Battese-Coelli time-decay model (3), which again may indicate a downward bias for the technical efficiencies estimated with the time-decaying model. The correlation between the true random-effects model and the time-invariant model is negative, which is not surprising, given that these models are based on very different assumptions. The rank correlations between the TRE and both cross-sectional and time-decaying model are significant and of similar magnitude.

Table 18: Spearman rank correlations by sector and model

Sector	10	13	14	20	21	23	24	25	27	28
					SFA (1)					
StoNED	.	.	.3367*	.9165*	.7060*	.	.3114*	.7687*	.	.8043*
SFA (2)	.5940*	.4945*	.7051*	.5843*	.6640*	.1667	.6394*	.6619*	.6360*	.5753*
SFA (3)	.2588*	.8681*	.6553*	.4098*	.5386*	.6000	.4741*	.4931*	.3095*	.4374*
SFA (4)	.4189*	.6044	.4639*	.3677*	.2354*	.	.1198	.2905*	.2662*	.4895*
					SFA (2)					
StoNED	.	.	.2082	.4457*	.5507*	.	.2797*	.3468*	.	.4270*
SFA (3)	-.3830*	.5220	.2654*	-.1016*	.0426	-.3667	.0220	.0108	-.1778	-.1334*
SFA (4)	-.2046*	.1593	-.0155	-.2181*	-.1805	.	-.2758*	-.2404*	-.2099*	-.111*
					SFA (3)					
StoNED	.	.	.2418	.5272*	.2493	.	-.0262	.4793*	.	.2461
SFA (4)	.3698*	.6868*	.3649*	.3307*	.3520*	.	.2176*	.3714*	.3623*	.4093*
					SFA (4)					
StoNED	.	.	.2025	.4167*	.2924*	.	.2323	.4507*	.	.5113*

Compared to the aggregated correlations, the sector-specific correlations give a clearer picture of how the estimated efficiencies relate to one another. Most correlations are found to be significant, with the exception of the TRE model, which again correlates weakly when compared to the other specified models. The nonparametric StoNED correlates most with the cross-sectional SFA model, undoubtedly due to the fact that both methods are estimated as cross-sections. Interestingly, comparing the correlations to the graph of estimated mean efficiencies, we find that even in wood production (sector 20), where the nonparametric efficiency estimate is significantly lower than the parametric estimate the rank correlation between the two classes of models is still very high. This result indicates that in this case, even with divergent results from these two models, the real culprit behind this difference may be our subsampling of the data and not model misspecification. This notion is given more support by our estimated robustness checks for our nonparametric models, which indicate that the nonparametric model is not systematically biased. For the quarrying sector (14), which is the other case where the parametric and nonparametric estimates differ markedly, the rank correlation is estimated to be very low, and in this case we cannot infer that the difference in estimated efficiency would be due to subsample selection alone.

The nonparametric efficiencies correlate weakly with the time-invariant efficiency estimates, and estimated correlations with the TRE model are low, and in some cases insignificant. The parametric estimates exhibit some correlation between the model types, but due to the different specifications, most of the rank correlations are estimated to be quite low. The time-decay model (3) correlates weakly with both StoNED and the time-invariant model, which is to be expected due to the differing assumptions of inefficiency in these models. Significant positive correlations are estimated with both the TRE –model and the simple cross-sectional model, however, for most sectors these remain quite low.

6. Conclusions and discussion

In this thesis we have estimated the technical efficiency of Finnish heavy industry with both parametric and nonparametric models. The estimated efficiencies vary markedly between different models, with the time-decay model and time-invariant models yielding significantly lower estimates than the other models. Our comparisons of the estimated parameters alongside the robustness checks for nonparametric models indicate that our estimated models are not subject to considerable misspecification. The output elasticities of our chosen inputs are found to be mostly comparable both between our chosen models and of similar magnitude than with cost component estimates from Statistics Finland. For the majority of our parametric models, the hypothesis of constant returns to scale production is rejected, with only a few sectors satisfying this hypothesis.

Overall, the estimated technical efficiency in the Finnish industrial sector seems to be rather high, but with significant differences between individual sectors. Mean efficiencies upward of 75% were quite commonly estimated, which given the time period under consideration we consider a very decent performance indeed. The sectors that were most consistently estimated to be operating with high efficiency in most models are the wood and wood products industry (20), the quarrying industry (14), the fabricated metal products industry (28) and metal refining industry (27). Surprisingly, we find that in our analysis, traditional behemoths such as the paper and pulp industry were estimated to be in the middle of the pack in terms of estimated technical efficiency. Among the parametric models, the true random effects model that allows for heterogeneity and flexible development of technical efficiency over time seems to provide the best fit to our data, yielding quite accurate output elasticity estimates and identifying notable trends in the technical efficiencies. The nonparametric and parametric technical efficiency estimates were compared using Spearman rank correlations, and were found to be of similar magnitudes, bar for the quarrying sector where the difference in mean efficiency was significant, but apparently not caused by either model misspecification or subsampling.

