TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Wirtschaftswissenschaften

Lehrstuhl für Volkswirtschaftslehre – Umweltökonomie und Agrarpolitik

Univ.-Prof. Dr. Klaus Salhofer

On the Measurement of Efficiency and Productivity Under Firm Heterogeneity

Magnus Antonius Kellermann

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.) genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Axel Börsch-Supan

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Klaus Salhofer
2. Prof. Dr. Giannis Karagiannis
University of Macedonia, Thessaloniki, Griechenland

Die Dissertation wurde am 21. 07. 2014 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15. 09. 2014 angenommen.

Acknowledgements

During my doctoral studies I have received a lot of support, encouragement, wisdom and friendship from a large group of people. At the end of this journey, I would like to express my sincere gratitude to all of them; they made this thesis possible.

First of all I would like to thank my advisor and teacher, Prof. Klaus Salhofer. He inspired my interest in economics as an undergraduate and supported me in every aspect during my doctoral studies. I am deeply grateful for his valuable feedback and the close collaboration, and especially appreciate his faith in me.

I am also very grateful to my co-advisor Prof. Giannis Karagiannis from the University of Macedonia, Thessaloniki, for the excellent support and the stimulating collaboration which has improved my work substantially. He encouraged me to push my limits and provided tireless guidance throughout my time as a graduate student. I also thank him for making me aware of the importance of the works by Zvi Griliches. Although they did not influence much of the work in this thesis, every single article I read taught me how economic research really should be done.

I owe a special thanks to Prof. Spiro Stefanou for his support and the valuable discussions during my stay at the Penn State University. Moreover, I thank Prof. Axel Börsch-Supan for serving as chairperson at the doctoral examination committee.

I would also like to thank my colleagues Thilo, Roswitha and Gisela at the Environmental Economics and Agricultural Policy Group for the friendly atmosphere. I really enjoyed working with you. My doctoral student colleagues and friends Andreas, Paul and Stefan deserve special mention. They made these years a special, unforgettable time; thank you! I would also like to thank my friends Thanasis and Maria, from the University of Macedonia, for the great time at many conferences and during my stays in Thessaloniki.

Finally, I want to express my gratitude to my parents, Anton and Edigna, and my sisters Edigna and Irmgard for their love and their support in every aspect of life. Thank you so much for everything! I am also indebted to the best parents-in-law you could imagine, Elisabeth and Michael.

Most importantly, I wish to thank my wife and best friend Yvonne. Thank you, for your love and patience and for the encouragement during the last years, especially during some hard times of the doctoral studies.

This thesis is dedicated to Jakob and Stefan, who taught me about the important things in life.

TABLE OF CONTENTS – OVERVIEW

TABLE OF CONTENTS - DETAILS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The thesis addresses firm-heterogeneity in micro-level panel datasets and the necessity to take this heterogeneity into account to obtain unbiased measures of productivity (growth) and technical efficiency in a parametric frontier framework. Various existing and newly introduced econometric specifications of stochastic frontier models are discussed. The focus lies on these models' assumptions about heterogeneity and how they attempt to distinguish heterogeneity from inefficiency. The effects of these assumptions on the results of an analysis of productivity growth are examined. In addition, the thesis describes extensions of existing econometric models to account for unobserved heterogeneity. Two studies tackling empirical questions complete the thesis. Both elaborate on various forms of firm-heterogeneity that can complicate empirical work. In an examination of German breweries, a detailed decomposition of sectoral labor productivity growth into between and within firm effects is provided. In a productivity analysis of Bavarian dairy farms a group- and chain-linked multilateral productivity index is introduced, which allows to compare the productivity of groups of firms over time.

ZUSAMMENFASSUNG

Die Dissertation behandelt Heterogenität in Paneldatensätzen auf Firmenebene. Es wird aufgezeigt, dass eine Berücksichtigung von Heterogenität unerlässlich ist, um Produktivität und technische Effizienz in parametrischen Frontiermodellen unverzerrt messen zu können. Verschiedene gängige und neu eingeführte ökonometrische Spezifikationen von „Stochastic Frontier“ Modellen werden diskutiert. Dabei liegt der Fokus auf den Annahmen der einzelnen Modelle bezüglich möglicher Heterogenität und darauf, inwieweit diese eine Abgrenzung zu technischer Effizienz ermöglichen. Es wird untersucht, inwiefern die Ergebnisse einer Analyse von Produktivitätsveränderungen durch diese Annahmen beeinflusst werden. Darüber hinaus beschreibt die Arbeit Erweiterungen bisheriger ökonometrischer Modelle, die eine Berücksichtigung von Heterogenität ermöglichen. Zwei empirische Studien beschließen die Arbeit. Darin wird auf unterschiedliche Formen von Heterogenität und die dadurch entstehenden empirischen Probleme eingegangen. Anhand einer Untersuchung deutscher Brauereien wird gezeigt, wie Arbeitsproduktivität auf sektoraler Ebene in verschiedene Effekte zwischen und innerhalb von Firmen zerlegt werden kann. In einer Untersuchung der Produktivität bayerischer Milchviehbetriebe wird zusätzlich ein Produktivitätsindex eingeführt, der es ermöglicht, die Produktivität von Gruppen von Firmen über die Zeit zu vergleichen.

1. INTRODUCTION

1.1. Productivity and Efficiency

The terms “productivity” and “efficiency” are commonly used in relation to the technical and economic performance of production units, which may include firms or plants on the micro-level as well as whole industries or countries on the macro-level. Despite the common use of these terms by economists as well as managers, auditors and controllers, some ambiguity exists in regard to the exact meanings associated with them. In this thesis, the term “productivity” denotes the ratio of a firm’s[1] produced output over the respective input used in the production process, where a higher value of this ratio naturally implies a better performance. This concept builds an intuitive framework to monitor the performance of firms over time or compare the performance between firms. A firm that manages to increase its output by more than its input or decrease its input by more than its output from one time period to another will therefore exhibit positive productivity growth. Likewise, a firm compared to another firm in the same period will be labeled as more productive if it uses less input to produce the same amount of output, and vice versa.[2] In contrast, *efficiency* compares the amounts of input and output observed for a particular firm to the respective optimal amounts. That is, the efficiency of a firm is obtained by comparing its present productivity with the best practice frontier productivity for a given level of output (output orientation) or input (input orientation). This highlights the key property of the efficiency concept – the need to determine the technical frontier production potential for each producer.

Measures of productivity and efficiency, i.e., measures of the productive performance of all types of firms, have been used intensively to investigate a great variety of relevant questions concerning the production side of the economy, both on the micro as well as the macro level. Examples for such investigations are the effects of R&D (Griliches and Mairesse, 1991; Crepon et al., 1998; Kumbhakar et al., 2011), policy measures such as production- or export subsidies (Demidova and Rodríguez-Clare, 2009;

[1] Throughout the thesis, the term *firm* encompasses any type of decision-making unit that carries out a production process. In the empirical parts of the thesis we address farms and breweries; however, we can find empirical applications concerning the performance of almost any type of productive entity in many different sectors, such as manufacturing, health care provision, public transport or education. See table 1.1 in Fried et al. (2008) for a list of recent applications.

[2] Balk (2003) distinguishes the comparison of a firm in the two dimensions (over time and cross-sectional) as “monitoring” and “benchmarking”.

Kumbhakar and Lien, 2010; Rizov et al., 2013), regulatory policies (Nicoletti and Scarpetta, 2003; Haney and Pollitt, 2009) or trade liberalization (Pavcnik, 2002; Melitz, 2003; Hossain and Karunaratne, 2004). As Fried et al. (2008, p. 11) note, one of the central challenges that hampers the examination of such sources of productivity and efficiency differentials is “to level the playing field” and “to separate their effects from those of the operating environment”. That is, the empirical researcher has to ensure that any identified measures of productivity or efficiency (change) are not biased due to unobserved or unaccounted factors.

Most empirical studies find efficiency or productivity measures to be widely dispersed in many economic sectors. Studies by Dhrymes (1991), Oulton (1998), and Syverson (2004) are examples of studies on total factor- and labor productivity, while Pitt and Lee (1981), Tzouvelekas et al. (2001) and Kellermann et al. (2011) examine technical efficiency. In the temporal dimension, substantial but unexplained productivity growth has been found since the early beginnings of empirical productivity analysis (Abramovitz, 1956; Solow, 1957). Specific issues that have to be considered to be able to meaningfully interpret such results concern the correct measurement of inputs and outputs, unobserved heterogeneity in production conditions, and misspecifications of the production technology and others. These considerations motivate my work. Specific questions that are covered in this thesis include how unobserved heterogeneity can be taken into account and how this affects the measurement of efficiency and productivity (growth).

1.2. Measurement of Total Factor Productivity

The calculation of a productivity measure is a straightforward task if firms just produce a single output using a single input. However, this is a simplified, unrealistic setting. In most cases, firms use several inputs to produce several outputs, which brings up the need to aggregate those inputs and outputs in “some economically sensible fashion […], so that the productivity remains the ratio of two scalars” (Lovell, 1993 p.3). In this case, the concept of total factor productivity (TFP)[3] can be useful. TFP describes the ratio of (an index of) total outputs produced over (an index of) total inputs used in the production process. It therefore has to be distinguished from partial productivity measures, which can also be used in a multi-input multi-output setting. The most widely used partial

[3] Sometimes TFP is also referred to as MFP (multifactor productivity) “to signal a certain modesty with respect to the capacity of capturing *all* factors’ contribution to output growth” (italics in original) (OECD, 2001).

productivity measure is labor productivity, which is defined as the ratio of (an index of) total outputs over the labor input.[4]

The measurement of productivity is based on the notion that in a given dataset, part of the variance in total output cannot be explained by the variance in total input. A fundamental prerequisite to identify this “residual” of output variance is an accurate measurement of inputs and outputs (Jorgenson and Griliches, 1967) and the knowledge of the substitutional relationship of inputs and outputs embedded in the production technology (Arrow et al., 1961). For the sake of consistency with seminal parts of the literature and simplicity in notation, I begin with the production function as representation of a single-output multiple-input production technology with a Hicks-neutral shift parameter.

$$Y^t = A^t \, F(\boldsymbol{X}^t) \tag{1.1}$$

The function $F(\cdot)$ describes the properties of the production technology, i.e., the ways in which the inputs in $\boldsymbol{X}$ are combined to produce the maximum feasible output Y in period t. The total factor productivity of a firm in time period t is then defined as

$$TFP^t \equiv A^t = \frac{Y^t}{F(\boldsymbol{X}^t)}. \tag{1.2}$$

By taking logarithms[5] of the left- and right-hand sides, equations (1.1) and (1.2) can be rewritten as follows:

$$y^t = a^t + f(\boldsymbol{x}^t) \tag{1.3}$$

and

$$tfp^t \equiv a^t = y^t - f(\boldsymbol{x}^t) \tag{1.4}$$

where the lower-case letters denote the natural logarithm of the variables, i.e., tfp^t is the log-measure of total factor productivity TFP^t in time period t. Equations (1.1) to (1.4) are used as departure points for further investigations.

1.2.1 Productivity growth

Starting from equation (1.1), we can write the representation of the production process in the precedent time period s as $Y^s = A^s \, F(\boldsymbol{X}^s)$. This specification indicates that changes in the output can occur due to input changes and through the neutral shift component A, but

[4] The major advantages of partial productivity measures such as labor productivity are that they are simple to calculate (no aggregation of inputs required) and to interpret. However, the growth rate of labor productivity usually reflects the combined effects of productivity changes and the substitution of labor through other inputs.

[5] Throughout the thesis, I only use the natural logarithm.

the production technology $F(\cdot)$ remains unaltered. Changes in total factor productivity over the periods t and s can now be written in the following way:

$$\frac{TFP^t}{TFP^s} = \frac{A^t}{A^s} = \frac{Y^t/F(\boldsymbol{X}^t)}{Y^s/F(\boldsymbol{X}^s)} \tag{1.5}$$

The first attempts to get a grasp on a measure of productivity and its development over time for a whole economy based on aggregated time-series data can be found in papers by Tinbergen (1942), Abramowitz (1956) and Solow (1957).[6] Following Solow (1957) and assuming a continuous time passage, full technical efficiency and Hicks-neutral technical change, equation (1.3) can be totally differentiated with respect to time:

$$\dot{y} = \dot{a} + \sum_{k=1}^{K} \frac{\partial f(x_k^t)}{\partial x_k^t} \dot{x}_k \tag{1.6}$$

where the dots over a variable indicate its continuous growth rate and the subscript k denote different inputs in use. Assuming further homogeneity of degree one in $f(\cdot)$, cost-minimizing behavior without allocative inefficiency and perfect competition on input and output markets, (i.e., input factors are paid their marginal products), we can substitute the output elasticities by the inputs' factor shares, and equation (1.6) can be rewritten in the following way:

$$\dot{y} = \dot{a} + \sum_{k=1}^{K} s_k \, \dot{x}_k \tag{1.7}$$

In equation (1.7), s represents the input factor shares, which are used to non-parametrically estimate the slope of the production function at the observed input-output combinations.

$$s_k = \frac{\partial f(x)}{\partial x} = \frac{W_k X_k}{YP} = \frac{W_k X_k}{\sum_{k=1}^{K} W_k X_k} \tag{1.8}$$

where W_k is the price of the k-th input, and P is the output price. Slightly rearranging equation (1.7) in regard to (1.4) yields the familiar growth accounting formulation put forth by Solow (1957), Jorgenson and Griliches (1967) and others.

$$t\dot{f}p \equiv \dot{a} = \dot{y} - \sum_{k=1}^{K} s_k \, \dot{x}_k \tag{1.9}$$

[6] This short list is certainly not complete. For a historical overview on the early developments of this field see Griliches (1996) and Hulten (2001).

This continuous framework corresponds to a Divisia index of productivity growth (see, e.g., Richter, 1966). The "residual" of TFP growth results from the percentage growth in (aggregated) output, which remains unexplained by the percentage growth in the aggregated inputs. Subject to the several assumptions stated above, Solow's approach allows a separation of the sources of output growth in movements along the production function $f(\cdot)$ due to the changes in input use and the residual $\dot{a}$ itself, representing "*any kind of shift* in the production function (italics in the original)" (Solow, 1957), which are then interpreted as "technical change". Note that we do not have a superscript t in the factor shares s_k and in equation (1.8). This emphasizes the basic case of neutral technical change with time-constant production elasticities. However, the extension to the non-neutral case was already mentioned by Solow and eventually introduced by Brown and Popkin (1962). It basically allows for changes in the production elasticities over time and accounts for the output changes through factor augmentation.

An important contribution to the measurement of productivity in a non-parametric framework was by Caves et al. (1982). The formulation in equation (1.9) denotes instantaneous productivity changes. However, the data on inputs and outputs is almost exclusively available for discrete time periods. Based on the theory of exact and superlative index numbers by Diewert (1976), Caves et al. (1982) propose a Törnqvist TFP index, which provides a discrete approximation to the continuous Divisia index if the underlying production technology of $f(\cdot)$ is translog. Starting from equation (1.9), the Törnqvist TFP index can easily be derived by taking the first differences of the logarithms of inputs and output instead of continuous growth rates and using arithmetic means of current and lagged factor shares.[7] Then, the total factor productivity change between two subsequent time periods s and t is approximated by:

$$tfp^{st} = (y^t - y^s) - \sum_{k=1}^{K} \frac{1}{2}(s_k^t + s_k^s)\,(x_k^t - x_k^s) \qquad (1.10)$$

Note that we use the superscripts s and t instead of the dot (tfp^{st} vs. $\dot{tfp}$) to indicate the discrete approximation to the continuous change. The Törnqvist TFP index requires a point of reference for its construction. In a time-series application, the choice of the reference point is naturally the preceding time period, with the first time period in the dataset as a starting point.

[7] The Törnqvist index allows for non-neutral technical change

In bilateral applications, the Törnqvist index in (1.10) can be used for base-invariant comparisons of two firms. In that case, the superscripts s and t would denote two different firms (in the same time period) instead of the same firm in two consecutive periods. In a multilateral comparison (with more than two units to be compared and no natural ordering as in a time-series case), however, this index is not base-invariant, and no natural candidate for a reference point exists. Caves et al. (1982a) provide a solution by defining a hypothetical representative firm (h) that is used as a reference point. This hypothetical unit is constructed using the arithmetic means of the cost (and revenue) shares and of the log inputs and outputs of all units in a cross-section. Hence, a multilateral Törnqvist productivity index can be written as follows:

$$tfp^{ih} = \left(y^i - \bar{y}\right) - \sum_{k=1}^{K} \frac{1}{2}\left(s_k^i + \overline{s_k}\right)\left(x_k^i - \overline{x_k}\right) \tag{1.11}$$

where a bar over a variable indicates its arithmetic mean over all firms in the sample, i.e., $\bar{z} = \frac{1}{N}\sum_i^N z^i$. The main advantage of this index is that it is transitive because all firms are compared indirectly by relating them to the same hypothetical reference firm. A disadvantage of the index is its sample dependency. As the sample changes, the hypothetical firm has to be recalculated and by association the productivity measure of all firms. Good et al. (1997) and Delgado et al. (2002) extend the multilateral TFP index for comparisons over time and across groups. In chapter 5 of this thesis, I combine both approaches and introduce a group- and chain-linked multilateral TFP index.

1.2.2 The econometric approach

An alternative to the non-parametric index number approach is the econometric estimation of the parametric representation of the production technology $y^t = a^t + f(\boldsymbol{x}^t)$. Then, the production elasticities can be derived directly from the parameters of the estimated production function as $\varepsilon_k = \frac{\partial f(\boldsymbol{x}^t, t)}{\partial x_k}$, and the TFP level can be calculated from $TFP^t = \exp\left(y^t - f(\boldsymbol{x}^t)\right)$. One of the major advantages of the econometric approach concerns the data, as information on prices are not required, which is difficult to come by in many empirical applications. On the other hand, we only need two observed data points for the index number calculations, whereas the econometric estimation of the firms' technology requires a large number of observations; however, with the currently increasing availability of firm-level panel data sets, this aspect has lost part of its pressing relevance.

Furthermore, the econometric estimation of the elasticities also allows for the relaxation of some of the aforementioned assumptions on market structure and optimizing behavior, which are needed to approximate the input factor elasticities by their cost shares, as in equation (1.8). This in turn enables researchers to identify sources of productivity growth other than pure technical change. The effects of economies of scale were discussed by Brown and Popkin (1962) in the primal framework and Otha (1975) in the dual framework. Other components include the effects of market power (Hall, 1988; 1990; Denny et al., 1981; Basu and Fernald, 2002) capacity utilization and adjustment costs (Morrison, 1992; Prucha and Nadiri, 1996). Another strand of the literature attempted to accommodate deviations from efficient production in the measurement of productivity growth. Nishimizu and Page (1982) distinguished effects of technical change and technical efficiency change, while Bauer (1990) also allowed for allocative inefficiency and a scale effect. The effects of non-constant returns to scale and technical inefficiency on productivity growth are discussed below.

However, the advantages associated with the econometric estimation all come at some cost. First of all, a functional form for the representation of the production technology has to be specified. Several different functional forms have been proposed and discussed in the literature (see, e.g., Fuss et al., 1978, Berndt and Khaled 1979, Chambers 1988, Giannakas et al., 2003). Two of the most common functional forms in empirical work are the "transcendental logarithmic" (Christensen et al., 1973), typically abbreviated as "Translog", which is a generalization of the well-known "Cobb-Douglas" functional form. Fitting a function of the Cobb-Douglas form to the data yields fixed output elasticities across all data points. In contrast, the Translog form is flexible and allows for varying output elasticities. As Lau (1978) notes, the appealing feature of flexibility can cause major issues concerning the assumptions of monotonicity and (quasi-) convexity. However, those assumptions have to be met to exploit the duality theorems of production theory and for a meaningful economic interpretation of estimated elasticities.[8] Hence, the specification of a flexible functional form requires either to impose restrictions on the estimated parameters to ensure the validity of the production technology or to test the theoretical consistency after the estimation at all data points.[9] Other concerns relate to the

[8] For example, violations of monotonicity cause incorrectly signed elasticities. As a consequence, a firm could improve its (measured) productivity by increasing its input usage with fixed outputs.

[9] O'Donnell and Coelli (2005), Sauer et al. (2006) and Henningsen and Henning (2009) discuss implications of violations of the monotonicity and curvature conditions, in particular in a stochastic frontier framework and provide possible solutions.

reliability of the econometric estimates themselves. Endogenous decisions on input factor usage by a firm that observes its productivity and/or is aware of its probability to cease production give rise to the so-called "simultaneity" and "attrition" bias. Unobserved price dispersion across firms may also bias the results of the econometric models. A broad body of literature has emerged around these issues, providing a theoretical analysis of their effects and a variety of techniques to address them. Marschak and Andrews (1944), Hoch (1962) and Mundlak and Hoch (1965) discuss why OLS single equation estimates of production functions may be biased and offer some alternatives, such as fixed-effects estimation (see also Mundlak, 1961), while Zellner, Kmenta and Dreze (1966) provide conditions under which the OLS estimates are tenable. Further approaches include dynamic panel models (Chamberlain, 1982; Griliches and Hausman, 1986; Blundell and Bond, 2000). More recently, Olley and Pakes (1996) approached the simultaneity problem from a completely different angle. One major contribution of their article is the attempt to find a proxy for the unobservable (to the econometrician, but not to the firm) productivity component of the error term. They use the firms' observed investment decisions as a proxy. In addition, they also approach the attrition bias related to firm exits. Their proxy approach was pursued by Levinsohn and Petrin (2003), Wooldridge (2009) and more recently by De Loecker (2011), who, following Klette and Griliches (1996), also accounts for unobserved price dispersion. Because the main focus of this thesis lies in the parametric methods that take technical inefficiency into account, I do not elaborate on this strand of the literature and refer to Griliches and Mairesse (1998) and Van Beveren (2012) for survey articles.

1.2.3 The micro-macro linkage of productivity measures

Regardless of whether index number techniques or econometric estimations are used, having micro-level data available makes it possible to measure firm-specific levels of productivity. Compared to the results from macro-level studies based on time-series data, these data can provide valuable information on the sources and components of the (aggregate) industry productivity level and its development. An aggregated measure of sectoral productivity, for example, can be obtained from a weighted average of the single firms' productivity levels:

$$TFP_I^t = \sum_{i=1}^{N} \varphi_i^t \; TFP_i^t \quad \forall\, t \tag{1.12}$$

where the superscript I signifies the aggregated "industry" productivity for period t, and the subscript i now denotes the individual firm.

Depending on which weights (φ) are used in (1.12), the interpretation of aggregate productivity changes. In the simplest case, each firm is assigned the same weight, and we obtain the arithmetic mean of productivity over all firms in the sample. This measure corresponds to the productivity of the representative average firm in the sample. Olley and Pakes (1996) provide a decomposition of aggregated productivity levels that has been used for the analysis of patterns in sectoral productivity in many empirical studies (Nishimura et al. 2005; De Loecker and Konings, 2006; Eslava et al. 2004; Bartelsman et al., 2009). Assuming a Cobb-Douglas production function, they calculate the firms' productivity level from $TFP = \exp\big(y - f(\boldsymbol{x})\big)$ after econometrically estimating the parameters of $f(\cdot)$. In the following, they decompose aggregated productivity in two components: the average, unweighted productivity of the representative firm and a covariance term that reflects how output shares are allocated across the distribution of differently productive firms:

$$TFP_I^t = \sum_{i=1}^{N} \varphi_i^t \; TFP_i^t = \overline{TFP}^t + \sum_{i=1}^{N} (\varphi_i^t - \bar{\varphi}^t)\,(TFP_i^t - \overline{TFP}^t) \;\; \forall\, t \tag{1.13}$$

This cross-sectional decomposition shows that aggregate productivity only equals the unweighted average productivity if: all firms hold the same share of the industry; all firms exhibit the same productivity; the firms' productivities and shares are uncorrelated. In all other cases, the average productivity over- or underestimates industry productivity. The average measure (over-) underestimates the aggregate measure if larger (smaller) firms are more productive and smaller (larger) firms are less productive. In their initial application of this decomposition, Olley and Pakes (1996) use output (market) shares as weights and identify significant productivity gains in the US telecommunications industry due to a reallocation of output from less productive to more productive firms. It has remained a question of interest how the aggregating weights should be constructed. As a general statement, Bartelsman and Doms (2000) note that the weights should mirror the importance of each firm in the industry. While e.g., Griliches and Regev (1995) and Olley and Pakes (1996) use output shares, others such as Bartelsman and Dhrymes (1998) and Bartelsman et al. (2013) use input shares. Van Biesebroeck (2008) and Fox (2012) discuss the effects of aggregation weights and the resulting effects on the monotonicity and interpretation of aggregated productivity measures.