We find significant time-varying effects in the technical efficiencies of the sectors under analysis, particularly in models where firm-specific heterogeneity is taken into account. The financial crisis and its broad effects are visible in the efficiencies of the industries under analysis, with most showing a slight downturn during this period. Given our definition of technical efficiency, when faced with a large exogenous demand shock such as the one brought on by the events in 2008-2009, firms may be slow to adjust their input usage accordingly. Particularly for inputs such as labor, frictions and rigidities in the markets may lead to situations where firms operate at a non-optimal input level, thus rendering them technically inefficient by our definition. Encouragingly most sectors seem to have recovered from this shock by the end of 2009. Given the substantial impacts on output volume (cf. Figure 1), *prima facie* intuition would dictate that this shock had also major impacts to the technical efficiencies of the sectors. However, for many sectors the impact of the financial crisis seems to have been not quite that drastic, at least from an efficiency standpoint.

Another event which might have had major impacts on the sectors was the 2005 labor shutout in the paper and forestry industry. Our estimated time-varying technical efficiencies show that the effects of this shutout possibly spread throughout some of the sectors considered in our analysis, with most sectors experiencing a slight decrease in their efficiencies during this period. However, based on our analysis, it is still quite difficult to fully determine whether this event indeed was the major cause of the aforementioned drop in estimated technical efficiency due to many exogenous factors that might have also influenced these sectors during this period.

Our chosen approach of aggregating industries by their operating sector allows us to analyze things from a rather broad perspective. However, interpreting the results should be done keeping in mind certain things. Firstly, our chosen output variable of company turnover is not a totally accurate measure of a firms’ output, given that the firms within each sector most likely produce products that cannot be considered as completely homogenous. Same reasoning applies to our measure of labor input – a more accurate characterization of the input would have been, for instance the actual utilized labor hours. The second consideration relates to the first in the way that we consider the firms production as a single-output process. This is generally a restrictive assumption, as most industrial production is at least in some form joint production. Unfortunately, this is one assumption we cannot do without, at least within the methodology utilized in this thesis. As already touched upon previously, the key question is whether these assumptions can be considered feasible. Given the relative validity of our results, we propose that our characterization of the production process is, if not entirely correct, then at least not systematically incorrect.

A promising avenue of further research would be to incorporate contextual variables in this type of model framework. Particularly the nonparametric models allow for inclusion of a wide spectrum of factors that can affect inefficiency, but as always in econometrics, the true difficulty lies in finding suitable data to evaluate these effects. Specifically concerning our analysis, measures such as managerial practices or factors characterizing the operating environment of the firms would have been very worthwhile extensions to our models. Another interesting approach would be to apply the StoNED model in the panel data context (cf. models presented in Kuosmanen et al., 2014) to see whether nonparametric and parametric estimates differ with regards to the temporal variability of technical efficiencies. For the macroeconomically oriented researcher, following the approach of Koop (2001) is yet another interesting alternative. Especially considering our results regarding the development of technical efficiency in time-varying models, a feasible idea would be to investigate the links between efficiency and the business cycle. The methodological divide between parametric and nonparametric efficiency analysis methods is now rapidly closing with the introduction of true unifying models such as StoNED. The scope of applications for this new approach is sizable, with research interest in StoNED likely to increase in the future as further research is published and the models gain further exposure in the fields of economics and econometrics and the model is developed further.

In conclusion, we find that things are looking somewhat good in the Finnish industrial sector, at least from our chosen viewpoint of technical efficiency. Most sectors are operating on average with a reasonably high degree of efficiency, maintained even during adverse market conditions, such as those encountered from 2008 onwards. In our view, the deep-seated rural heritage of Finland’s industrial production is reflected in our analysis in subtle ways. The high level of expertise and know-how accumulated during the past decades in mining, forestry and metal refining and fabrication is evident in our results, as these sectors are consistently estimated as most efficient in our models.

The latest projections for the Finnish economy (ETLA, 2014) indicate that industrial production is indeed picking up after a rather lengthy period of stagnation and decline. While our analysis covers only a significantly turbulent period during the years 2003 to 2009, the results indicate that the industrial sector as a whole remains quite efficient and is well positioned to benefit when market conditions improve.

7. References

Afriat, S. N. (1972): “Estimation of Production Functions”. *International Economic Review*, Vol. 13, No. 3, 568-598.

Aigner, D. & Chu, S. F. (1968): “On Estimating the Industry Production Function”. *The American Economic Review*, Vol. 58, No. 4, 826-839.

Aigner, D., Lovell, K. & Schmidt, P. (1977): “Formulation and Estimation of Stochastic Frontier Production Function Models”. J*ournal of Econometrics,* Vol. 6, 21-37.

Andor, M. & Hesse, F. (2014): “The StoNED age: the departure into a new era of efficiency analysis? A Monte Carlo comparison of StoNED and the “oldies” (SFA and DEA)”. *Journal of Productivity Analysis*, Vol. 41, 85-109.

Badunenko, O.; Henderson, D. J. & Kumbhakar, S. (2011): “When, where and how to perform efficiency estimation”. *Discussion paper series*, Forschungsinstitut zur Zukunft der Arbeit, No. 5997.

Banker, R. D. & Maindiratta, A. (1992): ”Maximum likelihood estimation of monotone and concave production frontiers”. *Journal of Productivity Analysis*, Vol. 3, No. 4, 401-415.

Banker, R. D. & Maindiratta, A. (1986): ”Piecewise Loglinear estimation of Efficient Production Surfaces”. *Management Science*, Vol. 32, No. 1, 126-135.

Banker, R. D., Charnes, A. & Cooper, W. W. (1984): “Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis”. *Management Science*, Vol. 30, No. 9, 1078-1092.

Battese, G. E. (1991): “Frontier production functions and technical efficiency: a survey of empirical applications in agricultural economics”. *Agricultural Economics*, Vol. 7, No. 3. 185-208.

Battese, G. E. & Coelli, T. J. (1992): “Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India”. *Journal of Productivity Analysis*, Vol. 3, 153-169.

Battese, G. E. & Coelli, T. J. (1988): “Prediction of Firm-level Technical Efficiencies With a Generalized Frontier Production Function and Panel Data”. *Journal of Econometrics*, Vol. 38, 387-399.

Bauer, P. (1990): ”Recent Developments in the Econometric Estimation of Frontiers”. *Journal of Econometrics*, Vol. 46, 39-56.

Cameron, A. C., & Trivedi, P. K. (2005): “Microeconometrics: Methods and applications”. *Cambridge University Press.*

Charnes, A., Cooper, W. W. & Rhodes, E. (1978): “Measuring the Efficiency of Decision Making Units”. *European Journal of Operational Research*, Vol. 2, 429-444.

Cobb, C. W. & Douglas, P. H. (1928): “A Theory of Production”. *The American Economic Review*, Vol. 18, 139-165.

Coelli, T. J, Prasada Rao, D. S, O’Donnell, C. J. & Battese, G. E. (2005): “An Introduction to Efficiency and Productivity Analysis (2nd ed.)”. *Springer Science+Business Media.*

Coelli, T. J. (1995): “Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis”. *Journal of Productivity Analysis*, Vol. 6, 247-268.

Cook, W. D. & Seiford, L. M. (2009): “Data Envelopment Analysis (DEA) – Thirty years on”. *European Journal of Operational Research*, Vol. 192, 1-17.

Cornwell, C. & Schmidt, P. (2008): “Stochastic Frontier Analysis and Efficiency Estimation”. In Mátyás, L. & Sevestre, P. (eds.) *The Econometrics of Panel Data*, Springer-Verlag, Berlin.

Cornwell, C., Schmidt, P. & Sickles, R. (1990): “Production frontiers with cross sectional and time series variation in efficiency levels”. *Journal of Econometrics*, Vol. 46, 185-200.

Debreu, G. (1951): “The Coefficient of Resource Utilization”. *Econometrica*, Vol. 19, No. 3, 273-292.

Elinkeinoelämän Tutkimuslaitos ETLA. (2014): “Suhdanne: Vuosikirja 2014:1”. *Unigrafia Oy.*

Eskelinen, J. & Kuosmanen, T. (2013): ”Intertemporal efficiency analysis of sales teams of a bank: Stochastic semi-nonparametric approach”. *Journal of Banking & Finance*, Vol. 37, 5163-5175.

Fan, Y., Li, Q. & Weersink, A. (1996): “Semiparametric Estimation of Stochastic Production Frontier Models”. *Journal of Business & Economic Statistics*, Vol. 14, No. 4, 460-468.

Farrell, M. J. (1957): “The Measurement of Productive Efficiency”. *Journal of the Royal Statistical Society, Series A (General)*, Vol. 120, No. 3, 253-290.

Fried, H. O., Lovell, K. & Schmidt, S. S. (2008): “Efficiency and Productivity”. In Fried, H. O., Lovell, K. & Schmidt, S. S. (eds.), *The Measurement of Productive Efficiency and Productivity Growth*, 3-91.

Fritsch, M. & Stephan, A. (2004): “What causes cross-industry differences of technical efficiency? An empirical investigation”. *Freiberg Working Papers 2004, No 13.*

Färe, R., Grosskopf, S., Norris, M. & Zhang, Z. (1994): ”Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries”. *American Economic Review*, Vol. 84, No. 1, 66-83.