Moving from productivity levels to productivity growth, the issue of aggregation naturally applies as well. Following Olley and Pakes' (1996) line of thought, the development of the aggregate productivity of an industry can be attributed to the individual productivity change of each firm, shifts in the firms' relative contribution to the industry as well as the entry and exit of firms to and from the industry. Baily et al. (1992), Griliches and Regev (1995) and Foster et al. (2001), among others, propose a decomposition of aggregate productivity growth into these four components. Baily et al. (1992) express the aggregated log industry productivity level as the weighted average of the log productivity measure tfp_{it} similar to equation (1.12) in the following way:

$$tfp_I^t = \sum_{i=1}^{N} \varphi_i^t \; tfp_i^t \;\; \forall\, t \tag{1.14}$$

The sole difference between equations (1.12) and (1.14) is that productivity is defined in logs such that the productivity change between periods is measured in percentage changes. The change in aggregated industry productivity is then:

$$tfp_I^{st} = tfp_I^t - tfp_I^s = \sum_{i=1}^{N} \varphi_i^t \; tfp_i^t - \sum_{i=1}^{N} \varphi_i^s \; tfp_i^s \tag{1.15}$$

From here, as a common starting point, the approaches by Baily et al. (1992), Griliches and Regev (1995) and Foster et al. (2001) keep track of the changes in the individual firms' productivity as well as its contribution (share) to the industry. They differ, however, in their exact construction, mainly in regard to the reference point of comparison.[10] The decomposition by Griliches and Regev (1995) is employed in chapter 4 of this thesis and can be written as follows:

$$\begin{aligned} tfp_I^{st} = {} & \sum_{i\epsilon C} \tilde{\varphi}_{it} \; tfp_i^{st} + \sum_{i\epsilon C} \varphi_{it} \left(\widetilde{tfp}_{it} - \widetilde{tfp}_t^I\right) + \sum_{i\epsilon N} \varphi_{it} \left(tfp_{it} - \widetilde{tfp}_t^I\right) \\ & - \sum_{i\epsilon X} \varphi_{is} \left(tfp_{is} - \widetilde{tfp}_t^I\right) \end{aligned} \tag{1.16}$$

In this decomposition, a tilde over a variable denotes the arithmetic mean of the variable in the present and preceding time period ($\tilde{z}_{it} = (z_{it} + z_{it-1})/2$), and the double superscript st denotes the change from one period to another. Furthermore, C denotes continuing firms (present in t and s), N denotes entering firms (present in t but not in s), and X denotes exiting firms that leave the sample (not present in t but present in s).

[10] See Balk (2003) and Melitz and Polanec (2013) for an extensive discussion. Balk advocates the use of the Griliches and Regev approach, while Melitz and Polanec propose another "dynamic Olley and Pakes" decomposition.

Hence, in equation (1.16), the first two terms on the right-hand side capture the contribution of firms that are continuously observed in the dataset – termed “stayers” by Baily et al. (1992). The first term tracks the individual firm’s productivity growth weighted by the arithmetic mean of the industry share over the respective periods. The second term measures the effect of shifts in the weights of firms with higher or lower than average productivity. The third and the fourth terms contain the contribution of firms that enter or leave the sample. For example, firms with above-average productivity that enter the sample contribute positively to aggregated productivity as well as firms with below-average productivity that drop out of the sample.

1.3. Technical Efficiency

The term “efficiency” in this thesis denotes observed productivity over the maximum productivity, which is technically feasible. This includes an output expanding-perspective as well as an input-conserving perspective. Hence, an output-oriented efficiency measure that complies with this definition can be written as observed output over maximal output with a given level of input. The respective input-conserving measure is then defined by minimum input over observed input with a given level of output. This notion of technical efficiency dates back to a series of papers from the 1950s. Koopmans (1951 p. 60) provided a formal definition of technical efficiency as follows: “A possible point […] in the commodity space is called efficient whenever an increase in one of its coordinates (the net output of one good) can be achieved only at the cost of a decrease in some other coordinate (the net output of another good)”. Putting this definition in context with the use of inputs in the production process, Lovell (1993 p. 10) describes an inefficient producer in less technical words: “Thus a technically inefficient producer could produce the same outputs with less of at least one input, or could use the same inputs to produce more of at least one output”. The works by Debreu (1951) and Shepard (1953), who theoretically derived output- and input-oriented distance measures, and a remarkable paper by Farrell (1957), who proposed specific measures of technical and allocative efficiency, are generally viewed as the starting points of the efficiency literature. Farrell was also the first to empirically calculate measures of efficiency in an application to agricultural production in the US. For surveys on the historical developments of the efficiency literature, I refer to Fried, Lovell and Schmidt (2008) and Kumbhakar and Lovell (2000).

To allow for technical inefficiency, equation (1.1) is rewritten in the following way:

$$Y^t \leq A^t\, F(\boldsymbol{X}^t) \tag{1.17}$$

The Farrell-Debreu output-oriented measure of technical efficiency for this single-output case is then defined as (Kumbhakar and Lovell, 2000 p. 46):

$$TE_O^t = [\max\{\phi\colon \phi Y^t \leq A^t\, F(\boldsymbol{X}^t)\}]^{-1} \tag{1.18}$$

Equation (1.18) can be rearranged to:

$$TE_O^t = \frac{Y^t}{A^t\, F(\boldsymbol{X}^t)} \tag{1.19}$$

Following equation (1.18 and 1.19), $0 < TE_O^t \leq 1$ provides a measure of the shortfall of observed output Y from the technically feasible output, predicted by the production technology for a given level of input $Y^* = A\,F(X)$. Figure 1-1 illustrates this output-oriented measure of technical efficiency.

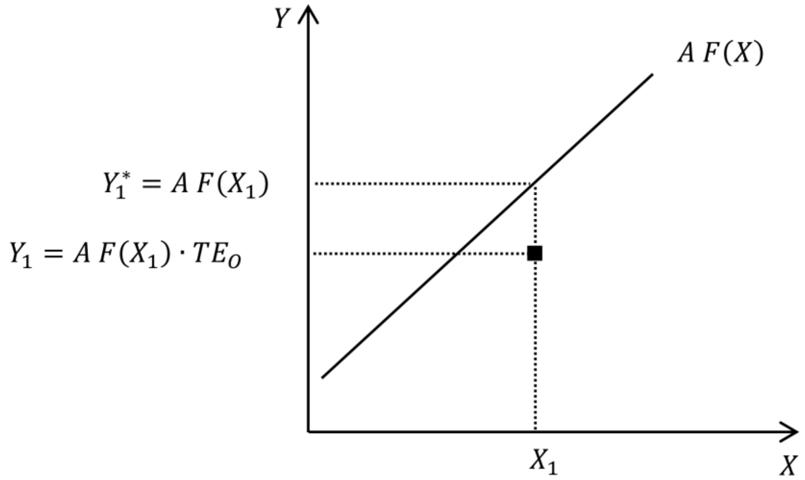


Figure 1-1: Output-oriented measure of technical efficiency in the single-input, single-output case.

In figure 1-1, $A\,F(X)$ represents the production frontier of the technology $F(\cdot)$, and the ■ represents the observed output Y_1 of firm "1" at input level X_1. The output-oriented measure of technical efficiency for a specific firm is the ratio of observed output over the respective frontier output, i.e.:

$$\frac{Y_1}{Y_1^*} = \frac{A\,F(X_1) \cdot TE_{O|1}}{A\,F(X_1)} = TE_{O|1} \tag{1.20}$$

This measure aims at a vertical expansion of output towards the "best practice" (Farrell, 1957) technology, keeping inputs fixed. If a firm produces at the production frontier,

$Y = Y^*$ and consequently, $TE_O = 1$. An input-oriented measure would aim at a horizontal contraction of inputs towards the "best practice" technology, keeping the level of output fixed.

Following Farrell's work, Charnes, et al. (1978) formulated the calculation of efficiency measures as a mathematical programming problem. The resulting data envelopment analysis (DEA) is a well-established and widespread methodology to analyze the technical efficiency of all types of firms, allowing for multiple-output, multiple-input production technologies in a non-parametric framework. See Thanassoulis (2001) and Cook and Seiford (2009) for an extensive treatment.

Some of the first approaches to the measurement of technical efficiency of a cross-section of firms in a parametric framework were conducted by Aigner and Chu (1968) and Timmer (1971). They reformulated the log equation (1.3) to accommodate technical inefficiency as follows:

$$y = f(\boldsymbol{x}) - u \tag{1.21}$$

where u is now a non-negative random variable representing the shortfall of observed output from the frontier production technology, i.e., inefficiency. The respective parameters of $f(\boldsymbol{x})$ were derived through linear programming methods. A series of papers that allowed an estimation of the frontier econometrically led to a significant advancement of the parametric approach to efficiency analysis. Aigner et al. (1977), Meeusen and van den Broeck (1977) proposed and Battese and Corra (1977) applied a model of the form:

$$y = f(\boldsymbol{x}) - u + v \tag{1.22}$$

This model includes an additional component v, which accounts for statistical noise. To estimate the resulting composed error model with maximum likelihood methods, Aigner et al. (1977) specified inefficiency as a non-negative term with $u \sim iid\ N^+(0, \sigma_u^2)$ and the symmetric statistical noise term as $v \sim iid\ N(0, \sigma_v^2)$.[11] Together with the paper by Jondrow et al. (1982), who proposed a way to estimate firm-specific efficiency scores, these papers lay the foundation of the stochastic frontier analysis (SFA) approach. Several excellent surveys on the estimation of stochastic frontier models, their properties and applications exist; see, e.g., Greene (1993), Kumbhakar and Lovell (2000), Murillo-Zamorano (2004), Coelli et al. (2005) and Greene (2008). I also discuss the properties of a variety of stochastic frontier models in chapter 2 of this thesis. Hence, I abstain for now from a

[11] Aigner et al. (1977) and Meeusen and van den Broeck (1977) also proposed an exponential distribution for the non-negative efficiency term.

further treatment of technical details and a broad survey on subsequently developed models.

1.3.1 Productivity growth taking technical efficiency into account

Allowing for the possibility of technical inefficiency we can write the production technology of two adjacent time periods s and t in the following way:

$$Y^s = A^s\, F(\boldsymbol{X}^s) TE_O^s$$

and (1.23)

$$Y^t = A^t\, F(\boldsymbol{X}^t) TE_O^t$$

Rearranging both equations to compare the productivity in the two time periods, similarly to equation 1.5 yields:

$$\frac{TFP^t}{TFP^s} = \frac{Y^t /\, F(\boldsymbol{X}^t)}{Y^t /\, F(\boldsymbol{X}^t)} = \frac{A^t}{A^s} \frac{TE_O^t}{TE_O^s} \tag{1.24}$$

From equation (1.24), we obtain a simple expression that shows that productivity growth can be decomposed into two sources. The first ratio on the right-hand side represents the change in the production frontier, i.e., technical change, while the second ratio represents the change in technical efficiency.

To decompose total factor productivity growth within a parametric framework, the way in which technical change and technical efficiency can vary over time has to be specified. The most common form used to account for technical change is the introduction of a time trend in the production function (first applied by Tinbergen (1942)), often augmented by a squared trend variable to allow for non-linearity over time and interaction terms of the trend variable and inputs to allow for non-neutral technical change (e.g., Gollop and Jorgenson, 1980). Yearly dummy variables allow for flexible but purely neutral technical change; however, Baltagi and Griffin (1988) introduce a flexible general index of technical change, which may be non-neutral and scale-augmenting. Numerous different specifications of time-varying technical efficiency have been proposed for parametric frontier models, which also include the use of linear and non-linear trends as well as time dummies (see, e.g., Kumbhakar and Lovell (2000) and Karagiannis and Tzouvelekas (2010)).

A generalized representation of the production technology that allows for technical inefficiency and non-neutral and scale-augmenting technical change can be written as follows:

$$Y^t = G(\boldsymbol{X}^t, t)TE_O^t \tag{1.25}$$

Following Solow's (1957) theory of growth, Bauer (1990) builds on the work by Denny et al. (1981) and Nishimizu and Page (1982) to decompose total factor productivity growth (see also Lovell, 1996). After taking logs of both sides, the total differential of equation (1.25) with respect to time yields:

$$\dot{y} = \sum_{k=1}^{K} \frac{\partial g(x_k^t, t)}{\partial x_k^t} \dot{x}_k + \frac{\partial g(x_k^t, t)}{\partial t} + \frac{\partial \ln TE_O^t}{\partial t} \tag{1.26}$$

By subtracting aggregate input growth denoted by $\sum_{k=1}^{K} s_k \, \dot{x}_k$ and rearranging, we yield the following expression for TFP growth:

$$\dot{tfp} = \sum_{k=1}^{K} \left(\frac{\partial g(x_k^t, t)}{\partial x_k^t} - s_k \right) \dot{x}_k + \frac{\partial g(x_k^t, t)}{\partial t} + \frac{\partial \ln TE_O^t}{\partial t} \tag{1.27}$$

Assuming allocative efficiency and competitive input markets, equation (1.27) can be rewritten as

$$\dot{tfp} = \frac{\partial g(x_k^t, t)}{\partial t} + \frac{\partial \ln TE_O^t}{\partial t} + (\varepsilon - 1) \sum_{k=1}^{K} \left(\frac{\varepsilon_k}{\varepsilon} \right) \dot{x}_k \tag{1.28}$$

where $\varepsilon_k = \frac{\partial g(x^t, t)}{\partial x_k}$ is the production elasticity of the k-th input, $\varepsilon = \sum_{k=1}^{K} \varepsilon_k$ is the scale elasticity and $s_k = \frac{\varepsilon_k}{\varepsilon}$ (Chan and Mountain, 1983). This decomposition equals the one derived by Bauer (1990) and allows the decomposition of total factor productivity growth into technical change, changes in technical efficiency and a scale change effect.[12] If all firms are technically efficient or if technical efficiency is time-invariant and the technology exhibits constant returns to scale, TFP growth again equals technical change.

Färe et al. (1994) developed an almost equivalent decomposition of TFP growth based on the non-parametric Malmquist index. Empirically, they derive the distance measures required to calculate the index using the Data Envelopment approach. However, the Malmquist productivity index can also be calculated based on econometrically estimated parametric frontier technologies. Balk (2001), Fuentes et al. (2001), Orea (2002)

[12] Having data on input prices W_k available, an additional component, namely allocative efficiency, can be identified (Bauer, 1990; Lovell, 1996).

and Coelli et al. (2005) provide extensive discussions on this approach; see Newman and Matthews (2006) and chapter 5 in this thesis for empirical applications. The Malmquist index and the derivative based approach yield almost identical results if they are applied to the same parametric representation of a production technology.

1.3.2 Efficiency and heterogeneity

The notion of technical inefficiency usually raises suspicion among neoclassical economists, as it implies the opportunity of a "free lunch". In the discussion following Farrell's presentation of his 1957 paper, several issues were immediately recognized, which are still relevant in efficiency analysis, such as the following: i. which inputs to include; ii. how to aggregate inputs; iii. how to distinguish between the heterogeneity of inputs and production conditions and inefficiency; iv. how to account for different production technologies; v. how to account for random productivity shocks, such as weather (Farrell, 1957; Hall and Winsten, 1959). In his discussion of the problems and possible benefits of efficiency analysis, Lovell (1993) refers to Knight (1933), who noted that truly taking all inputs and outputs into account would inevitably cancel out all productivity dispersion between firms.[13] Stigler (1976) and de Alessi (1983) also follow this line of thought in their criticism of the work on X-inefficiency by Leibenstein (1966, 1976). Rather, they assign any "perceived" productivity dispersion to an incomplete model specification with missing or non-homogeneous inputs and outputs, unobserved constraints and an incorrectly specified objective function but not to a general deviation from maximizing behavior, as proposed by Leibenstein. Førsund et al. (1980) in an early survey article on frontier production functions and more recently, Fried et al. (2008) acknowledge these objections; however, they also emphasize the benefits and potential uses of technical efficiency as a "partial" measure of performance:

> "What we have seen are simplified […] models of the firm in which measured performance differentials presumably reflect variation in the ability to deal with the complexities of the real world. Indeed, performance measures based on simplified models of the firm are often useful and sometimes necessary. They are useful when the objectives of producers, or the constraints facing them, are either unknown or unconventional or subject to debate. […] The use […] has proven necessary in a number of contexts for lack of relevant data." (Fried et al., 2008 p. 9).

[13] An identical suggestion made by Zvi Griliches is discussed by Schultz (1959).

Nevertheless, it has become one of the major research objectives on frontier functions to reduce the proportion of unexplained productivity dispersion, which is then called inefficiency. This is also the overarching subject of this thesis. In all four chapters, the impact of heterogeneity on the resulting performance measures is addressed. Drawing on the example of two farmers, which was already used by Stigler (1976) and Førsund et al. (1980), I illustrate the conceptual difference between heterogeneity and inefficiency as follows: We observe two farmers using the same technology and identical amounts of inputs in their production but yield consistently different amounts of the (measured) output corn. Let us consider the case of substantial differences in the land quality of the farms. This would surely be regarded as a lack of homogeneity in inputs or, in other words, heterogeneity in production conditions. In this context it is commonly accepted that the notion of inefficiency to "explain" the shortfall in production is misplaced. The farmer has only limited potential to improve the land quality and faces serious constraints in the form of transaction costs for relocating his business. In addition, it can be argued that (in principle) the quality of land could be measured or approximated in empirical studies. Let us now consider a case in which one of the farmers daydreams occasionally during work and therefore makes technical mistakes. Again, the effect of these mistakes could be assigned to a lower quality of an input, labor this time. Another explanation in line with neoclassical theory would be differences in the farmers' objective functions. The daydreaming farmer might efficiently maximize utility instead of corn output, leading to a misspecification of the simple production frontier model.[14] However, it is highly unrealistic to ever obtain a measure of daydreaming and correct the quality of the labor input, or even to specify a model that allows for an output mix of corn *and* welfare through daydreaming. In that case, the concept of technical inefficiency can be worthwhile to obtain a measure of the shortfall of production from a best-practice production frontier, even if it is impossible to attribute this shortfall to a specific input or to identify the misspecification of the model. Despite the apparent difficulties in defining inefficiency and heterogeneity unambiguously,[15] it has become common practice in the frontier literature to attribute heterogeneity to "factors outside the firms' control" or "environmental factors" and inefficiency to factors that seem manageable by the firm. Examples of this distinction can be found in Perelman and Pestieau (1988), Farsi et al. (2005), Greene (2005), Abrate et al. (2011), and Emvalomatis (2012), among others.

[14] A similar point was also made by Alchian and Kessel (1962) in the context of a comparison of monopolists and competitors.

[15] See also Hall and Winston (1959) for an exhaustive discussion.

In this thesis, three distinct manifestations of heterogeneity are addressed, although not with equal weight. These are heterogeneity in production conditions, heterogeneous technologies and output price heterogeneity.

Heterogeneous production conditions

The idea behind heterogeneous production conditions is to account for disadvantages that cannot, or only with great difficulty, be addressed by the firms. If the differences in production conditions are observed or can be approximated, control variables can be easily included in the production function part of the stochastic frontier model (Sherlund et al., 2002; Coelli et al., 1999). The use of regional dummy variables as shifters of the production function is a widely used example of this practice, especially in empirical applications to the agricultural sector (see, e.g., Hadley (2006) and chapter 5 of this thesis). Another part of the stochastic frontier literature concerns the incorporation of factors, which are assumed to directly affect the efficiency with which inputs are converted to outputs (Kumbhakar and Lovell, 2000). In these models, the variables are not included in the production function part of the model but in the specification of the composed error term. That way, the effects on the mean and the variance of the inefficiency term or the variance of the stochastic noise term can be modeled. The age and education of farmers as well as the location of farms are examples of variables that are frequently assumed to directly affect the inefficiency of an agricultural production process. For the various specifications of such models and the interpretation of their results, see, e.g., Kumbhakar et al. (1991), Battese and Coelli (1995), Wang and Schmidt (2002), Karagiannis and Tzouvelekas (2005), Alvarez et al. (2006), Liu and Myers (2009) and, for a textbook treatment, chapter 7 in Kumbhakar and Lovell (2000).

However, if no information about the production conditions or firm characteristics is available, i.e., heterogeneity is unobserved, model requirements change. Assuming away any heterogeneity, Pitt and Lee (1981) and Schmidt and Sickles (1984) proposed stochastic frontier models for panel data with a one-sided firm effect. Inefficiency in those models contains by construction the effects of all time-invariant differences between the analyzed units. Greene (2005) discusses stochastic frontier models for panel data, which attempt to identify inefficiencies from the skewness of the composed error term, similarly to the original model by Aigner et al. 1977, but includes a firm-specific fixed or random effect in the model. These so-called “true-effect” models separate inefficiency from unobserved heterogeneity based on the assertion that any time-invariant differences in productivity represent heterogeneity rather than inefficiency. Kumbhakar et al. (2014)

further explore the skewness of the firm effects and use multi-step models to identify time-varying inefficiency as well as time-invariant "long-term" inefficiency.[16] In an attempt to mitigate the "heterogeneity bias" in technology parameters caused by the correlation of unobserved factors with input quantities, Farsi et al. (2005) incorporate Mundlak's (1978) formulation into stochastic frontier models. We build on their idea and propose stochastic frontier specifications to take observed and unobserved heterogeneity into account (chapter 3) and apply it in chapters 2 and 4.

Technological heterogeneity

The notion that neglecting potentially heterogeneous production technologies leads to a misspecified model and causes productivity dispersion among analyzed firms was discussed already by Marschak and Andrews (1944) and Griliches (1957). Nelson (1968) and Atkinson and Stiglitz (1969) also abandon the assumption that all firms use one identical production technology. They focus on "localized" technical change, the state of development and the diffusion of technological advances across firms and countries in the passage of time. Drawing on Nelson's paper, Hayami and Ruttan (1970) develop the concept of a meta-production function to model labor productivity growth in the agricultural sector in different countries. A more straightforward way to take technological heterogeneity in an empirical application into account is simply dividing the sample at hand and estimating separate production technologies. Hoch (1962) briefly notes this approach; he suspects technological heterogeneity due to the different locations of farms in southeastern and southwestern Minnesota. More recent empirical applications on diary production distinguish technologies according to characteristics such as the degree of specialization, organic versus conventional production, or milking technology (Newman and Matthews, 2006; Mayen et al. 2010; Alvarez et al. 2012). Moreira and Bravo-Uretha (2010) use a meta-frontier framework in the spirit of Hayami and Ruttan (1970) to analyze dairy production in three South American countries. Karagiannis et al. (2011) estimate a bilateral production frontier to identify the technological differences between organic and conventional dairy production.

If technological heterogeneity is suspected, which cannot be described by single or even multiple characteristics, or if the respective information is unobserved, econometric procedures such as latent class or random coefficient models can be applied (see Alvarez and del Corral (2010), Sauer and Morrison Paul (2013) and Emvalomatis (2012) for recent

[16] Colombi et al. (2011) suggest a one-step maximum likelihood model for the econometric implementation of the short-run and long-run inefficiency specification.

applications on dairy production). In chapter 5 of this thesis, we combine the two approaches in an analysis of German dairy farms in Bavaria. First, we distinguish farms that operate solely on permanent grassland from fodder-crop dairy farms. Second, to account for the remaining unobserved technological heterogeneity, we use a latent class model.[17]

Output price heterogeneity

Accounting data play an important role as a major data source in many empirical studies. However, the wide spread use of costs to approximate inputs and revenues to approximate outputs brings the need to address price changes over time as well as price dispersion between firms. The temporal dimension of price fluctuations is commonly taken into account by deflating all monetary variables using appropriate price indices, which are obtained from external data in most cases (e.g., statistical offices as Destatis in Germany or the Bureau of Labor Statistics in the US). However, whether the attempt to remove temporal price changes from the data is successful depends on the extent to which common industry or sector-based price indices display the overall price development in an actual empirical sample. If the sample that is used by the authority to construct the price index deviates systematically from the researchers' sample, the deflator may fail to represent the unobserved average price development. Similar issues can occur if the index is constructed on a different level of product aggregation or if an index for a seemingly related product has to be used as a proxy if prices for the actual product under consideration are not recorded. However, the researcher cannot assess the application-specific quality of a price index directly if the firms' prices are unobserved in his dataset. Therefore, much of the judgment of whether a deflator at hand is suitable depends on the availability of detailed information on how it was constructed.

Even if the deflator price index perfectly represents the overall average price development in the sample at hand, a common industry or sector-based price index cannot account for price dispersion between firms. The crucial points here are as follows: what does the price dispersion reflect and can it be considered as an *iid* white noise term that is uncorrelated with input and output quantities? If price dispersion reflects actual quality differences in produced products or input factors, the use of deflated revenues and costs is a convenient way of taking heterogeneity in outputs or input factors into account. An

[17] Flexible functional forms such as the translog also allow for a certain degree of technological heterogeneity, as the production elasticities for the individual firms vary according to their level of input use (Sauer and Morrison Paul, 2013).

illustration of this case is the use of deflated revenues as output measures instead of the physical quantity in studies concerning dairy production. Milk is a rather homogeneous product, but its quality can vary according to protein and fat content, which are seldom observed in a dataset. Dairy processors, however, observe the milk quality and adjust prices accordingly (see Reinhard et al., 1999 and Emvalomatis, 2012). An example on the input side is provided by Fox and Smeets (2011), showing the advantage of using the deflated wage bill as a measure of labor input compared to the number of employees or hour-based measures.

However, unobserved quality-adjusted price heterogeneity may have serious adverse effects. The monotonicity property naturally requires the correlation between inputs and outputs to be non-negative. Following the usual assumptions of a downward-sloping demand curve and negatively correlated output quantities and prices, prices and inputs are negatively correlated. These considerations also fit with the results of an empirical analysis of US manufacturers by Roberts and Supina (2000), who consistently find that large producers charge lower prices. Klette and Griliches (1996) discuss the implications of these relationships for the econometric estimation of production and cost functions. They show that the resulting omitted variable bias leads to downward-biased coefficients and scale elasticities in most cases. They also propose a potential solution for the omitted price bias by approximating unobserved price dispersion by the firms' market share. Under the assumption of monopolistic competition and horizontal product differentiation, they integrate a CES demand system in the production function and solve for the firm-level prices. More recently, De Loecker (2011) built on this approach to reduce the omitted price bias in productivity measures. He estimates the resulting model using a semi-parametric proxy estimator (Olley and Pakes, 1996) to account for simultaneity and allows for multi-product firms and segment-specific demand elasticities.

Foster et al. (2008) examine the resulting bias in the calculated productivity measures if the price dispersion between firms is not taken into account by comparing revenue-based and quantity-based TFP. Thereby, they avoid the econometric estimation of a production technology and calculate TFP using index numbers techniques. They find that the use of a common price index for the firms in a sample leads to a substantial bias of TFP measures. Inefficient firms produce less physical output using a given set of inputs, implying higher marginal costs. In an imperfectly competitive environment, firms can pass these costs along, resulting in a firm price above the industry average. Deflated revenues are then an upward-biased measure of the firms' output quantities and consequently lead to an upward-biased productivity measure. Likewise, for highly productive firms charging lower than average prices, deflated revenues underestimate the physical output. Hence, output prices are inversely correlated with physical productivity (see also De Loecker, 2011). Consequently, revenue-based TFP has a lower variation than quantity-based TFP, and the firms appear to be a more homogenous group if deflated revenues are used as output measure. This also translates into an inflation of the estimated technical efficiency scores in frontier applications. In chapter 4 of this thesis, we use firm-specific output prices to obtain a quantity type output measure for a sample of German breweries. In chapter 5, however, we use revenues deflated by a common price index for all firms in the sample to allow for quality differences in the output of dairy farms.[18]

The discussion above shows how important it is to appreciate the theoretical and methodological foundations of the measurement of productivity and efficiency. Numerous studies have provided empirical evidence of unexplained productivity change over time or productivity dispersion between firms. To be able to draw useful conclusions from their results, it is necessary to be aware of the underlying assumptions made, possible shortcomings of the data and the methodology in use, and potential misspecifications of the empirical model. Only then can these partial measures of firm performance be used to further investigate the sources of productivity change over time or efficiency dispersion between firms, and only then it is sensible to test hypotheses, e.g., concerning the effect of policy measures on productivity.

1.4. Outline of the Thesis

The main body of the thesis is composed of four articles that address methodological as well as empirical research questions. In the first article (chapter 2), I discuss various specifications of stochastic frontier models and examine the effect of model choice on the measurement of productivity growth in an empirical application. Thereby, the focus particularly lies in the way in which models account for (unobserved) heterogeneity in the data and how they attempt to distinguish heterogeneity from inefficiency. I elaborate not only on the efficiency scores but also on the estimated representation of the production technology itself, as the corresponding production elasticities and returns to scale

[18] Note that I do not observe firm specific prices on the input side of production. Hence, even measures of "physical" productivity may still contain price effects on the input side, i.e., firms that face higher factor prices will appear to utilize a relatively higher level of inputs and less productive as a consequence. However, as Foster et al. (2008) note, using quantity output, productivity reflects firms' "idiosyncratic cost components, both technological fundamentals and factor prices". See also Ornahghi (2006) and Katayama et al. (2009) for a discussion of the effects of input price differences.

measures are elementary parts of every TFP decomposition. The results show that the relative contribution of the components to TFP growth is quite sensitive to the choice of the econometric model, which brings up the need to select the “right” model. I apply various statistical tests to narrow down the range of applicable models and offer further criteria with which to choose between non-nested models.

In the second article (chapter 3), we follow a suggestion by Farsi et al. (2005; 2005a) and propose two alternative models in the spirit of Mundlak (1978) and the “true” random effects model by Greene (2005). In both of them, we try to improve the ability of the true random-effects models to account for heterogeneity by further enlarging the set of potential correlates to increase the portion of measured heterogeneity and squeeze the impact of heterogeneity bias on the estimated technology parameters and technical efficiency. In addition to the true effect types of models, accounting for endogenous individual effects gives rise to several new variants of standard stochastic frontier models. We present two such variants, making no distributional assumptions that can be viewed as extensions of the fixed and the random effects stochastic frontier models introduced by Schmidt and Sickles (1984) and developed further by Good et al. (1990). We also discuss three models that make distributional assumptions about the inefficiency term, which complement previous attempts by Coelli et al. (1999) and Sherlund et al. (2002) to account for environmental factors in a maximum-likelihood type of stochastic frontier model. We outline the estimation procedures and compare the effectiveness of the discussed models in an empirical application. The results show the ability of the proposed specifications to take observed and unobserved heterogeneity into account and to reduce the respective contamination of the efficiency scores.

In the third article (chapter 4), we investigate the evolution of labor productivity in the German brewing sector. To measure the technical efficiency of the breweries and at the same time account for unobserved heterogeneity, we apply one of the stochastic frontier models as proposed in chapter 3. In addition, we take price heterogeneity into account and use a firm-specific price index to obtain a quantity type measure of output. We provide a method to decompose aggregate industry labor productivity growth into seven distinct components: input deepening, technical change, technical efficiency, scale effect, between-firm reallocation and effects from exits and entry. The first four components measure the productivity growth that takes place within a firm. The latter three components capture industry dynamics. Our results show that within-firm effects and particularly technical change and the scale change effect clearly dominated the effects of industry restructuring.

In the fourth article (chapter 5), we compare the productive performance of dairy farms that operate solely on permanent grassland and dairy farms using fodder crops from arable land. Using a latent class stochastic frontier model, we allow for heterogeneous production technologies and identify more intensive and extensive production systems for both types of farms. Thereby, our notion of intensive vs. extensive dairy production is based on differences in stocking density and milk yield per cow and year. We try to take heterogeneity in production conditions into account by introducing regional dummy variables for various agricultural production areas. To be able to compare the productivity levels and productivity developments of the various groups of farms, we develop a group- and chain-linked multilateral productivity index. Our results show that the intensive classes in both groups of farms are more productive and are also able to increase their productivity to a greater extent over the observed period. We find technical progress to be by far the most important component of TFP growth for all classes. Our calculations of the multilateral Törnqvist index reveal that both the FC classes are more productive than their PGL counterparts. However, our distinction between intensive and extensive classes shows that there are highly productive grassland farms that can keep up with the fodder-crop farms and that those farms are predominantly found in the more intensive class of farms. The more substantial productivity gap exists between the intensive and extensive classes, and this gap widens because of the higher productivity growth rates in the intensive classes.

Each of the chapters in this thesis can be viewed as an independent contribution to the literature in the broad field of productivity analysis; however, they all are connected by addressing the overarching research question of how productivity and efficiency can be measured most accurately in the presence of observed and unobserved heterogeneity.

2. TOTAL FACTOR PRODUCTIVITY DECOMPOSITION AND UNOBSERVED HETEROGENEITY IN STOCHASTIC FRONTIER MODELS

Abstract

This paper examines in an empirical comparison how different econometric specifications of stochastic frontier models affect the decomposition of total factor productivity growth. We estimate nine different stochastic frontier models, which have been in wide use in empirical investigations on the sources of productivity growth. Our results show that the relative contribution of the components to TFP growth is quite sensitive to the choice of the econometric model, which brings up the need to select the “right” model. We show various statistical tests to narrow down the range of applicable models and offer further criteria to choose between non-nested models.

Keywords: heterogeneity, panel data, stochastic frontier, TFP decomposition

JEL Classification: C23, D24, O12

This chapter is based on the article *Total factor productivity decomposition and unobserved heterogeneity in stochastic frontier models* by Magnus Kellermann. The article is accepted for publication at the *Agricultural and Resource Economics Review*. The author of this dissertation is the sole author of the article.

2.1. Introduction

Productivity analysis has become a major instrument to better understand and describe economic developments in many economic sectors (Fried, Lovell, and Schmidt 2008). Departing from the seminal paper by Solow (1957), authors tried to minimize the unexplained residual of technical change and provide a more detailed structure of total factor productivity (TFP) growth. Essential component of such attempts is the estimation of a parametric representation of the underlying production technology.[19] One of the first papers in this regard was Griliches (1964), who integrated education as well as research and extension services as explanatory variables in an aggregate agricultural production function. Denny, Fuss, and Waverman (1981), Nishimizu and Page (1982) and Bauer (1990) relaxed assumptions such as of full technical efficiency and constant returns to scale.[20] By this means, they separated the respective effects from technical change, resulting in a decomposition of TFP growth. In many empirical studies following this approach, the relevant question of interest relates to the relative importance of the contributing factors to total factor productivity growth. Those results then often build the basis for recommendations on regulatory or support policies.[21] Therefore, it becomes crucial to be aware how potentially sensitive those results can be with regard to the choice of particular methods.

The focus of this study lies on the econometric models used to estimate parametric representations of a production technology in a stochastic frontier framework and how different models can influence results in regard to the sources of productivity growth. Consequently, we elaborate not only on the efficiency scores, but also on the estimated representation of the production technology itself, as the corresponding production elasticities and returns to scale measures are an elementary part of every TFP decomposition. We also call attention to the fact that we can only draw inferences from the results of a TFP decomposition, if the underlying estimate of the production technology fulfills the requirements of microeconomic theory. These features distinguish our work from the existing studies which compare the results of different stochastic

[19] Färe et al. (1994) show how TFP growth can be decomposed using non-parametric methods.

[20] Denny, Fuss, and Waverman (1981) and many others also identified further contributing factors to total factor productivity such as market imperfections and allocative efficiency (e.g. Morrison 1992).

[21] Examples are Fan (1991), Brümmer, Glauben, and Thijssen (2002), Saal, Parker, and Weyman-Jones (2007), Key and McBride (2007), Goto and Sueyoshi (2009) and Tovar, Ramos-Real, and de Almeida (2011).

frontier models mainly in regard to efficiency scores (e.g. Ahmad and Bravo-Uretha 1996, Hallam and Machado 1996, and Abdulai and Tietje 2007).

We compare a variety of stochastic frontier models that have been most widely applied in empirical TFP growth studies. We focus in particular on the way how models account for (unobserved) heterogeneity[22] in the data and how they distinguish heterogeneity from inefficiency. It appears that there are no clear-cut criteria available to guide the researcher when choosing “the” appropriate model, since seemingly valid models are not nested altogether, which complicates the choice purely based on econometric specification tests. However, we are able to provide some guidance on how to choose an appropriate econometric model for a specific empirical application.

For our application, we use a data set of just under 1000 dairy farms in an unbalanced panel, covering the years 2000 to 2008. A translog output-oriented distance function is utilized to represent the production technology. In order to make the results of the different models comparable, we keep the data, the specification of variables and the functional form identical across the econometric specifications. Based on the estimated parameters and inefficiency estimates, we decompose productivity growth into the three components that are most commonly found in empirical applications: technical change, changes in technical efficiency and the scale change effect.

2.2. Stochastic Frontier Models

Technical efficiency is considered as the ability of a firm to produce the maximal possible output using a given level of inputs.[23] A firm’s potential inefficiency is the shortfall in observed production as against a best practice frontier. The econometric estimation of a function that represents this maximal possible expansion of output for given inputs is the objective of all models we discuss in this paper. Several excellent surveys on the concepts of (technical) efficiency and the estimation of stochastic frontier models exist (e.g. Greene 1993, Kumbhakar and Lovell 2000); hence, we limit our overview to a short description of the respective models we use in our application and their main properties. Thereby, we focus mainly on the assumptions these models impose on the residual error term, whether

[22] Models that take parameter- or technological heterogeneity into account are excluded from this work. In this regard see works by Tsionas (2002), Orea and Kumbhakar (2004), and Greene (2005) for random parameter and latent class models in the context of stochastic frontier models and Emvalomatis (2011) for a recent application.

[23] This statement corresponds to the concept of output-oriented technical efficiency. Input oriented efficiency aims for the minimal feasible use of inputs to produce a given level of outputs.

they account for heterogeneity between the firms and the way estimates of inefficiency are derived. These properties are summarized in table 2-1.

We start our overview with the pooled model (I) based on the original normal half-normal model proposed for cross-sectional data by Aigner, Lovell and Schmidt (1977) which treats every observation in a sample as independent of each other. Two examples for TFP studies concerning the agricultural sector that used the pooled model are Fan (1991) and Key, McBride, and Mosheim (2008). In order to keep the notation simple we start from a production function, the single-output special case of the output oriented distance function. Assuming a log-linear functional form we can write this model as:

$$y_{it} = \boldsymbol{x}'_{it}\boldsymbol{\beta} + e_{it} \tag{2.1}$$

where $e_{it} = v_{it} - u_{it}$ is a composed error term, y_{it} is the log output, $\boldsymbol{x}$ is a vector of log inputs and $\boldsymbol{\beta}$ represents the vector of all technology related regression coefficients. The subscripts i and t denote firms and time periods, respectively. Model I contains a composed error term e_{it}, where $u_{it} \sim iid\ N^{+}(0, \sigma_u^2)$ is a non-negative term, representing inefficiency (u), while $v_{it} \sim iid\ N(0, \sigma_v^2)$ is a symmetric term that captures statistical noise e.g. from exogenous productivity shocks beyond the control of the analyzed units or measurement errors. Both components of e_{it} are assumed to be uncorrelated with input quantities and each other. Especially the assumption that the firms' inefficiency is uncorrelated with the used input quantities requires further reasoning. If their own technical inefficiency in t is not known to the firms at the time they make their input decisions and the firms maximize expected profits, Zellner, Kmenta, and Dreze (1966) argue that the quantities of variable inputs are largely predetermined and can consequently be uncorrelated with technical inefficiency. The individual efficiency scores of the i-th analyzed unit can be obtained, using the mean (or the mode) of the conditional distribution of u_{it} given e_{it} as point estimator (Jondrow et al. 1982). However, since the variance of the mean (mode) of $(u_{it}|e_{it})$ for each unit is independent of the sample size, efficiency scores cannot be estimated consistently using the pooled model.

Model II is an inefficiency effects model following a concept as initially proposed by Kumbhakar, Ghosh, and MacGuckin (1991). The specification we use in this paper has been formulated by Battese and Coelli (1995) for the use with panel data, which has been used extensively in the analysis of productivity growth. Examples are Yao, Liu, and Zhang (2001), Brümmer, Glauben, and Thijssen (2002), Rae et al. (2006) and Jin et al. 2010. The main feature of model II is the incorporation of exogenous influences on the inefficiency term in a one-step approach.[24] Battese and Coelli (1995) achieve this by assuming the inefficiency term to have a truncated-normal distribution with mean $\mu_{it} = \boldsymbol{z}_{it}'\boldsymbol{\zeta}$ and variance σ_u^2 (see table 2-1). In this context, $\boldsymbol{z}_{it}$ is a vector of observed exogenous variables which may have an influence on the firm's inefficiency and $\boldsymbol{\zeta}$ is the corresponding vector of additional parameters to be estimated. Although model II is designed for the use with panel data, it is not a panel data treatment in the classical sense since inefficiency terms are assumed to be independent over time (Battese and Coelli 1995) and observations of a single firm in different time periods are treated as observations of separate firms (Abdulai and Tietje 2007 p.397) just as in model I.[25] There is an ongoing debate in the efficiency literature which can be traced back to one of the seminal papers on the topic by Deprins and Simar (1989) about the "right place" of these exogenous z-variables. The question on "where do we put the z's" (Greene 2008 p. 154) concerns whether these variables truly explain part of the variation in inefficiency or if they rather pick up heterogeneity and misspecifications of the production technology.[26] An intuitive example for this debate is the use of variables on the education and age of farmers or the farms location found in many agricultural studies (e.g. Battese and Coelli 1995, Tzouvelekas, Pantzios, and Fotopoulos 2001). It can be argued, that these variables should rather enter the production function part in an attempt to reduce heterogeneity and create homogenous measures of the inputs labor and land (Sherlund, Barrett, and Adesima 2002). We do not further elaborate on this question; however, it seems worthwhile to cover these concerns when specifying any stochastic frontier model.

The models III and IV are fixed- and random-effects panel models, developed by Schmidt and Sickles (1984), extended to allow for time-varying technical efficiency by Cornwell, Schmidt, and Sickles (1990).[27] Wu (1995) and Karagiannis, Midmore, and Tzouvelekas (2004) have used this model to decompose TFP growth. Models III and IV are closely related to the standard effects models known from panel data treatment. In their initial specification with time-invariant efficiency, the term e_{it} is assumed to be an $iid\,(0, \sigma_e^2)$ white noise error term; the additional effect ϑ_i is a constant firm-specific

[24] Huang and Liu (1994) propose a very similar specification

[25] However, in contrast to model I, the distribution of inefficiency u_{it} varies over $\boldsymbol{z}$. Hence inefficiency is not assumed to be identically distributed.

[26] See Kumbhakar and Lovell (2000) chapter 7 for a literature review and a detailed summary of different approaches to incorporate exogenous influences on efficiency.

[27] Cornwell, Schmidt, and Sickles (1990) also propose an efficient instrumental variables estimator as a generalization of the Hausman and Taylor estimator (Hausman and Taylor 1981). We do not use the H-T estimator in this study.

parameter or an $iid\ (0, \sigma_\vartheta^2)$ random effect, respectively. The fixed-effects model can be estimated by OLS using the "within-groups" transformation. Then slope coefficients are estimated consistently as N or $T \to +\infty$ and unbiased due to unobserved time-invariant heterogeneity, since all stable characteristics of the individual firms are controlled. The random-effects model can easily be estimated by FGLS. As is common for random-effects models, the individual effects ϑ_i are assumed to be uncorrelated with the explanatory variables. In case this assumption does not hold, we have to expect biased slope parameters. Schmidt and Sickles (1984) rely on the firm-specific means of the residuals e_{it} from the within-groups and the FGLS estimator to recover estimates of the individual effects ϑ_i. From there on, the firms' level of inefficiency is obtained using the normalization $u_i = \max(\hat{\vartheta}_i) - \hat{\vartheta}_i$. However, the a-priori assumption of time-invariant inefficiency appears to be rather restrictive and may even be implausible for productivity growth analysis, especially if the operating environment is competitive and the panel includes more than a few time periods. In order to allow for time varying inefficiency, Cornwell, Schmidt, and Sickles (1990) adapt this model and replace the constant firm effect $\vartheta_i = \frac{1}{T_i}\sum_{t=1}^{T_i} e_{it}$ by $\hat{\vartheta}_{it} = \hat{\theta}_{1i} + \hat{\theta}_{2i}t + \hat{\theta}_{3i}t^2$, varying as flexible function of time. Firm-specific estimates of the respective parameters are derived by regressing the residuals of the within-groups and the GLS estimator on a constant, t and t^2 as in $e_{it} = \theta_{1i} + \theta_{2i}t + \theta_{3i}t^2 + \xi_{it}$.[28] Again, we get the firms' level of inefficiency from: $u_{it} = \max(\hat{\vartheta}_{it}) - \hat{\vartheta}_{it}\ \forall\, t$.

The main feature of models III and IV is that they allow for time varying estimates of inefficiency, consistent for all i and t as $T \to +\infty$ (Cornwell, Schmidt, and Sickles 1990) without the need to make distributional assumptions. An important issue regarding models III and IV is the lack of distinction of unobserved heterogeneity and inefficiency. The inefficiency estimates obtained from the models III and IV contain by construction the effects of all time invariant differences between the analyzed units. This may lead to an overestimation of inefficiency for firms which are subject to unfavorable external conditions. As Farsi, Filippini, and Greene (2005 p. 77) note, this question may be even more serious for the FE model, since "…the firm-specific effects do not follow a single distribution and thus can have a relatively wide range of variation". In the random-effects model, a part of the heterogeneity, which might be correlated with the explanatory

[28] ξ_{it} is an additional error term that captures all remaining variance in the residuals that is left unexplained by the flexible function of time.

variables (contrary to the respective assumption), can partly be suppressed in biased slope coefficients leading to biased TFP decompositions.

Model V was proposed by Battese and Coelli (1992) and extends the maximum likelihood random-effects panel model of Pitt and Lee (1981) to allow for time varying inefficiency. It is one of the most popular stochastic frontier models used in empirical work on TFP growth (e.g. Kim and Han 2001, Coelli, Rahman, and Thirtle 2003, Newman and Matthews 2006, Rasmussen 2010). Under the assumption that the firm effect u_i has a truncated normal distribution $N^+(\mu, \sigma_u^2)$ it is modeled as time-variant inefficiency taking the following form:

$$u_{it} = \beta(t)\, u_i \tag{2.2}$$

where $\beta(t) = exp\big(-\eta(t-T)\big)$. If inefficiency appears to be time-invariant $(\eta = 0)$ and u_i is half-normal distributed $(\mu = 0)$ the specification simplifies to the model of Pitt and Lee (1981). Model V shares two important properties with the GLS random-effects model IV. Despite the fact that inefficiency is allowed to vary over time, the firm specific random effect u_i is still time-invariant and includes constant firm effects in the inefficiency term (Greene 2005). In addition, model V also relies on the assumption that the firm effects are uncorrelated with the explanatory variables. Regarding the use of this model in productivity analysis, two more aspects are noted. First, the function $\beta(t)$ that determines how inefficiency varies over time is not very flexible and can therefore only depict monotonous patterns of efficiency change. Inefficiency increases at an increasing rate if $\eta < 0$, decreases at an increasing rate if $\eta > 0$ and remains constant if $\eta = 0$. Second, unlike models III and IV, model V restricts the time path for efficiency change to be common to all firms.[29] As an advantage of its panel nature, the model yields consistent estimates of u_{it} as $T \rightarrow +\infty$, in equivalence to the Pitt and Lee model – its time invariant special case.