Gong, B-H. & Sickles, R. C. (1992): “Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data”. *Journal of Econometrics*, Vol. 51, 259-284.

Green, A. & Mayes, D. (1991): “Technical Inefficiency in Manufacturing Industries”. *The Economic Journal*, Vol. 101, No. 406, 523-538.

Greene, W. (2008): “The Econometric Approach to Efficiency Analysis”. In Fried, H. O., Lovell, K. & Schmidt, S. S. (eds.), *The Measurement of Productive Efficiency and Productivity Growth*, 92-250.

Greene, W. (2005a): “Reconsidering Heterogeneity in panel data estimators of the stochastic frontier model”. *Journal of Econometrics*, Vol. 126, 269-303.

Greene, W. (2005b): “Fixed and Random Effects in Stochastic Frontier Models”. *Journal of Productivity Analysis*, Vol. 23, 7-32.

Greene, W. (2003): “Econometric Analysis (5^{th} ed.)”. *Pearson Education*.

Greene, W. (1997): “Frontier Production Functions”. In Pesaran, M. H. & Schmidt, P. (eds.) *Handbook of Applied Econometrics, Volume II: Microeconometrics.* Oxford: Blackwell Publishers.

Greene, W. (1980): “Maximum likelihood estimation of econometric frontier functions”. *Journal of Econometrics*, Vol. 13, 26-57.

Hanson, D. L. & Pledger, G. (1976): “Consistency in Concave Regression”. *Annals of Statistics*, Vol. 4 No. 6, 1038-1050.

Hayashi, F. (2000): “Econometrics”. *Princeton University Press.*

Hildreth, C. (1954): “Point Estimates of Ordinates of Concave Functions”. *Journal of the American Statistical Association*, Vol. 49, 598-619.

Honkatukia, J., Ahokas, J. & Simola, A. (2014): ”Kriisien jälkeen –Suomen talouden rakenteellinen kehitys vuosina 2013-2030”. *VATT Tutkimukset 176*, Valtion Taloudellinen Tutkimuskeskus (VATT).

Honkatukia, J. & Sulamaa, P. (1999): ”Tekninen tehokkuus ja kokonaistuottavuus Suomen sähkönjakeluverkkotoiminnassa 1996 - 1998”. *ETLA Discussion Papers No. 692,* Elinkeinoelämän Tutkimuslaitos (ETLA).

Hyytiäinen, A., Viitanen, J. & Mutanen, A. (2010): ”Suomalaisten itsenäisten sahojen tuotannon tehokkuus 2000-luvulla”. *Working Papers of the Finnish Forest Research Institute, no. 167,* Finnish Forest Research Institute (METLA).

Johnson, A. L. & Kuosmanen, T. (2011): “One-stage estimation of the effects of operational conditions and practices on productive performance: asymptotically normal and efficient, root-*n* consistent StoNEZD method” *Journal of Productivity Analysis*, Vol. 36, 219-230.

Johnson, N. L., Kotz, S. & Balakrishnan, N. (1994): ”Continuous Univariate Distributions”. *John Wiley & Sons*.

Jondrow, J., Lovell, K., Materov, I. & Schmidt, P. (1982): “On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model”. *Journal of Econometrics*, Vol. 19, No. 2, 233-238.

Kim, Y. & Schmidt, P. (2000): “A Review and Empirical Comparison of Bayesian and Classical Approaches to Inference on Efficiency Levels in Stochastic Frontier Models with Panel Data”. *Journal of Productivity Analysis*, Vol. 14, 91-118.

Kontulainen, J. & Spolander, M. (2010): “Suomen talous mittavien haasteiden edessä”. *Kansantaloustieteellinen aikakauskirja*, Vsk. 106, Nro. 3, 306-314.

Koop, G. (2001): “Cross-Sectoral Patterns of Efficiency and Technical Change in Manufacturing”. *International Economic Review,* Vol. 42, No. 1, 73-103.

Koopmans, T. C. (1951): “An Analysis of Production as an Efficient Combination of Activities”. In Koopmans, T. C. (ed.) *Activity Analysis of Production and Allocation*, Cowles commission for Research in Economics, Monograph No. 13, John Wiley & Sons.

Kopsakangas-Savolainen, M. & Svento, R. (2011): ”Observed and unobserved heterogeneity in stochastic frontier models: An application to the electricity distribution industry”. *Energy Economics*, Vol. 33, 304-310.

Koski, H. & Pajarinen, M. (2013): “Subsidies, Shadow of Death and Productivity”. *ETLA Working papers,* No. 16.

Kumbhakar, S. C. (1990): “Production Frontiers, Panel Data and Time-varying Technical Efficiency”. *Journal of Econometrics*, Vol. 46, 201-211.

Kumbhakar, S. C. (1987): “The Specification of Technical and Allocative Inefficiency in Stochastic Production and Profit Frontiers”. *Journal of Econometrics*, Vol. 34, 335-348.

Kumbhakar, S. C., Park, B. U. , Simar, L. & Tsionas, E. G. (2007): ”Nonparametric Stochastic Frontiers: A local maximum likelihood approach”. *Journal of Econometrics*, Vol. 137, 1-27.

Kumbhakar, S. C. & Lovell, K. (2000): “Stochastic Frontier Analysis”. *Cambridge University Press.*

Kuosmanen, T., Johnson, A. L. & Saastamoinen, A. (2014): ”Stochastic nonparametric approach to efficiency analysis: A Unified Framework”. In Zhu, J. (Ed.) *Handbook on DEA* Vol. 2, Springer. (Article in press).

Kuosmanen, T., Saastamoinen, A. & Sipiläinen, T. (2013): ”What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods”. *Energy Policy*, Vol. 61, 740-750.

Kuosmanen, T. (2012): “Stochastic semi-nonparametric frontier estimation of electricity distribution networks: Application of the StoNED method in the Finnish regulatory model”. *Energy Economics*, Vol. 34, 2189-2199.

Kuosmanen, T. & Kortelainen, M. (2012): “Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints”. *Journal of Productivity Analysis*, Vol. 38, 11-28.

Kuosmanen, T., Kortelainen, M., Kultti, K., Pursiainen, H., Saastamoinen, A. & Sipiläinen, T. (2010): ”Sähköverkkotoiminnan kustannustehokkuuden estimointi StoNED –menetelmällä”. *Sigma-Hat Economics.*

Kuosmanen, T. & Johnson, A. L. (2010): ”Data envelopment analysis as nonparametric least squares regression”. *Operations Research*, Vol. 58, No. 1, 149-160.

Kuosmanen, T. (2008): “Representation theorem for convex nonparametric least squares”. *Econometrics Journal*, Vol. 11, 308-325.

Kuosmanen, T. (2006): “Stochastic Nonparametric Envelopment of Data: Combining Virtues of SFA and DEA in a Unified Framework”. *MTT Discussion Papers*, Economic Research Unit, MTT Agrifood Research.

Lee, C-Y., Johnson, A. L., Moreno-Centero, E. & Kuosmanen, T. (2013): “A more efficient algorithm for Convex Nonparametric Least Squares”. *European Journal of Operational Research*, Vol. 227, No. 2, 391-400.

Lee, Y. H. & Schmidt, P. (1993): “A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency”. In Fried, H. O., Lovell, K. & Schmidt, S. S., eds. *The Measurement of Productive Efficiency: Techniques and Applications*.

Mas-Colell, A., Whinston, M. D. & Green, J. R. (1995): “Microeconomic Theory”. *Oxford University Press.*

Meeusen, W. & van Den Broeck, J. (1977): ”Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error”. *International Economic Review,* Vol. 18, No. 2, 435-444.

Mekaroonreung, M. & Johnson, A. L. (2012): “Estimating the Shadow Prices of SO2 and NOX for U.S. coal power plants: A convex nonparametric least squares approach”. *Energy Economics*, Vol. 34, 723-732.

Nurmi, S. (2004): “Essays on plant size, employment dynamics and survival”. *Helsinki School of Economics*, HeSe Print (doctoral dissertation).

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

Richmond, J. (1974): “Estimating the Efficiency of Production”. *International Economic Review,* Vol. 15, No. 2, 515-521.

Romer, D. (1996): “Advanced Macroeconomics”. *McGraw-Hill*.

Schmidt, P. & Lovell, K. (1979): “Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers”. *Journal of Econometrics,* Vol. 9, 343-366.

Statistics Finland (2013): “Official Statistics of Finland”. *Edita Prima*.

Stevenson, R. E. (1980): “Likelihood Functions for Generalized Stochastic Frontier Estimation”. *Journal of Econometrics*, Vol. 13, No.1, 57-66.

Thanassoulis, E., Portela, M. C. S. & Despić, O. (2008): “Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis”. In Fried, H. O., Lovell, K. & Schmidt, S. S. (eds.), *The Measurement of Productive Efficiency and Productivity Growth*, 251-420.

Wang, W. S. & Schmidt, P. (2009): “On the distribution of estimated technical efficiency in stochastic frontier models”. *Journal of Econometrics*, Vol. 148, 36-45.

White, H. (1980): “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”. *Journal of the Royal Statistical Society, Series A (General)*, Vol. 48, No. 4, 817-838.

Winsten, C. B. (1957): “Discussion on Mr. Farrell’s paper”. *Journal of the Royal Statistical Society, Series A*, Vol. 120, No. 3, 282-284.

ELECTRONIC SOURCES:

Statistics Finland (2011): *Components of price for output 2008-2011 (data)*, retrieved from http://193.166.171.75/database/StatFin/kan/pt/pt_en.asp.