[29] Cuesta (2000) proposes a maximum likelihood model allowing the temporal pattern of inefficiency to vary across firms

Table 2-1: Econometric specifications of stochastic frontier models

Model	Residual error e_{it} (unexplained by production technology)	Specification of error components	Heterogeneity	Inefficiency
Model I (Pooled)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{MLE} = e_{it} = v_{it} - u_{it}$	$v_{it} \sim N(0, \sigma_v^2)$ $u_{it} \sim N^+(0, \sigma_u^2)$	-	$E[u_{it} \| e_{it}]$
Model II (BC95)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{MLE} = e_{it} = v_{it} - u_{it}$	$v_{it} \sim N(0, \sigma_v^2)$ $u_{it} \sim N^+(\mu, \sigma_u^2)$ with: $\mu_{it} = \boldsymbol{z}_{it}'\boldsymbol{\zeta}$	-	$E[u_{it} \| e_{it}]$
Model III (FE)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{W} = e_{it} = v_{it} + \vartheta_i$	$v_{it} \sim iid(0, \sigma_v^2)$ $\vartheta_i = fixed$	-	$e_{it} = \theta_{1i} + \theta_{2i}t + \frac{1}{2}\theta_{3i}t^2 + \xi_{it}$ $\hat{\vartheta}_{it} = \hat{\theta}_{1i} + \hat{\theta}_{2i}t + \hat{\theta}_{3i}t^2$ $u_{it} = \max_i(\hat{\vartheta}_{it}) - \hat{\vartheta}_{it} \; \forall \, t$
Model IV (GLS)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{GLS} = e_{it} = v_{it} + \vartheta_i$	$v_{it} \sim iid(0, \sigma_v^2)$ $\vartheta_i \sim iid(0, \sigma_\vartheta^2)$	-	$e_{it} = \theta_{1i} + \theta_{2i}t + \frac{1}{2}\theta_{3i}t^2 + \xi_{it}$ $\hat{\vartheta}_{it} = \hat{\theta}_{1i} + \hat{\theta}_{2i}t + \hat{\theta}_{3i}t^2$ $u_{it} = \max_i(\hat{\vartheta}_{it}) - \hat{\vartheta}_{it} \; \forall \, t$

Model IV-M (GLS + Mundlak)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{GLS} - \overline{\boldsymbol{x}}_{i\cdot}'\boldsymbol{\gamma}^{GLS} = e_{it} = v_{it} + \vartheta_i$ $\alpha_i = \overline{\boldsymbol{x}}_{i\cdot}'\boldsymbol{\gamma}^{GLS} + \vartheta_i$	$v_{it} \sim iid(0, \sigma_v^2)$ $\vartheta_i \sim iid(0, \sigma_\vartheta^2)$	$\hat{\alpha}_i = \overline{\boldsymbol{x}}_{i\cdot}'\hat{\boldsymbol{\gamma}}^{GLS}$	$e_{it} = \theta_{1i} + \theta_{2i}t + \frac{1}{2}\theta_{3i}t^2 + \xi_{it}$ $\hat{\vartheta}_{it} = \hat{\theta}_{1i} + \hat{\theta}_{2i}t + \hat{\theta}_{3i}t^2$ $u_{it} = \max_i(\hat{\vartheta}_{it}) - \hat{\vartheta}_{it} \ \forall\, t$
Model V (BC92)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{MLE} = e_{it} = v_{it} - u_{it}$	$v_{it} \sim N(0, \sigma_v^2)$ $u_{it} = \beta(t)u_i$ $\beta(t) = \exp(-\eta(t-T))$ $u_i \sim N^+(\mu, \sigma_u^2)$	-	$E[u_{it} \mid e_{it}]$
Model VI (TFE)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{MLE} - \boldsymbol{D}'\alpha_i^{MLE} = e_{it} = v_{it} - u_{it}$	$v_{it} \sim N(0, \sigma_v^2)$ $u_{it} \sim N^+(0, \sigma_u^2)$	$\hat{\alpha}_i^{MLE}$	$E[u_{it} \mid e_{it}]$
Model VII (TRE)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{MSL} - \alpha_i = e_{it} = v_{it} - u_{it}$	$v_{it} \sim iid\ N(0, \sigma_v^2)$ $u_{it} \sim iid\ N^+(0, \sigma_u^2)$	$\alpha_i \sim N(0, \sigma_\alpha^2)$	$E[u_{it} \mid \alpha_i + e_{it}]$
Model VII-M (TRE + Mundlak)	$y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}^{MSL} - \overline{\boldsymbol{x}}_{i\cdot}'\boldsymbol{\gamma}^{MSL} - \vartheta_i = e_{it} = v_{it} - u_{it}$ $\alpha_i = \overline{\boldsymbol{x}}_{i\cdot}'\boldsymbol{\gamma}^{MSL} + \vartheta_i$	$v_{it} \sim iid\ N(0, \sigma_v^2)$ $u_{it} \sim iid\ N^+(0, \sigma_u^2)$	$\hat{\alpha}_i = \overline{\boldsymbol{x}}_{i\cdot}'\hat{\boldsymbol{\gamma}}^{MSL} + \vartheta_i$ $\vartheta_i \sim N(0, \sigma_\vartheta^2)$	$E[u_{it} \mid \vartheta_i + e_{it}]$

In an attempt to approach the issue of (unobserved) heterogeneity between firms in a stochastic frontier framework, Greene (2005) proposed the so called "true" fixed- and "true" random-effects models (henceforth abbreviated with TFE and TRE). Both "true" effects models have also come to use in TFP growth applications. Saal, Parker, and Weyman-Jones (2007), Wetzel (2009) and Filippini, Horvatin, and Zoric (2010) are recent examples. The TFE model VI is a straightforward extension of the pooled model (I), where α_i is a firm specific fixed-effect, while v_{it} and u_{it} are the components of the normal half-normal error term, representing statistical noise and inefficiency, just as in the pooled model.[30] Maximum likelihood is used to estimate the slope parameters and additional N dummy variables for individual α_i. The virtue of this brute force approach lies in the application of a numerical maximization algorithm, able to handle a large number of parameters. As Greene (2005) points out, maximum likelihood estimators of nonlinear models can be inconsistent in the presence of fixed effects, due to the incidental parameters problem.[31] The main difference between the TFE model (VI) and the conventional fixed-effects model (III) is in the way inefficiency estimates are derived. In the TFE model (VI) α_i represents time-invariant unobserved heterogeneity while inefficiency is obtained as in the pooled model from the conditional mean of the inefficiency term as $E[u_{it}|e_{it}]$. Thus, the TFE model is a fixed-effects model including a composed error term with normal half-normal distribution. Hence, despite the panel characteristic of the TFE model, technical inefficiency is assumed to vary stochastically over time and we cannot derive consistent estimates of u_{it} even if N or $T \rightarrow +\infty$. It should be noted, that the use of the TFE model is only appropriate if the analyzed panel contains more than a few time periods, since the individual efficiency scores rely on the variation of efficiency within the observations of an individual firm. If the observed period is short, some firms may exhibit inertia in their inefficiency, which would then mistakenly be captured by the fixed effect. A feature of the TFE model (VI) is that it allows the fixed effects α_i to be correlated with the input quantities x_{it}. However, α_i and x_{it} are still assumed to be uncorrelated with both, u_{it} and v_{it}.

In the TRE model (VII) the firm specific effect is assumed to be an *iid* normal distributed random term, i.e. $\alpha_i \sim N(0, \sigma_\alpha^2)$. As in model (VI) time-invariant effects are

[30] See Polachek and Yoon (1996) for one of the first discussions of a fixed effects model, accounting for inefficiency using a composed error term.

[31] We refer the reader to Greene (2005) for a short discussion on the incidental parameters problem related to stochastic frontier models. Wang and Ho (2010) provide a within- and first-difference transformation approach to estimate stochastic frontier models including fixed effects, which is "immune to the incidental parameters problem".

treated as heterogeneity and captured by α_i, while technical inefficiency is estimated by the conditional mean of the inefficiency term $E[u_{it}|\alpha_i + e_{it}]$.[32] As Greene notes, this model can be seen as a special case of the random parameters model where only the constant is a random parameter. As common to all random-effects models, the firm-specific effect α_i is "assumed to be uncorrelated with everything else in the model" (Greene 2008 p. 207). In order to overcome the problem of heterogeneity bias in the slope parameters in case this assumption does not hold, Farsi, Filippini, and Greene (2005) and Farsi, Filippini, and Kuenzle (2005) propose the incorporation of Mundlak's (1978) adjustment in the TRE and in the GLS model.[33] The underlying assumption is that the individual effects are a linear function on the group means of input quantities. The effects are then expressed in an auxiliary equation as:

$$\alpha_i = \boldsymbol{\gamma}'\bar{\boldsymbol{x}}_i + \vartheta_i \tag{2.3}$$

In (2.3), $\boldsymbol{\gamma}$ is an additional vector of parameters to be estimated and $\bar{\boldsymbol{x}}_i$ is a vector of the group means of all input variables, i.e. $\bar{x}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} x_{it}$. Now we want to highlight briefly the benefits of Mundlak's adjustment, applied in the GLS and TRE stochastic frontier models. We consider the incorporation of Mundlak's adjustment into the specification of stochastic frontier models to be based on the notion that firms have adjusted their input decisions conditional on the constant operating conditions they are subject to. That way it provides a possibility to improve the econometricians' ability to take heterogeneity into account that is unobserved to them but not to the producers.[34] By substituting equation (2.3) into the respective specifications of models IV (GLS) and VII (TRE), we add two more models to our comparison (IV-M and VII-M).

In those latter models, the individual effect α_i is decomposed in two components: the first part, which is explained by the group-mean variables and the remaining, unexplained part ϑ_i, assumed to be orthogonal to the explanatory variables. The important difference lies in the way this remaining component ϑ_i is treated. In the TRE specification, as proposed by Farsi, Filippini, and Greene (2005) (model VII-M) ϑ_i is treated as residual heterogeneity, which cannot be explained by the group means of input use. Then, as

[32] Kumbhakar and Hjalmarsson (1993) propose a very similar model, which is, however, estimated in two steps.

[33] In contrast to our application, Farsi, Filippini, and Kuenzle (2005) estimate a GLS model with time-invariant technical inefficiency.

[34] It is undisputed that this approach does not save the researcher from making assumptions. For example, it could be argued whether it is realistic to assume that firms adjust to operating conditions but do not know their inefficiency.

intended by the TRE model, this residual heterogeneity is captured by the firm specific random effect, i.e. $\vartheta_i \sim N(0, \sigma_\vartheta^2)$. In the augmented GLS random effects-model (IV-M), we assume that the group-mean variables explain all heterogeneity between the firms in the sample. Then the term ϑ_i becomes part of the GLS random-effects model's iid error term. In consequence of the following procedure to derive the inefficiencies, ϑ_i is treated as part of the time varying inefficiency.

We incorporate these models in our comparison, because they are useful in two ways. First, the estimated slope parameters are free from heterogeneity bias to the extent that (2.3) can capture the correlations of the random effect with the explanatory variables. As noted above, this is important in regard to our estimates of production and scale elasticities. Second, by modeling an individual effect α_i through a function of observed variables, we can also mitigate the heterogeneity bias in the estimates of inefficiency. This is especially appealing in the case of model IV-M (GLS); this model provides an alternative way to derive (consistent) time-varying estimates of inefficiency, while taking unobserved heterogeneity into account.[35]

2.3. Empirical Application

We apply the different models discussed in the previous section to a panel of specialized German dairy farms. Based on the resulting estimates of technology parameters and inefficiency, we calculate rates of total factor productivity growth. Using farm data for our empirical application is beneficial in two ways. First, the methodology we analyze has been used in numerous empirical studies on the agricultural sector e.g. Brümmer, Glauben, and Thijssen (2002), Newman and Matthews (2006), Key, McBride, and Mosheim (2008), Rasmussen (2010) and many others, showing its relevance. Second, farms are natural candidates for firms that operate under heterogeneous production conditions affecting their feasible output (Sherlund, Barrett, and Adesima 2002).[36]

2.3.1 Data

The data for our empirical application is taken from German farm bookkeeping records maintained by the Bavarian Agricultural Research Institute (LfL). The dataset is an

[35] We did not include the "fixed management model" proposed by Alvarez, Arias and Greene (2006) in our comparison. This model also attempts to account for unobserved heterogeneity and is closely linked to the TRE model. Nevertheless, we applied the model to our dataset, using the Mundlak specification, as suggested in the respective paper and found that the results are almost identical to those obtained from model VII-M (TRE + Mundlak). Results can be obtained from the authors.

[36] Abdulai and Tietje (2007) also elaborate on the relevance of (unobserved) heterogeneity in regard to the use of different stochastic frontier models on agricultural data.

unbalanced panel with 7465 observations of 974 farms. It covers the years 2000 to 2008. We only considered specialized dairy farms generating at least 66% of their total revenues from dairy production. Farms with less than four consecutive observations were also excluded from the analysis.[37] The observations are evenly spread over the period under consideration with 7.7 observations per farm on average. We consider two outputs (*milk* and *other output*) and four inputs (*labor*, *land*, *intermediate inputs* and *capital*). Descriptive statistics by year are given in table 2-2.

Table 2-2: Summary statistics of variables by year

	2000	2001	2002	2003	2004	2005	2006	2007	2008
Milk output (1000 €)	56.69 (25.30)	59.96 (28.26)	60.36 (28.61)	62.07 (30.34)	64.26 (31.32)	65.81 (32.56)	66.72 (33.64)	69.86 (35.28)	69.50 (36.39)
Other output (1000 €)	30.59 (15.63)	30.04 (15.70)	31.16 (16.10)	31.98 (17.35)	31.19 (16.18)	31.63 (17.58)	32.92 (18.84)	35.10 (21.27)	33.70 (21.19)
Labor (mwu)	1.52 (0.44)	1.54 (0.44)	1.54 (0.43)	1.53 (0.44)	1.54 (0.45)	1.54 (0.46)	1.55 (0.46)	1.54 (0.46)	1.55 (0.44)
Land (ha)	40.08 (22.52)	41.14 (23.19)	42.02 (23.69)	42.76 (24.99)	43.42 (25.09)	44.85 (26.07)	45.32 (26.54)	46.74 (28.08)	47.71 (29.17)
Intermediate inputs (1000 €)	44.91 (25.01)	45.34 (26.85)	46.32 (26.36)	47.28 (27.32)	48.87 (27.45)	48.58 (26.62)	48.78 (27.56)	49.94 (27.71)	49.64 (28.41)
Capital (1000 €)	206.62 (110.89)	208.56 (117.87)	206.94 (120.55)	204.15 (121.63)	202.94 (121.65)	202.18 (124.02)	198.96 (123.03)	195.15 (121.51)	195.51 (128.50)

[1] Mean value and standard deviation (in parentheses)

The output *milk* is measured in total revenues from milk and milk products. This allows to account for quality differences, since the price that the individual farmer receives from the processor varies, depending on the fat and protein content in the delivered milk. The variable *other output* contains revenues from beef, crops and other commodities. The input variable *labor* subsumes family and hired labor in man working units (mwu). The variable *land* measures total cultivated land in hectare (ha). The *intermediate inputs* include expenses for forage and crop production (e.g. seed, fertilizer, pesticides, fuel, contractors) and animal production (e.g. veterinary, concentrates). The variable *capital* includes the end-of-year value of buildings, technical facilities, machinery and livestock. We use price indices from the German Federal Bureau of Statistics to deflate the aggregated monetary input and output variables using the year 2005 as base year. As *z*-

[37] We create this subsample to improve the panel character of the dataset. Especially the use of the models III and IV made this restriction necessary.

variables for the inefficiency effects model II we use regional dummies representing nine different agricultural production areas.

2.3.2 Specification

Dairy farms are a typical example of a multi-product firm. Even specialized dairy farms do in most cases not solely produce milk, but may also produce beef, veal and field crops as part of their integrated production process.[38] We model this multi-input multi-output technology using an output-oriented[39] distance function $D^O(\boldsymbol{x},\boldsymbol{y},t)$, where $\boldsymbol{x} = (x_1, x_2, \dots, x_k) \in R_+^K$ refers to a nonnegative vector of inputs used to produce a nonnegative vector of outputs $\boldsymbol{y} = (y_1, y_2, \dots, y_m) \in R_+^M$ and t denotes an exogenous time trend $t = 1,2, \dots, T$. We choose the common flexible translog functional form that limits the a priori restrictions on the relationships among inputs and outputs.

Hence,

$$\begin{aligned}\ln D_{it}^O(\boldsymbol{y},\boldsymbol{x},t) = {} & \beta_0 + \sum_{m=1}^{M} \alpha_m \ln y_{mit} \\ & + \sum_{k=1}^{K} \beta_k \ln x_{kit} + \frac{1}{2} \sum_{m=1}^{M}\sum_{n=1}^{M} \alpha_{mn} \ln y_{mit} \ln y_{nit} \\ & + \frac{1}{2} \sum_{k=1}^{K}\sum_{j=1}^{K} \beta_{kj} \ln x_{kit} \ln x_{jit} + \sum_{m=1}^{M}\sum_{k=1}^{K} \delta_{mk} \ln y_{mit} \ln x_{kit} + \tau_1 t \\ & + \frac{1}{2}\tau_2 t^2 + \sum_{m=1}^{M} \varsigma_{mt}\, t \, \ln y_{mit} + \sum_{k=1}^{K} \nu_{kt}\, t \ln x_{kit}\end{aligned} \tag{2.4}$$

The parameters of this function must satisfy the symmetry restrictions $\alpha_{mn} = \alpha_{nm}$ and $\beta_{kj} = \beta_{jk}$. In addition, homogeneity of degree one in output quantities ($\sum_{m=1}^{M} \alpha_m = 1$ and $\sum_{m=1}^{M} \alpha_{mn} = \sum_{k=1}^{K} \delta_{mk} = \sum_{m=1}^{M} \varsigma_{mt} = 0$) is imposed by normalizing the function by an arbitrarily chosen output quantity:

$$\ln\left(\frac{D_{it}^O(y,x,t)}{y_{2it}}\right) = TL\,(\boldsymbol{y}^*, \boldsymbol{x}, t) \;\text{ with }\; y_{mit}^* = {}^{y_{mit}}/_{y_{2it}} \tag{2.5}$$

[38] A farm that has outsourced all calf and heifer rearing, as well as all cereal and forage production might be such an example. However, we do not find farms like this in our sample.

[39] The choice of orientation has to be made individually for a specific application. For our case, we assume that the farms in our sample are less flexible in the adjustment of their inputs than their output. The labor input which largely contains the family workforce is one example for a rather inflexible input. On the other side, a very well established system for quota trading exists in Germany. Hence, output can be considered as unrestricted.

where TL indicates translog, and $TL\ (\boldsymbol{y}^{*}, \boldsymbol{x}, t)$ is the right hand side of (1) after dividing all output quantities by y_2.[40] The dependent variable $\ln D_{it}^{O}$ is unobservable so we rearrange the distance function for the estimation in a stochastic frontier framework. We add a random error term v_{it} and given that $\ln D_{it}^{O} \leq 0$, we replace $\ln D_{it}^{O}$ with $-u$ such that,

$$-\ln y_{2it} = TL\ (\boldsymbol{y}^{*}, \boldsymbol{x}, t) + u + v_{it} \tag{2.6}$$

To be more comparable with the standard stochastic production frontier, we slightly adapt equation (2.6) by multiplying both sides by -1. Hence, we use $\ln y_{2it}$ as dependent variable and reverse the sign of the regressors and the one sided inefficiency term u. The way the inefficiency u is modeled varies, depending on the applied econometric model (table 2-1).

2.3.3 Calculation and decomposition of TFP growth

Based on the estimated parameters and inefficiency estimates of the described models *I* - *VII*, we use the derivative-based approach to calculate and decompose total factor productivity growth.[41] See Denny, Fuss, and Waverman (1981) and Bauer (1990) for applications using production- and cost functions and Brümmer, Glauben, and Thijssen (2002) and Karagiannis, Midmore, and Tzouvelekas (2004) for applications using output- and input distance functions.[42] Keeping our application simple and comparable to the production function one-output special case of the output distance function, we assume allocative efficiency and perfect competition on input- and output markets.[43] In this setup we obtain the following expressions for technical change, efficiency change and the scale change effect by taking the total differential of (2.6) and relating it to the Divisia index of total factor productivity growth. Technical change is calculated by: $TC_{it} = \frac{\partial \ln y_{2it}}{\partial t} = \tau_1 + \tau_2 t + \sum_{m=1}^{M} \varsigma_{mt}\ \ln y_{mit} + \sum_{k=1}^{K} \nu_{kt}\ \ln x_{kit}$ (Morrison Paul, Johnston, and Frengley 2000).

---

[40] Despite the common use of distance functions as representations of multi-input multi-output representations of production technology, concerns exist about the exogeneity of the ratio of outputs used as a dependent variable in the estimation. Based on findings by Schmidt (1988) and Mundlak (1996), Coelli (2000) argues that the ratio form of outputs does not suffer from endogeneity assuming expected profit maximization. See also Kumbhakar and Lovell (2000) p. 94 for a discussion.

[41] Many empirical studies also use the Malmquist TFP index developed by Caves, Christensen and Diewert (1982). This alternative approach to decompose TFP growth is based on the same estimates of technology parameters and inefficiency. Hence we expect qualitatively identical results of our analysis.

[42] Lovell (1996) provides an overview that also includes the non-parametric approach to efficiency measurement.

[43] Hence, we exclude in this primal framework contributing factors to TFP growth that are “connected to market” and concentrate on those “connected to technology” (Brümmer, Glauben, and Thijssen 2002 p. 632).

This expression of technical change is firm- and time-specific according to the translog functional form of (2.4).

The way we calculate the effect of changes in technical efficiency varies across the different econometric models. For the models III, IV and IV-M, we follow Fecher and Pestieau (1993) and obtain farm-specific estimates for the change in technical efficiency from $TEC_{it} = \frac{\partial \vartheta_{it}}{\partial t} = \hat{\theta}_{2i} + \hat{\theta}_{3i} t$.[44] For the model of Battese and Coelli (1992) (V), we get $TEC_{it} = -\frac{\partial u_{it}}{\partial t} = \hat{u}_i \hat{\eta} \exp\left(-\hat{\eta}(t-T)\right)$. The remaining models do not specify an explicit way how technical inefficiency is allowed to vary over time. On the contrary, they include the assumption that inefficiency is independent across farms and time. Thus, for these models the change in technical efficiency is measured from its discrete changes from period t to $t+1$; i.e. $TEC_{it} = u_{it} - u_{it+1}$.[45]

Based on the distance elasticities with respect to the inputs, the scale elasticity and the changes in input usage, we calculate the scale change effect: $SC_{it} = (\varepsilon_{it} - 1) s_{ikt} \ln\left(\frac{x_{ikt}}{x_{ikt+1}}\right)$.[46] In this expression $\varepsilon = \sum_{k=1}^{K} \frac{\partial \ln D^O}{\partial \ln x_k}$ and $s_k = \frac{\partial \ln D^O / \partial \ln x_k}{\sum_{k=1}^{K} \partial \ln D^O / \partial \ln x_k}$. We observe a positive (negative) contribution to productivity change if $\varepsilon > 1$ and input usage is expanded (reduced) or if $\varepsilon < 1$ and input usage is reduced (expanded). In the case of constant returns to scale ($\varepsilon = 1$) or constant input quantities SC becomes zero.

2.4. Empirical Results

We present the estimated parameters of the nine different models in tables 2-8 and 2-9 in the appendix. All models have been estimated using LIMDEP 9.0 (Greene 2007) and EViews 6 (QMS 2007), respectively. The percentage of slope parameters, significantly different from zero at the 5% level, ranges from 54.2% in model IV-M to 80.0% in model VII with an average of 68.5%. Comparing the estimated coefficients, we find an apparent variation across some of the models while others are more similar.

[44] This approach implicates that technical change is associated with the trend variable in the technology part of the distance function and is common to all firms in a sector. In contrast, efficiency change is individual to producers. See Lovell (1996) for a short discussion of different interpretations.

[45] Karagiannis and Tzouvelekas (2005) show, that the marginal effects of time-varying variables in the inefficiency part of model II have to be taken into account in a decomposition of TFP growth. In our application, only time-invariant variables enter the inefficiency part. In the decomposition described by Zhu and Lansink (2010), this relates to a case where all discrete changes in technical efficiency over time are ascribed to “unspecified factors”.

[46] For a continuous approximation of the elasticities, the arithmetic means of two subsequent periods can be used.

The variation in the technology parameters carries over into the respective distance elasticities. Table 2-3 shows the sample average elasticities of inputs and outputs as well as the average returns to scale measure for all estimated models. The average elasticities have the right signs on the input as well as on the output side. Some patterns in the calculated average elasticities can be noted: for all estimated models the *intermediate inputs* show the highest average output elasticity and the returns to scale measure is below one; on the output side the elasticities reflect the high share of milk output in total production. Based on similar average elasticities we can identify three groups of models. The first group consists of the Pooled, the BC95 and the TFE models (I, II, VI). The three models altogether show comparably high average scale elasticities close to one, a high elasticity for *intermediate inputs* and a relatively low elasticity for the input *land*, close to zero.[47] As a slight variation, model II shows an output elasticity for the input *land* which is still low, but significantly higher compared to models I and VI.[48] We assign this finding to the incorporation of the regional dummy variables into the inefficiency effects model. That way (observed) information about heterogeneous production conditions, which is basically related to the productive potential of the used acreage, is included into the model, leading to a more reliable estimate.[49]

Table 2-3: Average distance elasticities

	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
Labor	0.174	0.165	0.058	0.093	0.058	0.090	0.177	0.088	0.056
Land	0.018	0.057	0.111	0.163	0.110	0.165	0.013	0.174	0.103
Inter. Inputs	0.628	0.611	0.335	0.436	0.335	0.417	0.632	0.399	0.317
Capital	0.134	0.128	0.057	0.092	0.056	0.085	0.134	0.092	0.062
Milk	-0.725	-0.724	-0.801	-0.790	-0.801	-0.790	-0.718	-0.801	-0.814
Other	-0.275	-0.276	-0.199	-0.210	-0.199	-0.210	-0.282	-0.199	-0.186
RTS	0.953	0.961	0.560	0.783	0.559	0.757	0.955	0.752	0.539

The second group includes the random effects models IV, V and VII. All three models produce highly similar elasticities, on the input and output side. The average scale elasticity is lower compared to the first group of models, ranging between 0.752 and 0.783. The *intermediate inputs* again show the highest average elasticity, followed by the

[47] However, the standard errors of the average elasticities calculated using the delta method show that all are significantly different from zero at least at the 1% level.

[48] We confirmed the significance of the difference in the means using a Welch test.

[49] As noted above it could be argued that information about land quality should be incorporated directly in the distance function.

input *land* and then *labor* and *capital*. The three models share the assumption that the firm specific component – specified as ϑ_i, u_i and α_i, respectively in table 1 – is uncorrelated with the explanatory variables.

Models III, IV-M and VII-M build the third group of models which is connected by the fact that they either assume that individual effects are correlated with the explanatory variables (as the fixed effects model III) or they take a possible correlation explicitly into account using Mundlak's (1978) adjustment. For this group of models we find the lowest scale elasticity in a range between 0.539 and 0.560. The similarities in the distance elasticities for the models in this group indicate that the group-mean variables pick up a large fraction of the correlation between the firm-specific effects and the explanatory variables. In fact, the fixed effects model (III) and an augmented GLS model such as model IV-M are assumed to be identical as described by Mundlak (1978).[50] This relationship does not hold for maximum likelihood stochastic frontier models such as VII-M. However, to the extent that the group-mean variables are able to capture the correlations of the firm specific random effect with the explanatory variables, we can mitigate the heterogeneity bias in the estimated technology parameters.[51]

A summary of the estimated efficiency scores from each model is presented in table 2-4. As noted above, the efficiency scores obtained from the models III, IV and V contain the effects of firm-specific unobserved characteristics. This leads to downward biased efficiency scores for farms with competitive disadvantages beyond the control of the farm's manager, e.g. unfavorable natural conditions. The fact that these models produce the lowest efficiency scores suggests that unobserved heterogeneity cannot be ignored in our dataset. Model III produces the lowest efficiency scores with a mean efficiency of less than 0.5. This would imply that on average all observed dairy farms could double their output without altering their inputs if they were fully efficient – a clearly unrealistic result. Compared to model III, the mean efficiency obtained from the models IV and V are higher. They lie in the range between 0.59 and 0.74. As Farsi, Filippini, and Kuenzle (2005a) note, this can be attributed to the correlation between the explanatory variables

[50] In our case, the GLS-M model is not entirely identical to the FE model, because we have left out the group-mean variables for inputs interacted with the trend variable. The results of a specification including the respective additional variables are almost identical. Hence, we prefer the present, more parsimonious specification.

[51] The Mundlak adjustment is certainly not a panacea for all problems associated with the estimation of production or distance functions when heterogeneity is unobserved. Griliches and Mairesse (1998) elaborate on the benefits and difficulties arising from use of panel techniques for the estimation of production functions. They also discuss the frequently documented reduction in estimated scale elasticities, just as found in our empirical application. However, especially in the context of stochastic frontier analysis the Mundlak adjustment has its appealing features.

and the firm-specific effects. That way the heterogeneity is partly captured in the slope parameters.

Table 2-4: Descriptive statistics of technical efficiency scores

	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
Mean	0.868	0.869	0.492	0.588	0.614	0.741	0.842	0.918	0.916
Standard deviation	0.058	0.059	0.121	0.112	0.099	0.130	0.032	0.045	0.048
Maximum	0.977	0.980	1.0	1.0	1.0	0.994	0.962	0.991	0.990
Minimum	0.522	0.538	0.128	0.225	0.264	0.256	0.506	0.544	0.504

The “true” fixed- and “true” random-effects models (VI and VII) produce rather high efficiency scores, of 0.842 and 0.918 at the mean. These models account explicitly for all time-invariant firm-specific effects and thus, the efficiency scores solely depend on the within variation of the firms. That means any potentially existent time-invariant inefficiency is suppressed in the firm-specific effect. Given that the European dairy sector can hardly be assumed to be highly competitive, we cannot rule out the possibility that farms with a certain amount of inertia in their inefficiency stay in the sector. On the other side the present dataset also has features that agree with the use of the “true” effects models. The panel encompasses the years from 2000 to 2008, a reasonable sized time frame. More importantly, dairy farmers had to adapt to several severe changes in their operating conditions during this period, such as policy changes and fluctuations in factor and output prices. This case of a potential upward bias in the efficiency estimates illustrates the analogue to the predictable downward bias in the models III, IV and V, who do not account for heterogeneity. The incorporation of the Mundlak variables in models IV-M and VII-M shows the expected results. The GLS+M specification accounts for heterogeneity as specified in table 1 and can therefore reduce the respective contamination of the efficiency scores. This leads to an increase in the mean and a reduction in the standard deviation of the efficiency score. In case of the TRE model (VII) the incorporation of the group-mean variables has a different effect, as this model already attempts to capture heterogeneity in its basic specification. Hence, possible time-invariant differences between the firms are captured in the random constant anyway and the efficiency estimates of the TRE model are free of time-invariant heterogeneity. The random constant is specified to be normally distributed with an additional parameter σ_α or σ_ϑ which is the standard deviation of this random parameter in the TRE and the TRE-M model, respectively. This additional parameter is a measure of the unaccounted variation between the farms. By including the Mundlak variables into the TRE model we partly account for this unobserved heterogeneity. As expected, this reduces the unaccounted variation between the firms from σ_α= 0.2327 to σ_ϑ= 0.1459.

The correlation between the efficiency scores obtained from the different models (table 2-5) support the interpretations of the varying results in table 2-4. The efficiency scores from models I and II are highly correlated (0.93) and show considerable correlations with the scores of all other models (0.47 – 0.79).

Table 2-5: Correlation matrix of technical efficiency scores

	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
Pooled	1.00								
BC95	0.93	1.00							
FE	0.60	0.57	1.00						
GLS	0.75	0.75	0.92	1.00					
GLS+M	0.78	0.74	0.52	0.70	1.00				
BC92	0.70	0.71	0.87	0.92	0.65	1.00			
TFE	0.54	0.47	0.15	0.20	0.19	0.03	1.00		
TRE	0.56	0.50	0.24	0.29	0.30	0.13	0.87	1.00	
TRE+M	0.54	0.48	0.23	0.28	0.35	0.11	0.80	0.97	1.00

Note: different shades of grey visualize the extent of correlation from high (dark) to low (light)
[1]Spearman rank correlation coefficients are highly similar to the displayed coefficients

Both models do not take the panel structure of the data into account. Hence, the efficiency scores from these models contain time varying as well as time invariant components. This explains the apparent correlation of the efficiency scores with both, the conventional and the “true” effects models. We find strong correlations between the “conventional” panel models III, IV and V (0.87 – 0.92). This is not surprising. As a common feature, in these models the inefficiency estimates include a time invariant fixed- (III) or random- (IV, V) effect, which also contains firm-specific heterogeneity. The correlation with the efficiency scores obtained from the Mundlak specification of model IV is lower and lies between 0.52 and 0.70. This finding can also be expected, as the Mundlak adjustment accounts for part of the unobserved heterogeneity and removes it from the efficiency scores. The correlation between the scores from the “true” models VI, VII and VII-M is also fairly high. Controlling for the heterogeneity, these models show similar abilities to identify time-varying inefficiencies. However, the correlation between the “conventional” and “true” panel models is rather low and varies between 0.02 and 0.35. This confirms that the way how heterogeneity is handled has a strong influence on the efficiency estimates. Our findings for the efficiency scores and the correlations between the scores obtained from different models agree in general with the findings of previous studies, comparing stochastic frontier models (e.g. Farsi, Filippini, and Greene 2005, Farsi, Filippini, and Kuenzle 2005, Abdulai and Tietje 2007).

Based on the estimates of the different econometric models, we measure and decompose total factor productivity for the observed dairy farms. We report in table 2-6 the average values for changes in total factor productivity (TFPC), technical change (TC), changes in technical efficiency (TEC) and productivity changes due to changes in the scale of operations (SC) (see table 2-10 in the appendix for yearly values) and their percentage share on TFPC. The results in table 2-6 show that technical change has the strongest influence on total factor productivity. In all models, TC has a positive effect on productivity throughout the observed time period but we find TC to be slightly increasing over time for the models I, II, VII and VII-M but more or less linear growth rates for the remaining models (see table 2-10). The average annual productivity growth induced by technical change ranges from 1.19% in model IV to 1.64% in model V.

Table 2-6: Average growth rates of decomposed total factor productivity and share of components

	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
	Average annual change rate (%)								
TFPC	1.24	1.20	1.03	1.07	1.03	1.11	1.28	1.03	1.03
TC	1.36	1.29	1.24	1.19	1.24	1.64	1.41	1.25	1.30
TEC	-0.09	-0.06	-0.08	-0.06	-0.08	-0.51	-0.09	-0.11	-0.15
SC	-0.04	-0.03	-0.13	-0.05	-0.13	-0.02	-0.04	-0.10	-0.12
	Share of TC, TEC and SC on TFPC (%)								
TC	109.86	108.02	120.65	110.65	120.86	147.81	110.17	120.76	126.26
TEC	-6.90	-5.36	-7.80	-5.64	-7.96	-46.34	-7.16	-11.09	-14.64
SC	-2.95	-2.66	-12.84	-5.01	-12.90	-1.47	-3.01	-9.68	-11.62

The average rates of technical efficiency change also vary considerably across specifications. The highest absolute change rates are found for the BC92 model (-0.51).[52] The change rates obtained from the remaining models are very low, in a range between -0.06% and -0.15% per year (see table 2-10 for information on the development over time). The scale change effect also has rather low but negative impact on productivity for all models. The magnitude of the effect heavily depends on the returns to scale, thus we find

[52] Some uncertainty exists on the exact reason for the comparably high rates of TEC in this model. A possible explanation is that the specification of time-varying inefficiency is rather inflexible. Low levels of efficiency are connected to high rates of efficiency change, subject to the parameter η which is common to all firms. We find that the high rates of negative TEC are offset by proportional higher rates of positive TC.

the largest effect for the models III, IV-M and VII-M with RTS between 0.54 and 0.56. The average annual growth rates of TFP also vary across models. Looking at the extreme cases, model VI (TFE) reports a growth rate which is more than 20% higher compared to models at the lower range.

Even more important are the differences in the importance of the components of TFP growth. In the lower part of table 2-6 we report the share of the single components; the differences in the single components TC, TEC and SC are striking in some cases. This finding is relevant especially for the case of empirical applications which base regulatory or policy recommendations on their calculations of TFP growth. As Grosskopf (1993, p. 169) points out, “a slowdown in productivity growth due to increased inefficiency suggests different policies than a slowdown due to lack of technical change”. Low rates of technical change could be interpreted as an indication of an insufficiently innovative sector lacking investments, implying suggestions for expenditures through governmental policies. Decreasing efficiency levels point towards a growing heterogeneity in the firms’ productive performance. What often is recommended in these cases are investments in extension service and consulting as well as the reduction of incentive problems to bring the firms back towards the frontier (Fan 1991, Bayarsaihan and Coelli 2003, Aiello, Mastromarco, and Zago 2010). Special attention is also given to the interpretation of the returns to scale measure and the resulting scale effect on productivity growth; for instance Key, McBride, and Mosheim (2008) recommend revising legislation that limits the size or growth of hog farm enterprises. For our application, the substantial differences in the relative importance of the TFP growth components across the econometric models could have led to strongly differing or even contradictory policy advice.

2.5. Model Selection

Our empirical application shows that the results of an analysis of productivity growth, heavily relies on the choice of an econometric model, used to estimate the representation of the frontier production technology. The fact that different econometric models – imposing different assumptions on the data and the data generating process – lead to different results is not new. However, this is only a problem if we cannot choose reliably among the models. In the case of the stochastic frontier models presented in this study, we find the situation that not all models are nested; hence formal testing cannot reveal the “one” right model for each dataset. In the following section, we attempt to reduce the number of appropriate models by rejecting as many models as possible based on statistical tests. Subsequently, we discuss further options which might be of interest for empirical researchers to reduce the number of models.

We start with a test of the pooled model (I) against the inefficiency effects model BC95 (II). This is possible, since model I is nested in model II. The likelihood ratio test[53] gives statistic of LR = 241.12 exceeding the critical value at the 1% level $\chi^2_{(9)} = 21.67$ indicating that model II is preferred over model I.[54] In the specification of model I and II (Pooled, BC95) the panel structure of the data is ignored implying the assumption that no firm-specific effects are present. We approach this question using the Baltagi and Li (1990) form of the Breusch-Pagan Lagrange multiplier statistic for unbalanced panel data.[55] We clearly reject the null hypotheses of "no group effects" with a test statistic of $LM =$ 9293.65 against a critical value of $\chi^2_{(1)} =$ 6.64. Models I and II make the assumption that both error components, technical inefficiency and statistical noise, are independently distributed. Hence, this result stands against these specifications and has to be taken into account in regard to the use of models I and II. Another way of testing the presence of firm-specific effects in the data is to test the pooled model I against the "true" effects models VI and VII. Model I is a special case of the TFE model for $\alpha_i \equiv \beta_0 \ \forall \ i$. This hypotheses is rejected as the likelihood-ratio test gives a statistic of LR = 2071.28 which is much higher than the critical value of $\chi^2_{(973)}$ = 1078.55[56] (Greene 2008, p. 211). Finally we compare the log-likelihood of the TRE model (VI) against the pooled model (Greene 2008, p.207). The resulting test statistic is LR = 6615.32 against a critical value of $\chi^2_{(1)} = 6.64$[57].

For a straightforward check whether the explanatory variables are correlated with existing firm-specific effects ($E[x_{it}\vartheta_i] \neq 0$) we perform a Hausman test on the GLS random effects model. The test rejects the hypotheses of no correlation between the effects

---

[53] The likelihood ratio test statistic is given by $LR = 2(\ln L_R - \ln L_U)$, where $\ln L_R$ ($\ln L_U$) is the log-likelihood of the restricted (unrestricted) specification.

[54] We also check whether inefficiency is present at all in our empirical dataset by testing model I against a simple OLS model. The hypothesis of no inefficiency is clearly rejected.

[55] The test statistic is calculated based on the residuals (e_{it}) of a pooled OLS model: $LM = \left[\frac{(N\bar{T})^2}{(\sum_i T_i^2) - N\bar{T}}\right]\left[\frac{\sum_i(\sum_t e_{it})^2}{\sum_i \sum_t e_{it}{}^2} - 1\right]^2$ where $\bar{T} = \frac{N}{\sum_i\left(\frac{1}{T_i}\right)}$.

[56] The validity of this test is unclear. The incidental parameter problem can prevent the TFE and the pooled model to converge under the null hypotheses.

[57] We note that this is also a non-standard test. Under the null hypotheses (variance of the random constant equals zero), the test statistic is not asymptotically χ^2-distributed because the tested value is on the border of the feasible parameter space. However, for our application the issue is negligible since we only restrict a single parameter and the calculated LR statistic is about 1000 times the critical value anyway. For more on the topic see e.g. Self and Liang (1987).

and the used variables with a test statistic of 900.25 against a critical value of $\chi^2_{(27)} = 46.96$.[58] This is a strong indication, that all models, which assume no such correlation (IV, V, VII) produce biased slope parameters. Similarly we test the random effects models IV and VII against their respective Mundlak specifications. Using a Wald test, we can reject the hypotheses that the additional group-mean variables in the GLS + M specification are jointly equal to zero with a test statistic of F = 48.15 against a critical value of $F_{(20,6444)}$ = 2.38. The same applies for the TRE + M specification where we reject this hypotheses based on a likelihood ratio statistic of LR = 1028.66 against a critical value of $\chi^2_{(20)} = 37.57$. Based on the described statistical tests, we were able to exclude five (I, II, IV, V, VII) out of nine models, leaving us with the FE (III), the TFE (VI) and the two Mundlak specifications (IV-M, VII-M).[59]

As this in not entirely satisfactory, we look for alternatives to check which of the remaining models fits the data best. For the widely used translog functional form, we advise to take a closer look at how well the estimated representations of the production technology are in line with the requirements implied by microeconomic theory, namely monotonicity and quasi-concavity in inputs and concavity in outputs. It has been pointed out by several authors (e.g. O’Donnell and Coelli 2005, Sauer, Frohberg, and Hockmann 2006 and Henningsen and Henning 2009), how important this theoretical consistency is for a correct interpretation of the obtained parameters and efficiency scores and accordingly for the results of the decomposition of TFP growth. As shown in table 2-3, the distance elasticities resulting from all models show the right signs and therefore fulfill the monotonicity requirement at the sample mean. According to Sauer, Frohberg, and Hockmann (2006) this is the minimum requirement, which has to be fulfilled in any case to obtain meaningful results. Monotonicity violation on the input side would, for example, implicate that a reduction in inputs given a fixed level of output would reduce productivity. After checking for monotonicity for all observations we find some violations for all models in our application. However, as reported in table 2-7, the share of observations with present violations of monotonicity are more severe for some models and less for others. For example, we find that 40.3 % of the observations show the wrong sign in the distance elasticity of the input land if the TFE model (VI) is used for estimation.

[58] The test statistic is given by: $H = [\widehat{\boldsymbol{\beta}}^{W} - \widehat{\boldsymbol{\beta}}^{GLS}]'\boldsymbol{\Gamma}^{-1}[\widehat{\boldsymbol{\beta}}^{W} - \widehat{\boldsymbol{\beta}}^{GLS}]$ where $\boldsymbol{\Gamma}^{-1} = Var(\widehat{\boldsymbol{\beta}}^{W} - \widehat{\boldsymbol{\beta}}^{GLS})$ (Greene, 2003).

[59] Another issue that leads to further room for statistical testing is how technical efficiency is specified to vary over time. Karagiannis and Tzouvelekas (2010) provide some insights on this topic.

This can hardly be accepted. To check the curvature conditions of quasi-convexity in inputs and convexity in outputs, we construct the respective (bordered-) Hessian matrix for each data point and report the percentage of violations in the lower part of table 2-7.

Table 2-7: Violations of monotonicity and curvature conditions in percent

	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
					Monotonicity				
Labor	0.0	0.0	6.6	0.3	6.5	0.5	0.0	0.9	4.6
Land	37.0	14.4	0.2	0.0	0.2	0.0	40.3	0.0	0.3
Inter. Inputs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Capital	0.0	0.0	0.6	0.0	0.6	0.0	0.0	0.1	0.1
Milk	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Other	0.1	0.1	0.6	0.5	0.6	0.5	0.1	0.2	0.3
					Curvature				
Input	4.5	0.7	0.0	0.0	0.0	0.0	5.1	0.1	0.1
Output	31.3	31.5	9.3	13.6	9.2	12.3	26.3	40.9	38.2

On the input side, almost all models are perfectly in line with the curvature requirements whereas we find some curvature violations on the output side for all models. These are prominent in the TRE and the TRE+M models where 40.9% and 38.2% of the observations violate the curvature requirements on the output side. Hence, we could challenge two more econometric models based on how consistent the estimated technologies are with microeconomic theory.

Further aspects that should be taken into account when choosing an econometric model involves the distinction between inefficiency and heterogeneity. Based on expert knowledge about the sector under investigation, it should be reflected which assumptions are reasonable. Are the analyzed firms actually working under heterogeneous production conditions, which should be controlled? Can the existence of time-invariant inefficiency be ruled out generally, e.g. by a competitive operating environment, changes in the operating and management conditions (e.g. policy and regulation) and a sufficient number of time periods? These considerations argue against the use of the FE model (III) on our dataset. As noted, it includes all unaccounted time-invariant effects into the inefficiency term, resulting in an unrealistic low average efficiency below 0.5.

The criteria we used for choosing a suitable stochastic frontier model for a given dataset have helped to narrow down the range of applicable models from nine to only one, the GLS+M (IV-M) model. We would not consider this choice to be utterly irrefutable, however, considering our tests in combination with the knowledge on the production process, the general operating environment in the sector and the characteristics of the dataset at hand, empirical researchers should be able to make an educated choice.[60]

2.6. Concluding Remarks

In this paper, we compare the results of a decomposition of total factor productivity growth based on the estimates of nine different commonly used stochastic frontier models. Hereby we focus on the models' ability to take unobserved heterogeneity into account. The main conclusion from this comparison is not surprising: different econometric specifications can lead to very different results. For an unbalanced panel of 974 dairy farms, observed between the years 2000 and 2008, we find substantial differences in the estimated slope parameters of input-, output- and trend-variables, in the resulting distance elasticities and in the individual efficiency scores of the observed firms. These differences lead to an uncertainty in the interpretation of the results. Unstable distance elasticities raise questions about the importance of particular inputs for the production process. In our results, greatly varying returns to scale indicate almost constant returns to scale for some models, while strongly decreasing returns to scale for others. For all models, technical change is positive, but with increasing or constant rates of change. We find great differences in the average efficiency and thus in the potential for productivity improvement, as well as in the individual efficiency scores. While some models produce corresponding efficiency scores, the scores of others do not match at all.

Considering the widespread application of different econometric models for the analysis of productivity change (see above for some examples), we emphasize that the chosen methodology has to fit the characteristics and the structure of the dataset, as well as the purpose of the analysis. Especially if findings are used to state recommendations for regulation and policy, it is crucial to be aware of the consequences of choosing a particular econometric model. We also show how several statistical tests can be used to narrow down the range of appropriate models and hence facilitate the right choice of the model. Finally, the purpose of the study has to be taken into account. As described, the models have different virtues. Hence, the choice of the model depends also on whether the focus lies on individual efficiency scores and their development over time, the slope parameters or, as in case of an analysis of TFP change, both. Since the models are not nested

[60] In a similar situation Karagiannis and Tzouvelekas (2010) recommend to construct averages of the results from competing models. This approach was brought up by Coelli and Perelman (1999) in regard to efficiency analysis.

altogether, it is not possible to find the most appropriate model using formal statistical tests. However, it is possible to narrow the range of models and thereby to facilitate the choice in combination with the aforementioned aspects.

Appendix

Table 2-8: Regression results of the estimated stochastic frontier models

Coefficients	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
(Constant)	0.1022 [a] (0.0088)	0.1072 [a] (0.0106)	- -	-0.0615 [a] (0.0071)	-0.0373 [a] (0.0102)	0.2205 [a] (0.0058)	- -	0.0007 (0.0040)	0.0616 [a] (0.0042)
$\sigma(random\ const.)$	- -	- -	- -	- -	- -	- -	- -	0.2327 [a] (0.0013)	0.1459 [a] (0.0010)
β_1 (Labor)	0.1766 [a] (0.0191)	0.1707 [a] (0.0202)	0.0360 [a] (0.0131)	0.0782 [a] (0.0126)	0.0359 [a] (0.0130)	0.0838 [a] (0.0095)	0.1789 [a] (0.0155)	0.0676 [a] (0.0078)	0.0322 [a] (0.0088)
β_2 (Land)	0.0229 (0.0140)	0.0614 [a] (0.0142)	0.1156 [a] (0.0150)	0.1667 [a] (0.0118)	0.1162 [a] (0.0150)	0.1612 [a] (0.0071)	0.0223 [b] (0.0113)	0.1818 [a] (0.0055)	0.1097 [a] (0.0087)
β_3 (Interm. inputs)	0.5908 [a] (0.0131)	0.5708 [a] (0.0136)	0.3215 [a] (0.0108)	0.4286 [a] (0.0098)	0.3190 [a] (0.0108)	0.4284 [a] (0.0063)	0.5851 [a] (0.0117)	0.3970 [a] (0.0053)	0.3081 [a] (0.0069)
β_4 (Capital)	0.1378 [a] (0.0106)	0.1349 [a] (0.0106)	0.0222 [a] (0.0085)	0.0742 [a] (0.0079)	0.0221 [a] (0.0085)	0.0690 [a] (0.0053)	0.1385 [a] (0.0085)	0.0700 [a] (0.0040)	0.0271 [a] (0.0053)
β_{11}	-0.0494 (0.0418)	-0.0212 (0.0420)	-0.1060 [a] (0.0377)	-0.0829 [b] (0.0360)	-0.1056 [a] (0.0377)	-0.0880 [a] (0.0281)	-0.0530 [c] (0.0279)	-0.0571 [a] (0.0179)	-0.0747 [a] (0.0267)
β_{22}	-0.1136 [a] (0.0249)	-0.1140 [a] (0.0240)	-0.0609 [b] (0.0286)	-0.0599 [b] (0.0263)	-0.0596 [b] (0.0287)	-0.0271 [c] (0.0146)	-0.1188 [a] (0.0165)	-0.0162 (0.0100)	-0.0301 [b] (0.0148)
β_{33}	0.1339 [a] (0.0203)	0.1005 [a] (0.0193)	0.0646 [a] (0.0203)	0.0906 [a] (0.0192)	0.0636 [a] (0.0204)	0.0986 [a] (0.0095)	0.1315 [a] (0.0135)	0.1052 [a] (0.0079)	0.0927 [a] (0.0110)
β_{44}	-0.0487 [a] (0.0129)	-0.0526 [a] (0.0119)	-0.0302 [b] (0.0133)	-0.0481 [a] (0.0124)	-0.0300 [b] (0.0133)	-0.0396 [a] (0.0090)	-0.0474 [a] (0.0087)	-0.0388 [a] (0.0053)	-0.0191 [b] (0.0091)
β_{12}	0.0430 [c] (0.0259)	0.0223 (0.0262)	0.0101 (0.0238)	-0.0054 (0.0226)	0.0101 (0.0238)	0.0055 (0.0160)	0.0539 [a] (0.0175)	-0.0214 [b] (0.0108)	-0.0071 (0.0155)
β_{13}	-0.0210 (0.0230)	-0.0007 (0.0229)	0.0002 (0.0210)	0.0061 (0.0201)	0.0002 (0.0210)	0.0192 (0.0141)	-0.0253 (0.0156)	-0.0104 (0.0094)	0.0004 (0.0140)
β_{14}	-0.0496 [a] (0.0177)	-0.0459 [a] (0.0162)	0.0251 (0.0157)	0.0133 (0.0150)	0.0250 (0.0157)	0.0176 [c] (0.0099)	-0.0548 [a] (0.0122)	0.0005 (0.0068)	0.0158 (0.0099)
β_{23}	-0.0090 (0.0190)	0.0034 (0.0184)	0.0218 (0.0190)	0.0185 (0.0179)	0.0211 (0.0191)	0.0155 (0.0104)	-0.0077 (0.0128)	0.0086 (0.0077)	-0.0020 (0.0107)