Statistics Finland (2011): *Producer Price Indices 2000=100 (TOL 2002) (data),* retrieved from http://193.166.171.75/Database/StatFin/hin/thi/thi_en.asp.

Statistics Finland (2014): *Volume Index of Industrial Output 2003-2009 (data)*, retrieved from http://pxweb2.stat.fi/database/StatFin/teo/ttvi/ttvi_en.asp

Statistics Finland (2012): *Business Registry overview (document)*, retrieved from http://tilastokeskus.fi/til/syr/2012/syr_2012_2013-11-28_fi.pdf

Statistics Finland (2009): *ESA2005 Capital Stock database (data)*, (proprietary)

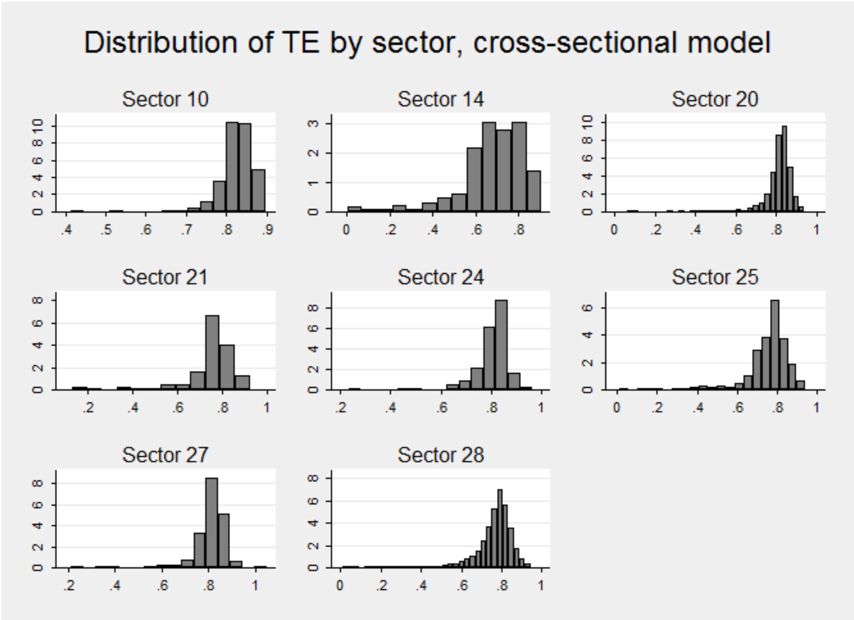
8. Appendix

APPENDIX 1: AIMMS CODE FOR THE CROSS-SECTIONAL CNLS REGRESSION:

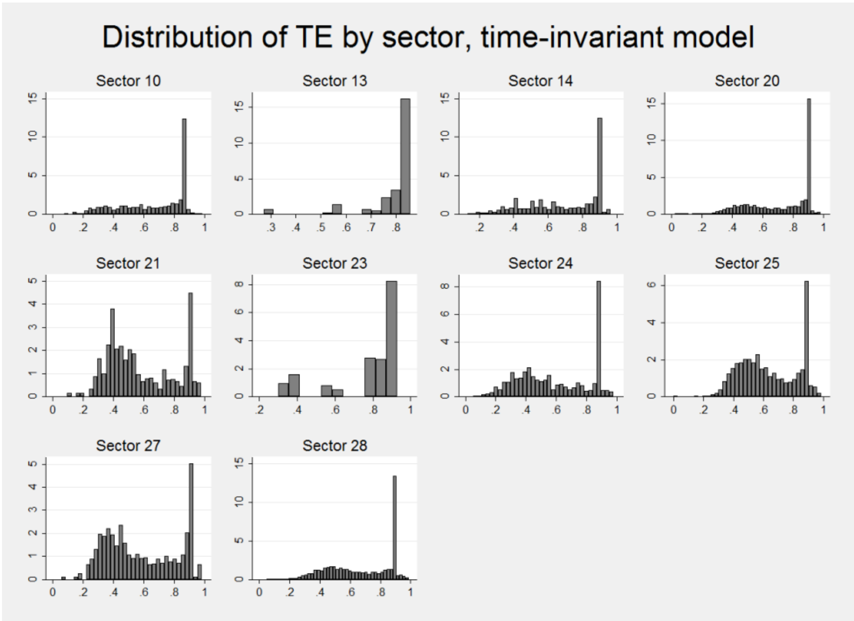
```
MAIN MODEL Cross_sectional_CNLS_estimation

 DECLARATION SECTION

  SET:
    identifier   :  DMUs
    indices      :  i, j ;

  SET:
    identifier   :  Inputs
    index        :  k ;

  PARAMETER:
    identifier   :  X
    index domain :  (i,k) ;

  PARAMETER:
    identifier   :  Y
    index domain :  (i) ;

  VARIABLE:
    identifier   :  z
    range        :  free ;

  VARIABLE:
    identifier   :  Epsilon
    index domain :  (i)
    range        :  free ;

  VARIABLE:
    identifier   :  Gamma
    index domain :  (i)
    range        :  nonnegative ;

  VARIABLE:
    identifier   :  Beta
    index domain :  (i,k)
    range        :  [1e-009, inf) ;

  CONSTRAINT:
    identifier   :  Obj
    definition   :  z=sum(i,Epsilon(i)^2) ;

  CONSTRAINT:
    identifier   :  C1
    index domain :  i
    definition   :  Gamma(i)=sum(k,Beta(i, k)*X(i, k))-1 ;

  CONSTRAINT:
    identifier   :  C2
    index domain :  (i,j)
    definition   :  sum(k,Beta(i, k)*X(i, k))<=sum(k,Beta(j, k)*X(i, k)) ;
```

```
    CONSTRAINT:
      identifier   :  C3
      index domain :  (i)
      definition   :  log(Y(i))=log(Gamma(i)+1)+Epsilon(i) ;

    MATHEMATICAL PROGRAM:
      identifier   :  CNLS
      objective    :  z
      direction    :  minimize
      constraints  :  AllConstraints
      variables    :  AllVariables
      type         :  Automatic ;

    PARAMETER:
      identifier   :  estimatedY
      index domain :  (i)
      definition   :  sum(k,Beta(i, k)*X(i, k)) ;

  ENDSECTION  ;

  PROCEDURE
    identifier :  MainInitialization

  ENDPROCEDURE  ;

  PROCEDURE
    identifier :  MainExecution
    body       :
      solve CNLS

  ENDPROCEDURE  ;

  PROCEDURE
    identifier :  MainTermination
    body       :
      return DataManagementExit();

  ENDPROCEDURE  ;

ENDMODEL Cross_sectional_CNLS_estimation ;
```