β_{24}	0.0767 [a] (0.0131)	0.0663 [a] (0.0119)	0.0047 (0.0138)	0.0260 [b] (0.0128)	0.0037 (0.0138)	-0.0018 (0.0078)	0.0768 [a] (0.0088)	0.0207 [a] (0.0054)	0.0089 (0.0085)
β_{34}	-0.0406 [a] (0.0119)	-0.0294 [b] (0.0117)	0.0008 (0.0124)	-0.0201 [c] (0.0116)	0.0022 (0.0124)	0.0082 (0.0068)	-0.0385 [a] (0.0079)	-0.0358 [a] (0.0047)	-0.0019 (0.0069)
α_1 (Milk)	-0.7186 [a] (0.0132)	-0.7175 [a] (0.0126)	-0.8287 [a] (0.0083	-0.8058 [a] (0.0082)	-0.8284 [a] (0.0083)	-0.7995 [a] (0.0056)	-0.7143 [a] (0.0118)	-0.8095 [a] (0.0048)	-0.8358 [a] (0.0049)
α_2 (Other)	-0.2814 -	-0.2825 -	-0.1713 -	-0.1942 -	-0.1716 -	-0.2005 -	-0.2857 -	-0.1905 -	-0.1642 -
α_{11}	-0.2146 [a] (0.0175)	-0.2150 [a] (0.0141)	-0.2125 [a] (0.0100)	-0.2088 [a] (0.0099)	-0.2125 [a] (0.0100)	-0.2113 [a] (0.0050)	-0.2214 [a] (0.0141)	-0.1683 [a] (0.0046)	-0.1631 [a] (0.0048)
α_{22}	-0.2146 -	-0.2150 -	-0.2125 -	-0.2088 -	-0.2125 -	-0.2113 -	-0.2214 -	-0.1683 -	-0.1631 -
α_{12}	0.2146 -	0.215 -	0.2125 -	0.2088 -	0.2125 -	0.2113 -	0.2214 -	0.1683 -	0.1631 -
δ_{11}	-0.0108 (0.0230)	-0.0194 (0.0210)	-0.0817 [a] (0.0152)	-0.0696 [a] (0.0149)	-0.0815 [a] (0.0152)	-0.0740 [a] (0.0091)	-0.0038 (0.0174)	-0.0736 [a] (0.0082)	-0.0778 [a] (0.0086)
δ_{21}	0.0517 [a] (0.0161)	0.0583 [a] (0.0131)	0.0571 [a] (0.0111)	0.0629 [a] (0.0109)	0.0577 [a] (0.0111)	0.0627 [a] (0.0054)	0.0461 [a] (0.0118)	0.0753 [a] (0.0049)	0.0655 [a] (0.0055)
δ_{31}	-0.0325 [b] (0.0137)	-0.0290 [a] (0.0103)	-0.0332 [a] (0.0095)	-0.0297 [a] (0.0093)	-0.0337 [a] (0.0095)	-0.0258 [a] (0.0038)	-0.0305 [a] (0.0100)	-0.0389 [a] (0.0034)	-0.0361 [a] (0.0039)
δ_{41}	0.0234 [b] (0.0119)	0.0216 [b] (0.0099)	0.0295 [a] (0.0082)	0.0298 [a] (0.0081)	0.0294 [a] (0.0082)	0.0302 [a] (0.0048)	0.0215 [b] (0.0087)	0.0275 [a] (0.0040)	0.0310 [a] (0.0045)
δ_{12}	0.0108 -	0.0194 -	0.0817 -	0.0696 -	0.0815 -	0.074 -	0.0038 -	0.0736 -	0.0778 -
δ_{22}	-0.0517 -	-0.0583 -	-0.0571 -	-0.0629 -	-0.0577 -	-0.0627 -	-0.0461 -	-0.0753 -	-0.0655 -
δ_{32}	0.0325 -	0.029 -	0.0332 -	0.0297 -	0.0337 -	0.0258 -	0.0305 -	0.0389 -	0.0361 -
δ_{42}	-0.0234 -	-0.0216 -	-0.0295 -	-0.0298 -	-0.0294 -	-0.0302 -	-0.0215 -	-0.0275 -	-0.031 -

τ_1 (Trend)	0.0078 [b]	0.0076 [b]	0.0146 [a]	0.0108 [a]	0.0146 [a]	0.0155 [a]	0.0123 [a]	0.0075 [a]	0.0097 [a]
	(0.0036)	(0.0037)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0014)	(0.0017)	(0.0016)
τ_2	0.0013 [c]	0.0012 [c]	-0.0001	0.0004	-0.0001	0.0003	0.0005	0.0011 [a]	0.0010 [a]
	(0.0007)	(0.0007)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
ν_{1t}	-0.0019	-0.0022	0.0050 [a]	0.0029	0.0050 [a]	0.0022	-0.0019	0.0030 [b]	0.0047 [a]
	(0.0033)	(0.0033)	(0.0019)	(0.0019)	(0.0019)	(0.0016)	(0.0028)	(0.0013)	(0.0013)
ν_{2t}	-0.0018	-0.0018	-0.0021	-0.0013	-0.0024 [c]	0.0000	-0.0026	-0.0016 [c]	-0.0023 [a]
	(0.0024)	(0.0024)	(0.0014)	(0.0014)	(0.0014)	(0.0010)	(0.0020)	(0.0009)	(0.0008)
ν_{3t}	0.0094 [a]	0.0100 [a]	0.0053 [a]	0.0040 [a]	0.0057 [a]	0.0016	0.0113 [a]	0.0025 [a]	0.0046 [a]
	(0.0023)	(0.0022)	(0.0013)	(0.0013)	(0.0013)	(0.0010)	(0.0020)	(0.0008)	(0.0008)
ν_{4t}	-0.0020	-0.0025	0.0059 [a]	0.0019 [c]	0.0059 [a]	0.0020 [a]	-0.0020	0.0023 [a]	0.0064 [a]
	(0.0018)	(0.0018)	(0.0011)	(0.0011)	(0.0011)	(0.0008)	(0.0015)	(0.0007)	(0.0007)
ς_{1t}	0.0010	0.0011	0.0074 [a]	0.0054 [a]	0.0073 [a]	0.0043 [a]	0.0013	0.0039 [a]	0.0063 [a]
	(0.0022)	(0.0021)	(0.0013)	(0.0012)	(0.0013)	(0.0009)	(0.0018)	(0.0008)	(0.0007)
ς_{2t}	-0.001	-0.0011	-0.0074	-0.0054	-0.0073	-0.0043	-0.0013	-0.0039	-0.0063
	-	-	-	-	-	-	-	-	-
$\lambda = \sigma_u/\sigma_v$	1.3287 [a]	1.1311 [a]				4.4007 [a]	1.4391 [a]	1.9320 [a]	2.1771 [a]
	(0.0367)	(0.0770)				(0.0073)	(0.0439)	(0.0707)	(0.0736)
$\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$	0.2258 [a]	0.2081 [a]				0.4022	0.3068 [a]	0.1258 [a]	0.1259 [a]
	(0.0000)	(0.0037)				-	(0.0026)	(0.0013)	(0.0012)
σ_u	0.1804	0.1559				0.3922	0.2519	0.1118	0.1145
	-	-				-	-	-	-
σ_ϑ			0.2467	0.1512	0.1407				
			-	-	-				
σ_v	0.1358	0.1378	0.0868	0.0868	0.0869	0.0891	0.1751	.0579	0.0526
	-	-	-	-	-	-	-	-	-
η						-0.0163 [a]			
						(0.0021)			
Log-likelihood	2500.77	2621.33				5712.47	3536.41	5808.43	6322.76

a,b,c: Significantly different from zero on the 1%, 5% and 10% level respectively
Standard errors in parentheses

Table 2-9: Auxiliary parameters of the inefficiency effects and the Mundlak models

BC95 (II)		GLS + Mundlak (IV – M)				TRE + Mundlak (VII – M)			
ζ_{reg2}	-0.1094 [a] (0.0265)	$\gamma_{\bar{x}1}$	0.1292 [a] (0.0263)	$\gamma_{\bar{x}14}$	-0.0697 (0.0534)	$\gamma_{\bar{x}1}$	0.1201 [a] (0.0080)	$\gamma_{\bar{x}14}$	-0.0546 [a] (0.0144)
ζ_{reg3}	-0.0720 [c] (0.0380)	$\gamma_{\bar{x}2}$	-0.1209 [a] (0.0207)	$\gamma_{\bar{x}23}$	-0.0211 (0.0556)	$\gamma_{\bar{x}2}$	-0.1175 [a] (0.0077)	$\gamma_{\bar{x}23}$	0.0265 [c] (0.0151)
ζ_{reg4}	-0.0660 (0.0673)	$\gamma_{\bar{x}3}$	0.3352 (0.0190)	$\gamma_{\bar{x}24}$	0.1062 [a] (0.0379)	$\gamma_{\bar{x}3}$	0.3583 [a] (0.0063)	$\gamma_{\bar{x}24}$	0.1225 [a] (0.0114)
ζ_{reg5}	0.0898 [a] (0.0169)	$\gamma_{\bar{x}4}$	0.0728 [a] (0.0145)	$\gamma_{\bar{x}34}$	-0.0663 [c] (0.0339)	$\gamma_{\bar{x}4}$	0.0686 [a] (0.0047)	$\gamma_{\bar{x}34}$	-0.0716 [a] (0.0095)
ζ_{reg6}	0.1225 [a] (0.0193)	$\gamma_{\bar{x}11}$	0.0331 [a] (0.1212)	$\gamma_{\bar{y}1}$	0.1317 [a] (0.0204)	$\gamma_{\bar{x}11}$	-0.0149 (0.0360)	$\gamma_{\bar{y}1}$	0.1227 [a] (0.0050)
ζ_{reg7}	0.1266 [a] (0.0226)	$\gamma_{\bar{x}22}$	-0.1178 (0.0728)	$\gamma_{\bar{y}11}$	-0.1321 [c] (0.0773)	$\gamma_{\bar{x}22}$	-0.1702 [a] (0.0201)	$\gamma_{\bar{y}11}$	-0.1673 [a] (0.0161)
ζ_{reg8}	0.0157 (0.0363)	$\gamma_{\bar{x}33}$	0.0809 (0.0591)	$\gamma_{\overline{xy}11}$	0.0924 (0.0853)	$\gamma_{\bar{x}33}$	-0.0083 (0.0158)	$\gamma_{\overline{xy}11}$	0.0589 [a] (0.0193)
ζ_{reg9}	0.2365 [a] (0.0283)	$\gamma_{\bar{x}44}$	-0.0237 (0.0369)	$\gamma_{\overline{xy}21}$	-0.0962 [c] (0.0582)	$\gamma_{\bar{x}44}$	-0.0348 [a] (0.0118)	$\gamma_{\overline{xy}21}$	-0.1539 [a] (0.0129)
		$\gamma_{\bar{x}12}$	0.0677 (0.0759)	$\gamma_{\overline{xy}31}$	0.0388 (0.0514)	$\gamma_{\bar{x}12}$	0.0560 [a] (0.0214)	$\gamma_{\overline{xy}31}$	0.0749 [a] (0.0109)
		$\gamma_{\bar{x}13}$	-0.0560 (0.0687)	$\gamma_{\overline{xy}41}$	0.0060 (0.0422)	$\gamma_{\bar{x}13}$	-0.0122 (0.0192)	$\gamma_{\overline{xy}41}$	0.0152 (0.0097)

a,b,c: Significantly different from zero on the 1%, 5% and 10% level respectively
Standard errors in parentheses

Table 2-10: Year to year average percentage TFP change

	Pooled (I)	BC95 (II)	FE (III)	GLS (IV)	GLS+M (IV-M)	BC92 (V)	TFE (VI)	TRE (VII)	TRE+M (VII-M)
Year	Total factor productivity change (TFPC)								
00/01	1.253	1.156	1.482	1.187	1.479	1.117	1.547	1.849	2.143
01/02	0.769	0.802	0.417	0.729	0.412	0.676	0.880	0.347	-0.029
02/03	1.031	1.000	0.665	0.859	0.662	0.913	1.298	0.972	0.883
03/04	0.481	0.492	0.325	0.666	0.322	0.773	0.458	0.253	0.106
04/05	1.671	1.619	1.623	1.457	1.625	1.346	1.718	1.517	1.661
05/06	1.544	1.499	1.253	1.280	1.254	1.257	1.543	1.257	1.352
06/07	2.168	2.056	0.941	1.065	0.941	1.190	2.216	2.097	2.184
07/08	0.965	0.966	1.531	1.342	1.534	1.610	0.604	-0.036	-0.040
Year	Technical change (TC)								
00/01	0.848	0.808	1.267	1.030	1.261	1.516	1.159	0.850	0.954
01/02	0.993	0.946	1.277	1.084	1.273	1.560	1.231	0.970	1.070
02/03	1.140	1.088	1.239	1.111	1.237	1.580	1.305	1.069	1.142
03/04	1.301	1.243	1.265	1.177	1.266	1.629	1.399	1.198	1.271
04/05	1.444	1.380	1.258	1.225	1.261	1.669	1.468	1.313	1.370
05/06	1.569	1.497	1.231	1.256	1.235	1.696	1.514	1.417	1.454
06/07	1.715	1.637	1.207	1.288	1.213	1.722	1.589	1.521	1.537
07/08	1.847	1.761	1.192	1.329	1.199	1.756	1.644	1.631	1.631
Year	Technical efficiency change (TEC)								
00/01	0.385	0.336	0.146	0.112	0.151	-0.493	0.368	1.004	1.140
01/02	-0.120	-0.057	0.018	0.019	0.021	-0.498	-0.245	-0.199	-0.232
02/03	-0.044	-0.032	-0.125	-0.065	-0.124	-0.509	0.057	0.155	0.193
03/04	-0.699	-0.654	-0.192	-0.161	-0.193	-0.510	-0.819	-0.522	-0.418
04/05	0.203	0.224	0.049	0.078	0.046	-0.525	0.229	0.068	-0.063
05/06	0.023	0.044	-0.040	0.008	-0.044	-0.526	0.078	-0.112	-0.178
06/07	0.494	0.457	-0.226	-0.199	-0.231	-0.524	0.671	0.633	0.662
07/08	-0.925	-0.833	-0.272	-0.277	-0.279	-0.531	-1.074	-1.944	-2.313
Year	Scale change effect (SC)								
00/01	0.019	0.012	0.068	0.045	0.067	0.094	0.021	-0.005	0.048
01/02	-0.104	-0.087	-0.878	-0.374	-0.881	-0.386	-0.106	-0.425	-0.867
02/03	-0.065	-0.056	-0.449	-0.187	-0.451	-0.157	-0.065	-0.252	-0.452
03/04	-0.121	-0.098	-0.748	-0.349	-0.751	-0.347	-0.122	-0.424	-0.747
04/05	0.023	0.015	0.317	0.154	0.318	0.203	0.021	0.136	0.354
05/06	-0.048	-0.042	0.062	0.017	0.063	0.087	-0.049	-0.049	0.076
06/07	-0.041	-0.037	-0.040	-0.025	-0.041	-0.008	-0.044	-0.057	-0.015
07/08	0.044	0.037	0.611	0.290	0.614	0.385	0.034	0.277	0.642

3. ACCOUNTING FOR ENDOGENOUS EFFECTS IN STOCHASTIC FRONTIER MODELS

Abstract

In this paper we propose several extensions for stochastic frontier models, to take firm heterogeneity into account and to reduce heterogeneity-induced biases in the estimated efficiency scores and technology parameters. We broaden the approach by Farsi et al. (2005) and consider the incorporation of an auxiliary equation with the group means of the input variables as an actual modeling of between-firm heterogeneity. A generalized specification of this auxiliary equation is proposed, including also environmental variables and considering that not all group-mean variables are necessarily correlated with firm heterogeneity, and adapted to three different stochastic frontier models for panel data. The results of an empirical application show that the proposed specifications help to reduce the heterogeneity bias in the estimated technology parameters and technical efficiency scores.

Keywords: heterogeneity, panel data, stochastic frontier

JEL Classification: C23, D24

This chapter is based on the article *Accounting for endogenous effects in stochastic frontier models* by Giannis Karagiannis and Magnus Kellermann. The authors share the main authorship.

3.1. Introduction

The true effects models developed by Greene (2005, 2005a) are undoubtedly one of the most important recent contributions in stochastic frontier analysis for panel data. Their main advantage is that they allow individual effects to exist alongside inefficiency in such a way that we can distinguish between heterogeneity and technical efficiency, thereby providing more accurate performance evaluations. Standard stochastic frontier models have failed in one way or another to address this issue.[61] On the one hand, stochastic frontier models making no distributional assumptions about the one-sided error term capturing technical inefficiency (i.e., Schmidt and Sickles, 1984) confound heterogeneity and inefficiency, as whatever is not accounted for by factor inputs is attributed to technical inefficiency (Greene, 2004). On the other hand, stochastic frontier models making distributional assumptions (e.g., Pitt and Lee, 1981) have simply ignored (assumed away) heterogeneity.

The direct consequence of these modeling limitations is inaccurate estimates of efficiency. In particular, for those stochastic frontier models making no distributional assumptions, not distinguishing between heterogeneity and efficiency may lead to an upward (downward) bias of technical inefficiency for production units that are subject to unfavorable (favorable) individual effects. In random effects specifications (Schmidt and Sickles, 1984), omitted variable biases in the technology parameters must be expected. For those stochastic frontier models making distributional assumptions, neglecting the heterogeneity that is asymmetrically distributed among production units also results in biased estimated technology parameters, significantly inflated estimates of technical inefficiency (Sherlund et al. 2002) and the increased dispersion of efficiency scores (Tybout, 2000).[62]

However, the true effects models that are able to distinguish between heterogeneity and inefficiency are unfortunately not without problems. On the one hand, true fixed effects models produce biased estimates of individual (fixed) effects and of firm-specific efficiency scores (Abdulai and Tietje, 2007).[63] On the other hand, true random effects models result in biased technology parameter estimates, as most of the unobserved factors

[61] Good et al. (1993), Coelli et al. (1999), and Sherlund et al. (2002) are notable exceptions, since they included environmental variables to account for heterogeneity.

[62] Sherlund et al. (2002) also mentioned that not controlling for unobserved heterogeneity also results in an upwards bias in the estimates of technical inefficiency for deterministic frontier models.

[63] For the purely estimation-related problems of the true fixed effect model, see Greene (2005a, b) and Wang and Ho (2010).

are likely to be correlated with the explanatory variables, i.e., input quantities. The inconsistent estimates of the technology (slope) parameters will then bias the estimated variance of the composed error term used in the Jondrow *et al*. (1982) procedure to estimate technical efficiency. In addition, because all time-invariant effects are incorporated within the individual effects, we cannot account for any persistent inefficiency; thus, technical inefficiency tends to be underestimated (Last and Wetzel, 2010).

Under these circumstances, the true random effects model with Mundlak's (1978) adjustment, proposed by Farsi et al. (2005) and Farsi et al. (2005a), appears to be a promising alternative, as it can reduce the heterogeneity bias in both, technology (slope) parameters and inefficiency estimates at the same time. The *Mundlak true random effects* (*M-TRE) model* is based on the assumption that unobserved heterogeneity is correlated with the group means of the explanatory variables, and in the case of production frontiers, this assumption refers to input quantities and the technical change index.[64] By controlling for (some of) the unobserved heterogeneity and separating the correlation effects, the M-TRE model tends to decrease the bias in the inefficiency estimates without affecting the consistency of the estimated technology parameters. In fact, when the error term in the estimated equation is a composite asymmetric term, as with stochastic frontier models, the heterogeneity bias will be reduced, given that the correlation between the individual effects and the explanatory variables is partly captured in the model (Farsi et al. 2005a).

This paper attempts to contribute to this strand of the literature by proposing two alternative models in the spirit of the M-TRE model.[65] In both of these models, we try to further improve the ability of the true random effects models to account for heterogeneity by further enlarging the set of potential correlates to increase the portion of measured heterogeneity and to squeeze the impact of heterogeneity bias on the estimated technology parameters and technical efficiency. The first model allows heterogeneity to be correlated with the group means of input quantities and with a set of relevant environmental variables that are beyond producers' control but account for the operating conditions with which the production units have to cope and which most likely tend to differ among firms. The second model takes even a boarder view by assuming that heterogeneity may not be correlated with all of the technology-related explanatory variables. For example,

[64] The explanatory variables in a cost frontier are input prices, output quantities and the technical change index; in a profit frontier, they include input and output prices and the technical change index.

[65] Both the proposed models and the M-TRE model itself fall in the boarder category of hierarchical or multilevel models, as referred to by Greene (2004).

heterogeneity is most likely uncorrelated with the neutral component of disembodied technical change as long as production units are not involved in R&D activities, as is the case in agriculture and several service industries. The same may also be true for a subset of environmental variables, such as those affecting the operating conditions in a uniform way, even if they change over time (e.g., policy variables). As a common feature of both of the proposed models, it is worth mentioning that each employs a more general specification of endogenous individual effects than the M-TRE model.[66]

Furthermore, the true-effects type of models, which account for endogenous individual effects, give rise to several new variants of standard stochastic frontier models. We present two such variants without distributional assumptions. These variants can be viewed as extensions of the fixed- and the random-effects stochastic frontier models introduced by Schmidt and Sickles (1984) and developed further by Good *et al*. (1993) to account for environmental variables and by Farsi et al. (2005a) to incorporate Mundlak's adjustments. Moreover, we propose three additional models for making distributional assumptions that complement previous attempts by Coelli et al. (1999) and Sherlund et al. (2002) to account for environmental factors in a maximum-likelihood type of stochastic frontier models.

The main reason for incorporating endogenous individual effects into applied efficiency analysis is that these effects may improve econometricians' abilities to account for the heterogeneity that is unobserved by them but is observed by producers, who adjust their input decisions in accordance with underlying environmental factors. We expect that at least part of what is considered to be individual heterogeneity may be accounted for by utilizing some control variables. These variables include environmental factors affecting firms' operating conditions and Mundlak's adjustment terms, which in the case of production frontiers, correspond to individual means of input quantities. The only difference is that in the proposed conventional stochastic frontier models, it is assumed that all heterogeneity is captured by the control variables, whereas in the true effects models, the included control variables account only for part of the individual effects. The remaining effects are treated as unobserved heterogeneity by means of random effects.

We provide an empirical evaluation of the proposed models using a dataset of a sample of specialized German dairy farms observed for the period 2003–2008. This example is especially appealing because agricultural sector firms (farms) are prone to

[66] The term "endogenous effects" has been used by Baltagi (1995, p. 116-120) to refer to the Mundlak (1978) and Hausman and Taylor (1981) types of panel data models.

heterogeneous production conditions. For this reason, models that fail to account for different production and environmental conditions have been criticized in the literature; see for example Sherlund et al. (2002) and Abdulai and Tietje (2007). We explicitly account for this and provide some empirical evidence as to the extent of the difference in the estimated technical efficiency scores by using alternative specifications to account for endogenous individual effects and/or different model assumptions.

The rest of paper is organized as follows: in the next section, we present the proposed model specifications and outline their estimation procedures. The data and the empirical model are described in the third section. The comparative empirical results are discussed in the fourth section. Concluding remarks follow in the last section.

3.2. Models Specification and Estimation

Following Greene (2005, 2005a) let the true effects stochastic production frontier model be given as follows:

$$y_{it} = \beta_0 + f(x_{it}) + \alpha_i - u_{it} + v_{it} \tag{3.1}$$

where *i* is used to index production or decision making units and *t* time periods, *y* refers to (the log of) output quantity and *x* to (the log of) input quantities, $f(\bullet)$ is the functional form of the production function not including the intercept term β_0, α_i represents firm-specific (individual) effects or unobserved heterogeneity, $-u_{it}$ is a one-sided non-negative error term measuring (the log of) technical efficiency, and v_{it} is a symmetric and normally distributed error term, which plays the role of statistical noise accounting for unanticipated production shocks that producers do not observe when making their input decisions and for econometricians' weaknesses related to omitted explanatory variables, measurement errors in the dependent variable, and functional form discrepancies. In addition, the following distributional assumptions are made for these error terms: that both are independent and identically distributed (*iid*) as $v_{it} \sim N(0, \sigma_v^2)$ and that $u_{it} \sim N^+(0, \sigma_u^2)$ (half-normal) and is also uncorrelated with (or distributed independently of) input quantities *x* and one another.

The former presupposes that technical efficiency is not known to producers before they make their input decisions.[67] In such a case where producers seek to maximize the expected profits, the quantities of (variable) inputs are largely predetermined and are

[67] If a firm knew its level of technical efficiency at the time it makes its production decisions, input choice would be affected (Schmidt and Sickles, 1984).

hence uncorrelated with technical efficiency (Zellner et al., 1966).[68] Moreover, technical efficiency is assumed to be stochastically (randomly) time-varying, as no particular form is specified for its time pattern. In other words, inefficiency is not persistent, and "each period brings about new idiosyncratic elements thus new sources of inefficiency. This is a reasonable assumption, particularly in industries that are constantly facing new technologies" (Farsi et al., 2005, p. 77). The implication of this assumption is that the observations from the same production unit are considered independent sample points.

Following Greene (2005, 2005a), there are alternative ways of modelling the firm-specific effects in (1). One such method is to treat α_i as fixed effects that are correlated with input quantities, even though α_i and x_{it} are assumed to be uncorrelated with both u_{it} and v_{it}. This specification corresponds to the *true fixed effects (TFE)* model. Another approach is to consider heterogeneity as an independent and identically distributed (*iid*) variable with $\alpha_i \sim N(0, \sigma_\alpha^2)$. This approach results in the *true random effects (TRE) model*, in which u_{it}, v_{it} and α_i are assumed to be uncorrelated with the input quantities and each other.