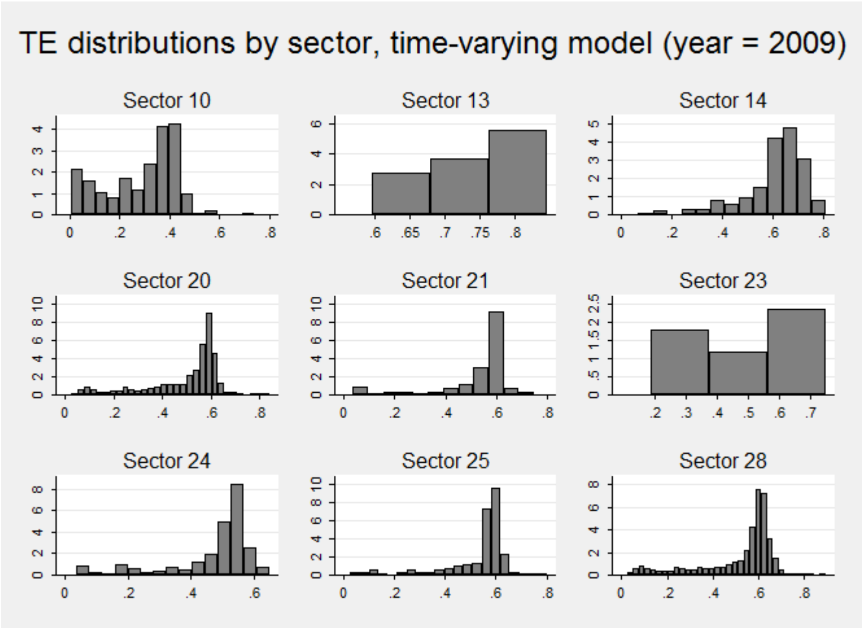
APPENDIX 2: DISTRIBUTIONS OF ESTIMATED TECHNICAL EFFICIENCIES



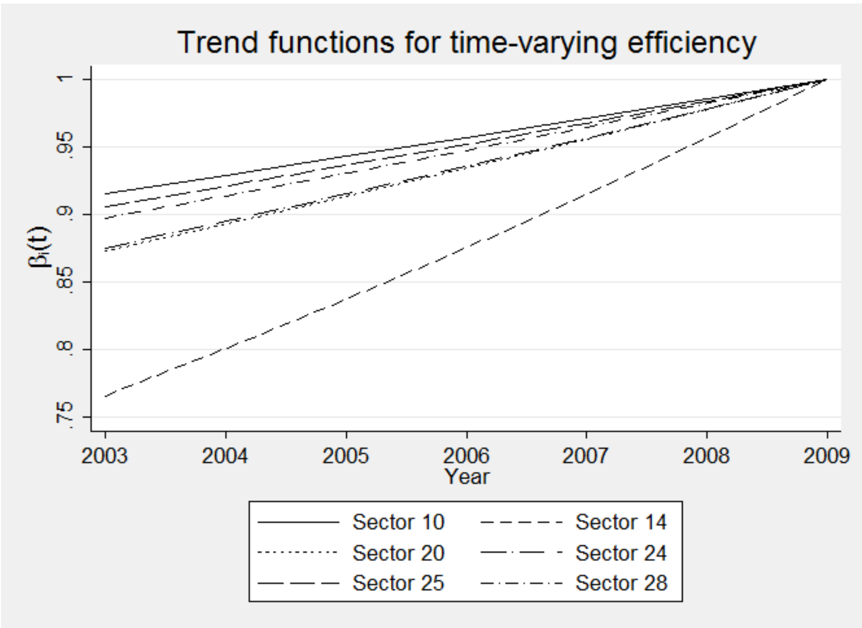
Appendix 2a: Cross-sectional technical efficiency distributions