On the other hand, the M-TRE model as given by Farsi et al. (2005) and Farsi et al. (2005a), is based on the assumption that individual effects are a linear function of the group means of all explanatory variables (i.e., input quantities) across time, namely:

$$\alpha_i = \pi' \bar{x}_i + \delta_i \tag{3.2}$$

where π are parameters to be estimated and a bar over a variable denotes its group mean, i.e., $\bar{x}_i = (1/T_i)\sum x_{it}$. By substituting (3.2) into (3.1), we obtain the estimable equation of the M-TRE model:

$$y_{it} = \beta_0 + f(x_{it}) + \pi' \bar{x}_i + \delta_i + e_{it} \tag{3.3}$$

where $e_{it} = -u_{it} + v_{it}$ is a composite asymmetric error term equal to the sum of two orthogonal error terms, one reflecting inefficiency and the other statistical noise. In estimating (3.3), Farsi et al. (2005) and Farsi et al. (2005a) have treated δ_i as pure unobserved or residual heterogeneity, reflecting that part the of firm-specific (individual) effects that cannot be explained by observed factors, namely, the group means of input use. To deal practically with this unobservable part the authors assumed that it was independent and identically distributed (*iid*) as $\delta_i \sim N(0, \sigma_\delta^2)$. As a random effects model,

[68] For comparison purposes, we maintain this assumption even for the GLS models presented later in this section, restricting ourselves to random effects GLS models.

(3.3) assumes that e_{it} and thus (v_{it} and technical inefficiency) are uncorrelated with pure unobserved heterogeneity δ_i and input use. However, with some fixed-effects elements being inherent through Mundlak's adjustment, α_i is found to be correlated with the group means of input use as in (3.2), which is expected to deteriorate the heterogeneity bias. Under this setup, the M-TRE model is estimated with simulated maximum likelihood (see Greene, 2005, 2005a), and technical efficiency estimates are obtained as $E(u_{it}|e_{it})$ using the Jondrow et al. (1978) estimator (Greene, 2004).[69]

A potential limitation of the M-TRE model is that it accounts only for heterogeneity reflected in the firms' level of input use. The impact of other variables affecting heterogeneity, such as those related to "environmental factors", is not accommodated in the M-TRE model.[70] We consider as environmental "those factors which (…) are taken as not within the management's field of choice" (Hall and Winsten, 1959, p. 72) and thus are not under its control. For instance, these factors account for the operating conditions with which the production units have to cope, as they most likely tend to differ among firms and perhaps even across time.

Although there are industries, such as banking and semiconductors, in which firms have considerable control over their operating conditions (Sherlund et al., 2002), several other examples exist for which this is not the case. For instance, the population density and per capita income might be considered as two environmental factors in evaluating the performance of retail distribution systems (Hall and Winsten, 1959). Similarly, according to Coelli et al. (1999), several geographic and demographic features of the regions or countries served are important to accurately estimate efficiency in the airline industry. Moreover, agro-ecological conditions, topography, pest infestation and (plant or animal) diseases are important environmental factors in estimating farm efficiency (see Mundlak, 1961; Sherlund et al., 2002).

To accommodate these concerns into (3.2), we extend the M-TRE model in two directions: *first*, we adopt a variant of Mundlak's model that first appeared in Maddala (1987) and includes time-invariant environmental factors within the auxiliary equation for explaining individual effects; and *second*, we incorporate time-varying environmental

[69] Notice however that the resulting estimated parameters are not the within estimates, as in the original Mundlak (1978) model because e_{it} is a composite asymmetric rather than a symmetric error term (Farsi et al., 2005a).

[70] The discussion on the role of "environmental factors" in efficiency measurement is dated back to Hall and Winsten (1959) and Mundlak (1961).

factors into the production function $f(\bullet)$. With these modifications the auxiliary equation (3.2) may be written as follows:

$$\alpha_i = \pi'\bar{x}_i^* + \gamma' z_i + \delta_i \tag{3.4}$$

where γ are parameters to be estimated, $x_{it}^* = (x_{it}, z_{it})$, and z is the set of the relevant environmental factors. Thus, with (3.4) we attempt to improve the ability of the TRE model to account for heterogeneity by including a set of relevant environmental factors along with firm-specific means of input quantities. By substituting (3.4) into (3.1), we obtain the following:

$$y_{it} = \beta_0 + f(x_{it}^*) + \pi'\bar{x}_i^* + \gamma' z_i + \delta_i + e_{it} \tag{3.5}$$

which we refer to as the *Mundlak-Maddala true random effects* (*MM-TRE) model*. As the M-TRE model, the MM-TRE model can be estimated with simulated maximum-likelihood, and technical efficiency estimates can be obtained as $E(u_{it} | e_{it})$ using the Jondrow et al. (1978) estimator. From (3.5) and (3.3), that the M-TRE is nested in the MM-TRE model can be verified. Hence, the decision as to which model should be used for the data at hand could be based on the likelihood ratio test.

Yet another modeling alternative emerges by noticing that the firm-specific effects might be correlated with some but not necessarily *all* technology-related and environmental variables. The idea of distinguishing between variables that are potentially correlated or uncorrelated with firm effects was introduced in a completely different setup by Hausman and Taylor (1981).[71] For example, individual effects are most likely uncorrelated with quasi-fixed inputs and neutral disembodied technical changes. According to Griliches and Mairesse (1998, p. 385), "if one accepts the notion that the quasi-fixed inputs are predetermined for the duration of the relevant observation period", then their quantity is uncorrelated with the individual effects. Similarly, the neutral part of disembodied technical change with a common impact (i.e., shift) on the production technology of all producers is most likely thought of as being uncorrelated with heterogeneity; the same though is not true for biased technical change, as it depends on input quantities. However, all symmetrically distributed time-varying environmental variables may have an impact on the position of the production function; however, this impact will be uniform for all producers. Such variables could be dummy variables

[71] Hausman and Taylor (1981) use the exogenous variables to derive valid instruments for the endogenous variables. We, however, just borrow the notion to split the set of variables in exogenous and endogenous variables.

reflecting policy changes and even rainfall measures for tightly defined geographical areas. However, under certain conditions the uncorrelatedness assumption can also apply to input quantities. A concrete example is a study on the US airline industry by Cornwell et al. (1990), who assumed that labor and material inputs were uncorrelated with individual effects for their illustration of a Hausman-Taylor type, efficient instrumental variable estimator.

We accommodate these aspects by partitioning the x_{it}^* vector into a group of technology and environmental variables $x_{1it}^* = (x_{1it}, z_{1it})$ that are correlated with heterogeneity (firm-specific effects) and another group $x_{2it}^* = (x_{2it}, z_{2it})$ that are uncorrelated. Thus, we can write the auxiliary equation as follows:

$$\alpha_i = \pi_1{}'\bar{x}_{1i.}^* + \gamma' z_i + \delta_i \tag{3.6}$$

where π_1 and γ are parameters to be estimated. By substituting (3.6) into (3.1), we obtain the following:

$$y_{it} = \beta_0 + f(x_{it}^*) + \pi_1'\bar{x}_{1i}^* + \gamma' z_i + \delta_i + e_{it} \tag{3.7}$$

which we refer to as the *Hausman-Taylor true random effects* (*HT-TRE) model*. As for the other two models, the HT-TRE model is estimated with simulated maximum-likelihood and the technical efficiency estimates are obtained as $E(u_{it}|e_{it})$ using the Jondrow et al. (1978) estimator. The empirical validity of the HT-TRE model can be tested against both the M-TRE and the MM-TRE models using the likelihood ratio test.

In the models above, we have followed Greene (2005, 2005a) by interpreting the δ_i term in (3.3), (3.5) and (3.7) as pure unobserved heterogeneity. Therefore, we have assumed that the group means of input quantities and environmental factors only partially explain their individual effects. Alternatively, the auxiliary equations in (3.2), (3.4) and (3.6) could be thought of as pure regression equations, with δ_i representing statistical noise and the individual effects being related to a vector of observed characteristics, including environmental factors and group means of input quantities, subject to statistical error. This is equivalent to saying that if all of the relevant variables used to control for heterogeneity in (3.2), (3.4) and (3.6) are included, then the measured heterogeneity can account for individual effects up to a statistical error.

Thus, $v_{it}+\delta_i$ is an independent and identically distributed (*iid*) error term with zero mean and constant variance i.e.: $(v_{it}+\delta_i) \sim iid\ N\left(0,\sigma_v^2+\sigma_\delta^2\right)$. Equations (3.3), (3.5) and (3.7) may be adjusted accordingly as follows:

$$y_{it}=\beta_0+f(x_{it})+\pi'\overline{x}_i+\varepsilon_{it} \tag{3.8}$$

$$y_{it}=\beta_0+f(x_{it}^*)+\pi'\overline{x}_i+\gamma' z_i+\varepsilon_{it} \tag{3.9}$$

$$y_{it}=\beta_0+f(x_{it}^*)+\pi_1'\overline{x}_{1i}^*+\gamma' z_i+\varepsilon_{it} \tag{3.10}$$

where $\varepsilon_{it}=(v_{it}+\delta_i)-u_{it}$ is a composite asymmetric error term equal to the sum of two orthogonal error terms, one reflecting normal distributed statistical noise (the sum of the two terms in the parenthesis) and the other reflecting half-normal distributed technical inefficiency. Because it is necessary to make distributional assumptions to estimate models (3.8)-(3.10), we have followed the same line of reasoning as above in referring to them as the M-MLE, MM-MLE and HT-MLE models, respectively. These models complement and enrich previous works by Coelli et al. (1999) and Sherlund et al. (2002) in accounting for environmental factors in maximum likelihood estimated (MLE) stochastic frontier models.

In contrast to the true effects models in (3.3), (3.5) and (3.7) and keeping track of the traditions of MLE stochastic frontier models for panel data, we assume that technical inefficiency is deterministically time-varying. In particular, we adopt the specification of Battese and Coelli (1992): $u_{it}=\exp(-\eta(t-T)u_i$, where η is a parameter to be estimated and u_i is time invariant technical efficiency that is assumed to be independent and identically distributed (*iid*) with $N^+(0,\sigma_u^2)$. If the estimated value of η is positive (negative), technical efficiency tends to improve (deteriorate) over time, whereas if $\eta=0$, technical efficiency is time-invariant. This specification allows inefficiency to evolve smoothly over time, with its movement being monotonic and standardized for all production units, depending on their level of inefficiency. Moreover, the most efficient firm does not change over time. Still, these two aspects of Battese and Coelli's (1992) specification of time varying technical efficiency are consistent with the stylized facts, as documented in Bartelsman and Doms' (2002) survey, on the uniformity of changes in efficiency across production units and on the persistence of efficiency differentials over time.

In the previous models, we have treated the δ_i-term as either pure unobserved heterogeneity or statistical noise. These are not, however, the only interpretations available. A third alternative offered by Hay and Liu (1997) views δ_i and u_{it} as two distinct components of technical inefficiency, namely long- and short-run. Long-run inefficiency reflects persistent differences in the quality of management due to innate abilities and business experience, as well as differences in the ability or lack of expertise to utilize the available technology. However, short-run inefficiency is allowed to vary over time because we expect management to raise its efforts in response to internal and external (i.e., competitive) pressures. The aforementioned interpretation of long-run inefficiency can be accommodated by treating δ_i as a one-sided error term that is independent and identically distributed (*iid*) with constant mean (which is positive) and variance.[72]

However, if the assumption of time-invariant efficiency is tenable, then we can assume away short-run (time-varying) technical inefficiency; that is, $u_{it} = 0$. The idea of time-invariant inefficiency is inherent in Jovanovic's (1982) "passive learning" model, wherein firms are "born" with a fixed efficiency level that they learn over time (Bartelsman and Doms, 2000). Then, the firms that are endowed with a relatively low efficiency level eventually have to exit the market, while the surviving firms exhibit efficiency persistence. Under these circumstances and following the initiative of Farsi et al. (2005a), who proposed (3.11) below, we can extend the class of random effect models suggested by Schmidt and Sickles (1984) with the following three models by adjusting (3.3), (3.5) and (3.7) accordingly as follows:

$$y_{it} = \beta_0^* + f(x_{it}) + \pi'\overline{x}_i + \omega_{it} \tag{3.11}$$

$$y_{it} = \beta_0^* + f(x_{it}^*) + \pi'\overline{x}_i^* + \gamma' z_i + \omega_{it} \tag{3.12}$$

$$y_{it} = \beta_0^* + f(x_{it}^*) + \pi_1'\overline{x}_{1i}^* + \gamma' z_i + \omega_{it} \tag{3.13}$$

where $\beta_0^* = \beta_0 - E(\delta_i)$ and $\omega_{it} = v_{it} - [\delta_i - E(\delta_i)]$ is an independent and identically distributed (*iid*) error term with zero mean and constant variance, and E refers to the expectation operator.[73] The above three models, which we refer to as M-GLS, MM-GLS, and HT-GLS, can be estimated with GLS along the lines suggested by Schmidt and

[72] Even though not considered here, this interpretation of long-run inefficiency can in principle be accommodated in Greene's (2005, 2005a) true effects models.

[73] Notice that (12) may result from the simple combination of the formulations of Good et al. (1993) and Farsi et al. (2005a).

Sickles (1984) without having to make distributional assumptions about the δ_i term, as in the true effects models. Moreover, the choice among the three aforementioned models can be made by a Wald test on the joint significance of the additional variables. Firm-specific estimates of technical efficiency are obtained by first computing the residuals ξ_{it} as $y_{it} - \hat{f}(x_{it}) - \hat{\pi}'\overline{x}_i$, $y_{it} - \hat{f}(x_{it}^*) - \hat{\pi}'\overline{x}_i^* - \hat{\gamma}'z_i$, and $y_{it} - \hat{f}(x_{it}^*) - \hat{\pi}_1'\overline{x}_{1i}^* - \hat{\gamma}'z_i$ using, respectively, the estimated parameters (denoted by a hat) of the models (3.11), (3.12) and (3.13). From those residuals ξ_{it}, we can recover estimates of the individual-firm intercepts in (3.11), (3.12) and (3.13) as $\hat{\beta}_i = (1/T)\sum \xi_{it}$ (Schmidt and Sickles, 1984). Then, using the normalization $\hat{\beta}_0 = \max(\hat{\beta}_i)$, we may derive estimates of the firms inefficiency from $\hat{\delta}_i = \hat{\beta}_0 - \hat{\beta}_i$.

3.3. Empirical Model and Data Description

We applied the three types of models discussed in the previous section to a dataset of 466 Bavarian farms specialized in dairy production, observed over the period 2003–2008 (2409 observations).[74] These three models comprised namely the "true" random effects model (TRE) (Greene, 2005; 2005a), the GLS random effects model (GLS) (Schmidt and Sickles, 1984) and the maximum likelihood random effects model (MLE), which allowed allowing for time-varying technical efficiency (Battese and Coelli, 1992). We estimated the three models in their well-known basic specifications (B), as well as the three discussed extensions: the Mundlak- (M), the Mundlak-Maddala- (MM) and the Hausman-Taylor- (HT) type specifications. Additional specifications have built on the work by Good et al. (1993) in the context of a GLS stochastic frontier model and by Coelli et al. (1999) in the context of MLE stochastic frontier models. Both models include environmental variables within the estimation equation to account for heterogeneous production conditions. Without commenting on those specifications in depth, we added this "Environmental-Variables"- (E) specification to all models to make our comparison more complete. Hence, we estimated a total of 15 different stochastic frontier models and have examined their respective results.

To keep the empirical application simple, we employed a single-output, multiple input Cobb-Douglas production function to represent the production technology. However, all of the discussed models are in no way limited a priori to this functional form.

[74] The term "specialized" states that at least 75% of the farm's revenue must come from dairy production.

We aggregated the farms' outputs into one variable and their inputs into four variables. The output variable included the farms total revenue from dairy production. As input variables we considered the following: i: *labor*, measured in full-time equivalents (including family labor and hired workers); ii: *land*, measured in ha; iii: *material*, including expenses for forage production, veterinary services, purchased feed and other related expenses; and iv: *capital*, including the end-of-year value of all farm-related machinery, equipment and buildings as well as the livestock. All monetary values were deflated using appropriate price-indices obtained from the German Federal Statistical Office (Destatis, 2012). We show the descriptive statistics of the input and output variables in the upper portion of table 3-1.

Table 3-1: Descriptive statistics of input-, output- and environmental variables

		Mean	S.D.	Min	Max
Output	Revenues (1000 €)	110.3	53.6	13.4	402.6
Inputs	Labor (fte)	1.61	0.49	0.30	3.86
	Land (ha)	52.6	28.8	5.6	318.3
	Material (1000 €)	57.7	30.7	5.7	247.1
	Capital (1000 €)	223.6	132.6	12.9	1063.5
Environmental variables	share of owned land (%)	48.59	24.10	0.00	100.00
	ag. Region 1[1]	8.3			
	ag. Region 2[1]	2.7			
	ag. Region 3[1]	33.7			
	ag. Region 4[1]	43.0			
	ag. Region 5[1]	6.9			
	ag. Region 6[1]	5.4			
	part time farming[1]	6.5			

[1]Dummy variable

The basic specification of all three model-types included just the four inputs alone as explanatory variables. For the Mundlak specification of the three models (as specified by equations (3.3), (3.8) and (3.11)) we constructed an additional set of variables $\bar{x}_{ik} = \frac{1}{T_i} x_{itk}\ \forall\ i$, representing the respective group means of the initial input variables, where $k = 1, \dots, 4$. For the Mundlak-Maddala specification (equations (3.5), (3.9) and (3.12)), we defined a set of environmental variables z to further improve the models' ability to account for heterogeneity. In our application, this set contains time-invariant dummy variables for part-time farming and dummy variables indicating the location of the farm in well-defined agricultural production regions in Bavaria[75]. We also included a time-varying variable z_{it}, which represents the share of owned land. Other examples of time-varying variables could be the average field size or the share of permanent grassland. The descriptive statistics for the environmental variables used are given in the lower part of table 1. These variables were also used for the “Environmental-Variables”-specification.

The intuition that only some and not all environmental and technology parameters are correlated with unobserved heterogeneity was the basis of the Hausman-Taylor type models specified in equations (3.7), (3.10) and (3.13). In principle, any technology- or environmental variable could be assumed to be independent of unobserved heterogeneity. However, for many variables – especially inputs – this is a strong assumption that has to be based on conclusive arguments. As previously discussed – there are also natural candidates, such as quasi-fixed inputs or the neutral part of disembodied technical change. We cannot justify the uncorrelatedness assumption for any of the above-mentioned input or environmental variables in our dataset. Hence, for illustrative purposes, we used this fact to construct a very simple example of the HT-type specification by simply adding a linear trend variable to the estimation equation, which by construction is uncorrelated with time invariant effects.

3.4. Discussion of Empirical Results

As a first assessment of the endogeneity of the individual effects, we tested the null hypotheses $H_0\colon \alpha_i \perp x_{itk}$ by performing a Hausman test on the basic GLS random effects model (B-GLS). This test clearly rejected the hypotheses of no correlation between the effects and the input variables, with a test statistic of 317.1 against a critical value of $\chi^2_{(4;0.01)} = 13.3$. Even if this test only applies to the GLS random effects model, the result indicates clearly that all models assuming exogenous random effects will generate biased coefficients for the technology parameters. We present the estimated coefficients, as derived from the discussed models (TRE, GLS and MLE), in tables 3-2 – 3-4. In each table, we have five columns. The first column contains the result of the basic specification, followed by the Environmental-Variables-, the Mundlak-, the Mundlak-Maddala- and the Hausman-Taylor-type specifications. In the third column in table 3-2, we show the respective coefficients of the additional Mundlak variables $\pi_{\bar{x}_k}$; all four of them are significantly different from zero.

[75] The agricultural production regions are defined by the Bavarian Agricultural Research Institute (LfL). We arbitrarily choose the Franconian region as reference group.

Table 3-2: Regression results of the estimated GLS stochastic frontier models

	Basic		Environmental Var.		Mundlak		Mundlak-Maddala		Hausman-Taylor	
	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
β_0 (Constant)	0.0046	(0.0072)	-0.0724	(0.0290) [b]	0.0134	(0.0067) [b]	0.0043	(0.0281)	-0.0510	(0.0285) [c]
β_{x1} (Labor)	0.0832	(0.0151) [a]	0.0772	(0.0152) [a]	0.0344	(0.0172) [b]	0.0422	(0.0172) [b]	0.0534	(0.0168) [a]
β_{x2} (Land)	0.2539	(0.0151) [a]	0.3132	(0.0178) [a]	0.3765	(0.0218) [a]	0.4456	(0.0270) [a]	0.3110	(0.0294) [a]
β_{x3} (Material)	0.5106	(0.0139) [a]	0.4980	(0.0140) [a]	0.3913	(0.0171) [a]	0.3906	(0.0171) [a]	0.3752	(0.0167) [a]
β_{x4} (Capital)	0.0687	(0.0091) [a]	0.0675	(0.0092) [a]	0.0150	(0.0109)	0.0168	(0.0109)	0.0574	(0.0113) [a]
$\pi_{\bar{x}1}$					0.1827	(0.0347) [a]	0.1478	(0.0358) [a]	0.1412	(0.0356) [a]
$\pi_{\bar{x}2}$					-0.4233	(0.0312) [a]	-0.4376	(0.0372) [a]	-0.3090	(0.0388) [a]
$\pi_{\bar{x}3}$					0.3066	(0.0291) [a]	0.2798	(0.0295) [a]	0.2937	(0.0292) [a]
$\pi_{\bar{x}4}$					0.1502	(0.0196) [a]	0.1528	(0.0201) [a]	0.1149	(0.0203) [a]
β_{z1} (owned land)			0.0013	(0.0004) [a]			0.0018	(0.0004) [a]	0.0012	(0.0004) [a]
$\pi_{\bar{z}1}$							-0.0012	(0.0005) [b]	-0.0005	(0.0005)
γ_1 (Region 1)			0.1297	(0.0352) [a]			0.0531	(0.0331)	0.0550	(0.0331) [c]
γ_2 (Region 2)			0.1450	(0.0467) [a]			0.0309	(0.0440)	0.0302	(0.0440)
γ_3 (Region 3)			0.0621	(0.0283) [b]			-0.0079	(0.0271)	-0.0092	(0.0270)
γ_4 (Region 4)			0.0034	(0.0274)			-0.0327	(0.0258)	-0.0331	(0.0258)
γ_5 (Region 6)			-0.1282	(0.0385) [a]			-0.0972	(0.0360) [a]	-0.0932	(0.0359) [a]
γ_6 (Part time)			-0.0673	(0.0291) [b]			-0.0389	(0.0289)	-0.0357	(0.0289)
β_t (Trend)									0.0154	(0.0015) [a]
σ_δ	0.1417	-	0.1368	-	0.1300	-	0.1261	-	0.1261	-
σ_v	0.1010	-	0.0102	-	0.1014	-	0.1010	-	0.0985	-

[a,b,c] parameter significant different from zero on the 1%, 5% and 10% level respectively

Under the assumption that all (still) unaccounted-for time-invariant differences between the firms are due to inefficiency, we obtain the respective efficiency scores from the group means of the residuals by applying the normalization described in Schmidt and Sickles (1984). The estimated standard deviation σ_δ is therefore equivalent to the standard deviation of the inefficiency term. If we compare σ_δ in column 1 and 3 in table 3-2, we can observe a decrease in the estimated standard deviation of the inefficiency term from 0.1417 to 0.130. This is to be expected, as in the M-GLS model, $\alpha_i = \pi\bar{x}_{i\cdot}$ is intended to capture all unobserved time-invariant heterogeneity that is correlated with the input variables. This helps to reduce heterogeneity bias in the inefficiency estimates. Hence, as in the Schmidt and Sickles (1984) initial fixed effects SFA model, the M-GLS specification provides in principle unbiased slope coefficients and therefore a better understanding of the underlying production technology without capturing all time-invariant differences between the firms in the fixed-effects inefficiency term.

If available, the introduction of environmental variables provides another method of accommodating heterogeneous production conditions in a production frontier. This approach corresponds to the E-specification (column 2), as in Good et al. (1993). We find that the technology-related coefficients of the input variables in this specification lie in between those of the B-GLS and the M-GLS model; six out of seven environmental variables are significantly different from zero. Comparing the standard deviation of the inefficiency term σ_δ in column 1 and 2 we find a reduction from 0.1417 to 0.1368. These findings confirm that the introduction of the environmental variables helps to reduce the heterogeneity bias in the technology parameters and in the inefficiency estimates.

In the MM-GLS specification, we combined the Mundlak approach with the additional set of environmental variables in the estimation equation, further improving our modeling of heterogeneity. In this specification, the term $\alpha_i = \pi_{\bar{x}}\bar{x}_{i\cdot} + \pi_{\bar{z}}\bar{z}_{i\cdot} + \gamma z_i$ takes into account not only the correlation of unobserved heterogeneity with input variables but also of observed (or rather approximated) heterogeneity, as measured by time-varying and time-invariant environmental variables. Accordingly, the estimated standard deviation of the time-invariant error component σ_δ is further reduced to 0.1261. In our example of an HT-GLS model, where we simply added a linear trend variable, we find no effect on σ_δ but, instead, an effect on σ_v – the estimated standard deviation of the idiosyncratic error component. This result must be the case because the trend variable has no in-between group variance. On the other side, we observe almost no effect on σ_v between the basic

GLS model and the M- and MM-GLS specification. This was also expected because only additional time-varying variable is the share of owned land[76].

The estimated parameters from the MLE random effects models (table 3-3) are similar to those obtained from the GLS models. The first order parameters of the input variables in the basic specifications have similar values, as well as the parameters of the Mundlak variables and the environmental variables. We also find the expected reactions in the estimated standard deviations of the inefficiency component σ_u, which is reduced from 0.2709 in the basic specification to 0.2101 in the Mundlak-Maddala specification. Interestingly, σ_u increases again to 0.2389 in the HT-MLE model. We assign this increase to the change in the sign of the parameter η. While this parameter is positive in the B-, E-, M- and MM-specification, indicating an increase in efficiency over time, it turns out to be negative in the *HT*-specification due to the incorporation of the trend variable within the production function part. This leads to a growing dispersion in inefficiency values over time and, therefore, to the higher value of σ_u.