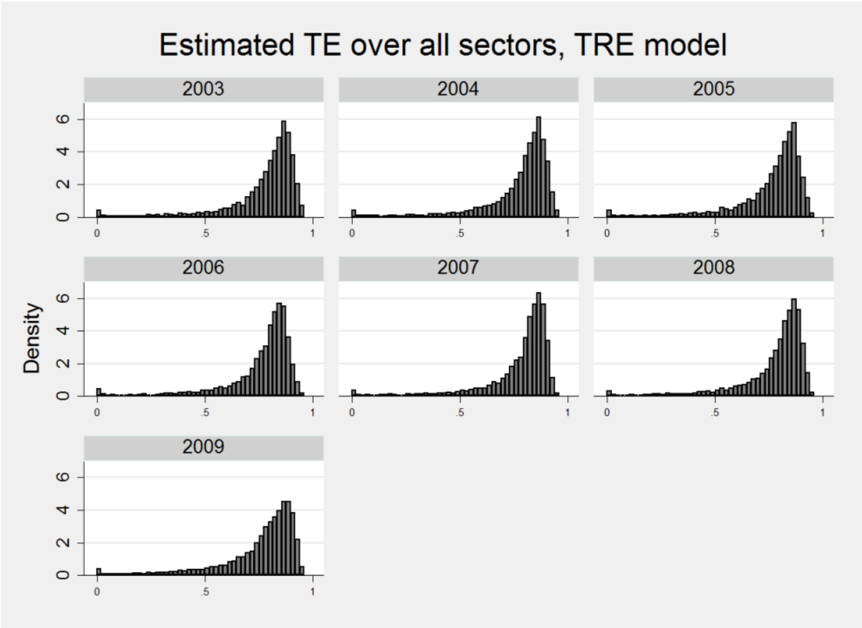
Appendix 2b: Time-invariant technical efficiency distributions



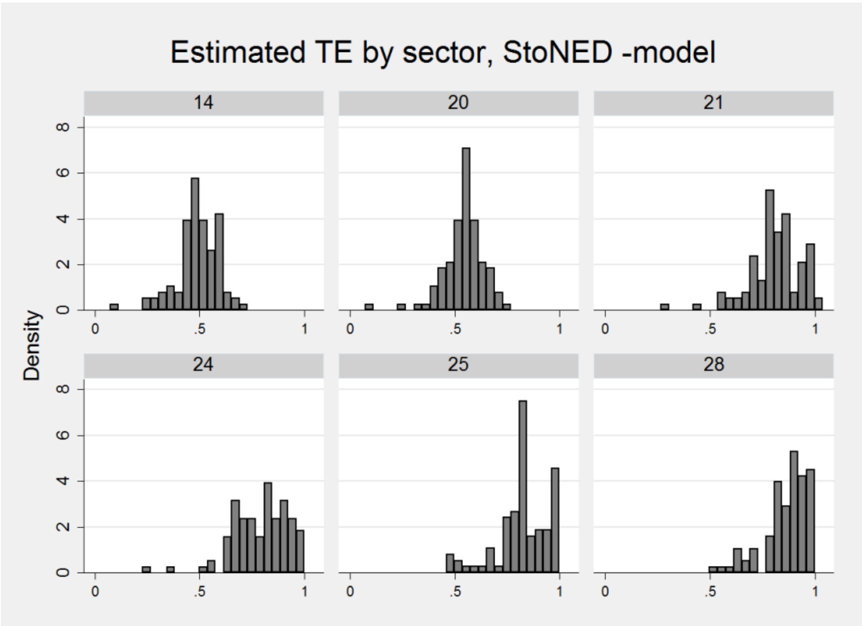
Appendix 2c: Time-varying technical efficiency distributions for 2009



Appendix 2d: Trend functions (cf. Table 9) for selected sectors



Appendix 2e: Technical efficiency distributions in the TRE model



Appendix 2f: Technical Efficiency distributions for the StoNED model