In contrast to the GLS and the MLE models, the TRE model (table 3-4) attempts to capture time-invariant heterogeneity in its basic specification. The TRE model is a random parameter model, with the constant as the only random parameter. This random parameter is specified to be normally distributed and can be interpreted as an individual random effect. The additional parameter σ_{β_0} is the standard deviation of this random parameter and herein a measure of the unobserved (or unaccounted) variation between the farms. By including the Mundlak variables into the TRE model, we partially account for unobserved heterogeneity, as in the M-GLS and the M-MLE models. As expected, the "unaccounted" variation between the firms is reduced from σ_{β_0}=0.1789 to 0.1378. Similarily, the inclusion of environmental variables also reduces the standard deviation of the random parameter, although not severly (from σ_{β_0}=0.1789 to 0.1700). Still, in the TRE model, the additional Mundlak variables and the environmental variables have no distinct effect on the inefficiency estimates. Comparing columns 1-5 in table 3-4, we find almost identical estimates for σ_u and σ_v. These findings are in line with our expectations. Because all time-invariant differences between the firms are captured by the random constant anyway, the inefficiency estimates do not contain any time-invariant component and are not contaminated with heterogeneity. Thus, we simply reduce the variation in the random constant.

[76] Despite being a time-varying variable, the between-group variance in the *share of owned land* accounts for most (93.2 %) of the variable's overall variance.

Table 3-3: Regression results of the estimated MLE stochastic frontier models

	Basic		Environmental Var.		Mundlak		Mundlak-Maddala		Hausman-Taylor	
	Coeff.	Std.Err.			Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
β_0 (Constant)	0.2545	(0.0059) [a]	0.1495	(0.0223) [a]	0.2137	(0.0055) [a]	0.1693	(0.0289) [a]	0.0870	(0.0303) [a]
β_{x1} (Labor)	0.0827	(0.0128) [a]	0.0746	(0.0132) [a]	0.0345	(0.0142) [b]	0.0403	(0.0145) [a]	0.0560	(0.0142) [a]
β_{x2} (Land)	0.2211	(0.0103) [a]	0.2754	(0.0132) [a]	0.3100	(0.0135) [a]	0.3718	(0.0180) [a]	0.3096	(0.0176) [a]
β_{x3} (Material)	0.4997	(0.0091) [a]	0.4988	(0.0092) [a]	0.3844	(0.0096) [a]	0.3838	(0.0095) [a]	0.3741	(0.0094) [a]
β_{x4} (Capital)	0.0970	(0.0079) [a]	0.0983	(0.0087) [a]	0.0435	(0.0104) [a]	0.0438	(0.0104) [a]	0.0559	(0.0102) [a]
$\pi_{\bar{x}1}$					0.1809	(0.0252) [a]	0.1277	(0.0283) [a]	0.1155	(0.0290) [a]
$\pi_{\bar{x}2}$					-0.3681	(0.0214) [a]	-0.3550	(0.0309) [a]	-0.2757	(0.0312) [a]
$\pi_{\bar{x}3}$					0.2836	(0.0206) [a]	0.2543	(0.0227) [a]	0.2453	(0.0234) [a]
$\pi_{\bar{x}4}$					0.1370	(0.0146) [a]	0.1365	(0.0173) [a]	0.1177	(0.0172) [a]
β_{z1} (owned land)			0.0013	(0.0003) [a]			0.0015	(0.0003) [a]	0.0012	(0.0003) [a]
$\pi_{\bar{z}1}$							-0.0007	(0.0005)	-0.0003	(0.0005)
γ_1 (Region 1)			0.1301	(0.0305) [a]			0.0787	(0.0281) [a]	0.0906	(0.0303) [a]
γ_2 (Region 2)			0.0951	(0.0583)			0.0112	(0.0487)	0.0289	(0.0527)
γ_3 (Region 3)			0.0730	(0.0220) [a]			0.0234	(0.0246)	0.0267	(0.0264)
γ_4 (Region 4)			0.0068	(0.0198)			-0.0111	(0.0236)	-0.0049	(0.0252)
γ_5 (Region 6)			-0.0824	(0.0278) [a]			-0.0943	(0.0307) [a]	-0.0730	(0.0369) [b]
γ_6 (Part time)			-0.0537	(0.0211) [b]			-0.0403	(0.0205) [b]	-0.0332	(0.0220)
β_t (Trend)									0.0184	(0.0020) [a]
λ	2.6086	(0.0167) [a]	2.4105	(0.0182) [a]	2.1631	(0.0188) [a]	2.0989	(0.0194) [a]	2.4162	(0.0179) [a]
σ_u	0.2709	(0.0011) [a]	0.2483	(0.0009) [a]	0.2184	(0.0006) [a]	0.2101	(0.0005) [a]	0.2389	(0.0008) [a]
σ_v	0.1038	-	0.1030	-	0.1010	-	0.1001	-	0.0989	-
η	0.0436	(0.0040) [a]	0.0458	(0.0045) [a]	0.0483	(0.0045) [a]	0.0480	(0.0049) [a]	-0.0166	(0.0086) [c]

[a,b,c] parameter significant different from zero on the 1%, 5% and 10% level respectively

Table 3-4: Regression results of the estimated TRE stochastic frontier models

	Basic		Environmental Var.		Mundlak		Mundlak-Maddala		Hausman-Taylor	
	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
β_0 (Constant)	0.0865	(0.0043) [a]	0.0090	(0.0101)	0.0986	(0.0046) [a]	0.0763	(0.0103) [a]	0.0383	(0.0104) [a]
σ_{β_0}	0.1789	(0.0024) [a]	0.1700	(0.0024) [a]	0.1378	(0.0019) [a]	0.1315	(0.0020) [a]	0.1303	(0.0020) [a]
β_{x1} (Labor)	0.0727	(0.0080) [a]	0.0635	(0.0084) [a]	0.0317	(0.0134) [b]	0.0362	(0.0134) [a]	0.0414	(0.0137) [a]
β_{x2} (Land)	0.2733	(0.0066) [a]	0.3265	(0.0077) [a]	0.3810	(0.0125) [a]	0.4243	(0.0129) [a]	0.3034	(0.0174) [a]
β_{x3} (Material)	0.5015	(0.0064) [a]	0.4958	(0.0066) [a]	0.3887	(0.0106) [a]	0.3892	(0.0105) [a]	0.3713	(0.0103) [a]
β_{x4} (Capital)	0.0658	(0.0048) [a]	0.0658	(0.0050) [a]	0.0167	(0.0093) [c]	0.0173	(0.0093) [c]	0.0601	(0.0095) [a]
$\pi_{\bar{x}1}$					0.2012	(0.0162) [a]	0.1521	(0.0167) [a]	0.1567	(0.0164) [a]
$\pi_{\bar{x}2}$					-0.4324	(0.0143) [a]	-0.4149	(0.0146) [a]	-0.3099	(0.0192) [a]
$\pi_{\bar{x}3}$					0.3235	(0.0130) [a]	0.2918	(0.0131) [a]	0.3050	(0.0127) [a]
$\pi_{\bar{x}4}$					0.1437	(0.0106) [a]	0.1578	(0.0109) [a]	0.1178	(0.0109) [a]
β_{z1} (owned land)			0.0013	(0.0003) [a]			0.0011	(0.0001) [a]	0.0011	(0.0003) [a]
$\pi_{\bar{z}1}$							-0.0004	(0.0003)	-0.0004	(0.0003)
γ_1 (Region 1)			0.1124	(0.0112) [a]			0.0447	(0.0111) [a]	0.0521	(0.0107) [a]
γ_2 (Region 2)			0.1467	(0.0153) [a]			0.0229	(0.0150)	0.0249	(0.0144) [c]
γ_3 (Region 3)			0.0418	(0.0090) [a]			-0.0242	(0.0090) [a]	-0.0221	(0.0087) [b]
γ_4 (Region 4)			-0.0104	(0.0087)			-0.0395	(0.0086) [a]	-0.0402	(0.0083) [a]
γ_5 (Region 6)			-0.1348	(0.0121) [a]			-0.0947	(0.0119) [a]	-0.0857	(0.0116) [a]
γ_6 (Part time)			-0.0394	(0.0093) [a]			-0.0722	(0.0083) [a]	-0.0686	(0.0084) [a]
β_t (Trend)									0.0163	(0.0011) [a]
λ	1.4723	(0.0993) [a]	1.4747	(0.1001) [a]	1.4853	(0.1114) [a]	1.4501	(0.1104) [a]	1.5996	(0.1145) [a]
σ_u	0.1154	-	0.1148	-	0.1127	-	0.1105	-	0.1134	-
σ_v	0.0784	-	0.0779	-	0.0759	-	0.0762	-	0.0709	-

[a,b,c] parameter significant different from zero on the 1%, 5% and 10% level respectively

Nevertheless, as in the M-GLS and the M-MLE models, the Mundlak variables in the M-TRE specification are able to control for possible correlations between the inputs and the unobserved heterogeneity, thereby mitigating the heterogeneity bias in the estimated coefficients of the input variables. We find that those coefficients in the TRE models to evolve in a similar manner as in the GLS and MLE models across the different specifications. Including additional environmental variables, as in the MM-TRE specification, leads to a further decrease in σ_{β_0}. Again, the inefficiency estimates are almost unaffected. In the HT-TRE specification, the coefficient of the trend variable is significant; however, the changes in the time-invariant and time-varying error components are only minor. The different specifications within the three models are nested. Hence, we are able to use Likelihood-Ratio and Wald tests to identify the most suitable specification for each model and the data at hand. The results of these tests are summarized in table 3-5.

Table 3-5: Model specification tests

Model		F- / LR-statistic	critical value (α =0.05)
GLS	B vs. M	87.53	$F_{(4,1935)} = 2.38$
	B vs. E	12.71	$F_{(7,1932)} = 2.01$
	M vs. MM	6.08	$F_{(8,1927)} = 1.94$
	E vs. MM	75.28	$F_{(5,1927)} = 2.22$
	MM vs. HT	104.55	$F_{(1,1926)} = 3.85$
MLE	B vs. M	276.39	$\chi^2_4 = 9.49$
	B vs. E	94.64	$\chi^2_7 = 14.07$
	M vs. MM	60.32	$\chi^2_8 = 15.51$
	E vs. MM	242.06	$\chi^2_5 = 11.07$
	MM vs. HT	52.59	$\chi^2_1 = 3.84$
TRE	B vs. M	319.82	$\chi^2_4 = 9.49$
	B vs. E	73.63	$\chi^2_7 = 14.07$
	M vs. MM	49.81	$\chi^2_8 = 15.51$
	E vs. MM	296.00	$\chi^2_5 = 11.07$
	MM vs. HT	119.32	$\chi^2_1 = 3.84$

Consistently across the three models, we find that all specifications that attempt to model heterogeneity are preferred over the basic specification. Additionally, in all three models, the HT-specification is preferred. However, the decision as to whether the GLS-, the MLE- or the TRE- model is the appropriate model for empirical application is not as straightforward. Instead this decision must be based on the researcher's knowledge about the sector under consideration, the characteristics of the dataset and the assumptions one is willing to make.

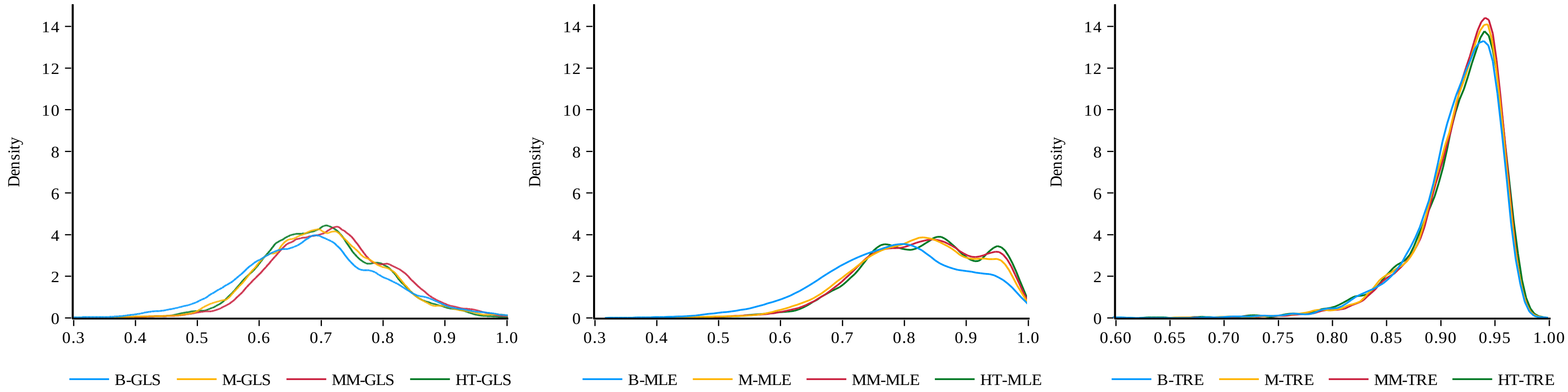
We present and compare the efficiency scores derived from the different models and specifications[77]. In addition, we calculate the predicted values of the individual effects α_i and examine how the discussed model extensions account for heterogeneity. In figure 3-1 we present the distributions and the descriptive statistics of the estimated technical efficiency scores from the GLS-, the MLE and the TRE models.

Figure 3-1 confirms what we already expected from the estimated variance components σ_δ and σ_v in the different GLS-model specifications. The additional Mundlak- and environmental variables account for heterogeneity and therefore help to reduce the respective contamination of the efficiency scores. In our example, the mean efficiency scores vary between 0.687 for the basic GLS model and 0.719 for the MM-GLS model. As we can see from figure 3-1 the same applies to the MLE-models. Both, the Mundlak- and the Mundlak-Maddala-MLE specification shift the distribution of the efficiency scores towards one and raise the mean efficiency from 0.786 in the Basic-MLE model to 0.821 and 0.828, respectively.

The efficiency results for the TRE-models also fit into this set of results. As expected, adding time-invariant Mundlak variables and mostly time-invariant environmental variables to the estimation equation has no substantial effect on the efficiency scores, as obtained from the TRE models. This finding might be different in empirical applications where more information about time-varying environmental production conditions is available. To examine how the efficiency scores obtained from the different models and specifications relate to one another we calculate their Pearson correlation coefficients. (table 3-6) For the models with time-varying efficiency, we use their group mean efficiency score over the observed period to calculate the correlation coefficients. The coefficients vary between 0.443 and 0.998. We find a consistently high correlation between all GLS and MLE specifications, in particular for the M-, MM- and HT-specifications that take heterogeneity into account. The correlation between the efficiency scores of those models and the efficiency scores from the TRE models is notably lower.

[77] From here on we omit the results of the Environmental-Effects specification for reasons of clarity and comprehensibility. The respective specifications perform as expected; the additional results are available upon request.

Figure 3-1: Distributions and descriptive statistics on technical efficiency scores from GLS-, MLE- and TRE models



	Mean	S.D.	Max.	Min.		Mean	S.D.	Max.	Min.		Mean	S.D.	Max.	Min.
B-GLS	0.687	0.112	1.0	0.332	B-MLE	0.786	0.113	0.990	0.368	B-TRE	0.915	0.039	0.987	0.602
M-GLS	0.704	0.095	1.0	0.399	M-MLE	0.821	0.098	0.990	0.429	M-TRE	0.917	0.039	0.986	0.662
MM-GLS	0.719	0.095	1.0	0.420	MM-MLE	0.828	0.096	0.986	0.444	MM-TRE	0.918	0.038	0.986	0.667
HT-GLS	0.701	0.092	1.0	0.408	HT-MLE	0.831	0.096	0.987	0.471	HT-TRE	0.916	0.041	0.987	0.637

Table 3-6: Pairwise Pearson correlation coefficients of technical efficiency scores

	B-GLS	M-GLS	MM-GLS	HT-GLS	B-MLE	M-MLE	MM-MLE	HT-MLE	B-TRE	M-TRE	MM-TRE	HT-TRE
B-GLS	1.0											
M-GLS	0.809	1.0										
MM-GLS	0.785	0.969	1.0									
HT-GLS	0.783	0.967	0.998	1.0								
B-MLE	0.985	0.823	0.801	0.799	1.0							
M-MLE	0.799	0.977	0.950	0.948	0.838	1.0						
MM-MLE	0.780	0.935	0.972	0.970	0.820	0.960	1.0					
HT-MLE	0.811	0.928	0.964	0.966	0.850	0.955	0.993	1.0				
B-TRE	0.519	0.462	0.449	0.443	0.539	0.483	0.475	0.473	1.0			
M-TRE	0.507	0.620	0.598	0.592	0.534	0.632	0.609	0.590	0.821	1.0		
MM-TRE	0.509	0.621	0.629	0.623	0.536	0.634	0.641	0.622	0.821	0.983	1.0	
HT-TRE	0.526	0.621	0.620	0.617	0.553	0.635	0.631	0.625	0.802	0.957	0.968	1.0

We examine our modeling of unobserved and observed heterogeneity and also present the predicted values for the individual effects, α_i. In table 3-7, we summarize how the predicted firm effects $\hat{\alpha}_i$ are calculated according to the respective model specification. As previously discussed, the basic GLS and MLE specifications do not contain any modeling of heterogeneity. In contrast, we also obtain an estimate of α_i from the basic TRE model.

Table 3-7: Specifications of calculated firm effects

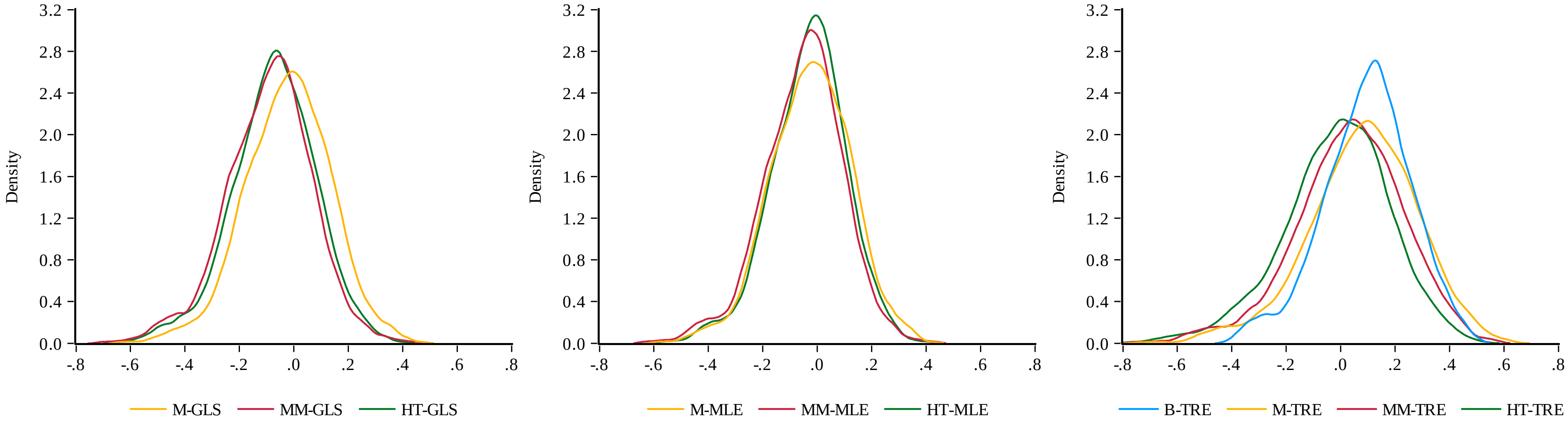
	GLS	MLE	TRE
Basic	-	-	$\hat{\alpha}_i = \hat{\delta}_i$
Mundlak	$\hat{\alpha}_i = \hat{\pi}\bar{x}_{i\cdot}$	$\hat{\alpha}_i = \hat{\pi}\bar{x}_{i\cdot}$	$\hat{\alpha}_i = \hat{\delta}_i + \hat{\pi}\bar{x}_{i\cdot}$
Mundlak-Madalla	$\hat{\alpha}_i = \hat{\pi}\bar{x}^*_{i\cdot} + \hat{\gamma}z_i$	$\hat{\alpha}_i = \hat{\pi}\bar{x}^*_{i\cdot} + \hat{\gamma}z_i$	$\hat{\alpha}_i = \hat{\delta}_i + \hat{\pi}\bar{x}^*_{i\cdot} + \hat{\gamma}z_i$
Hausman-Taylor	$\hat{\alpha}_i = \hat{\pi}_1\bar{x}^*_{1i\cdot} + \hat{\gamma}z_i$	$\hat{\alpha}_i = \hat{\pi}_1\bar{x}^*_{1i\cdot} + \hat{\gamma}z_i$	$\hat{\alpha}_i = \hat{\delta}_i + \hat{\pi}_1\bar{x}^*_{1i\cdot} + \hat{\gamma}z_i$

According to table 3-7, $\hat{\alpha}_i$ is calculated as the fitted value of our models of heterogeneity, which is obtained using the estimated coefficients – indicated by a hat ($\hat{\cdot}$) – and the respective data. However, in the TRE model, $\hat{\delta}_i$ is not a fitted value, but rather, the estimated mean of the conditional distributions of the random parameter, δ_i[78]. It is encouraging that the distributions and the descriptive statistics of $\hat{\alpha}_i$ obtained from the different models are quite similar. Figure 3-2 shows the distribution and descriptive statistics of $\hat{\alpha}_i$ – obtained from the GLS-, MLE- and TRE models. It should be noted that our lack of an estimate from the basic GLS- and MLE specifications is due to the model's non-accounting for heterogeneity.

Because similar distributions and statistics do not necessarily reveal a present relationship, we also calculate the correlation between these predictions of the firm effects; the respective coefficients are presented in table 3-8. We find that the described Mundlak-, Mundlak-Maddala and Hausman-Taylor specifications of the GLS and the MLE models are able to integrate and measure heterogeneity in a similar way to that of the TRE model. The relationship between the measures of heterogeneity becomes even closer as we compare the GLS- and MLE-models and those from the augmented TRE-models.

[78] Details on the estimation of the individual specific coefficient can be found in Greene (2007) and Train (2009)

Figure 3-2: Distributions and descriptive statistics of calculated firm effects from GLS-, MLE- and TRE models



	Mean	S.D.	Max.	Min.		Mean	S.D.	Max.	Min.		Mean	S.D.	Max.	Min.
										B-TRE	0.097	0.159	0.459	-0.370
M-GLS	-0.022	0.155	0.420	-0.584	M-MLE	-0.024	0.145	0.379	-0.544	M-TRE	0.082	0.197	0.576	-0.677
MM-GLS	-0.091	0.156	0.399	-0.657	MM-MLE	-0.051	0.144	0.392	-0.594	MM-TRE	0.031	0.196	0.509	-0.702
HT-GLS	-0.073	0.152	0.355	-0.608	HT-MLE	-0.031	0.136	0.374	-0.529	HT-TRE	-0.016	0.193	0.471	-0.718

Table 3-8: Pairwise Pearson correlation coefficients of calculated firm effects

	M-GLS	MM-GLS	HT-GLS	M-MLE	MM-MLE	HT-MLE	B-TRE	M-TRE	MM-TRE	HT-TRE
M-GLS	1.0									
MM-GLS	0.969	1.0								
HT-GLS	0.944	0.946	1.0							
M-MLE	0.995	0.957	0.962	1.0						
MM-MLE	0.968	0.990	0.969	0.965	1.0					
HT-MLE	0.953	0.959	0.993	0.965	0.985	1.0				
B-TRE	0.588	0.604	0.558	0.575	0.602	0.578	1.0			
M-TRE	0.786	0.782	0.764	0.782	0.787	0.777	0.955	1.0		
MM-TRE	0.781	0.795	0.760	0.773	0.792	0.772	0.955	0.995	1.0	
HT-TRE	0.751	0.746	0.790	0.65	0.766	0.786	0.917	0.975	0.969	1.0

3.5. Concluding Remarks

In this paper, we examine several extensions for stochastic frontier models, attempting to take heterogeneity, be it observable or unobserved, into account and to reduce heterogeneity-induced biases in the estimated efficiency scores and technology parameters. Therefore, we further develop the line of thought brought forth Farsi et al. (2005) in two directions. Following Mundlak (1978), these authors suggest the incorporation of the group-means of input variables in Greene's (2005) "true" random effects stochastic frontier model to mitigate the heterogeneity bias in the technology parameters.

We broaden this approach and consider the incorporation of the group mean variables, not as a sole auxiliary equation but as an actual modeling of between-firm heterogeneity. Extensions of this heterogeneity model include the incorporation of environmental variables and the notion that group-mean variables are only needed for inputs that are assumed to be correlated with heterogeneity. The second direction is that we adapt this modeling of heterogeneity, not only for TRE-type models but also for the MLE and GLS random effects models with time varying and time-invariant inefficiency. Thus, we extend the range of models and broaden the toolbox for the empirical researcher in a straightforward, useful way.

The results of our empirical application confirm the proposed specifications. The heterogeneity biases in the technology parameters are reduced in the sense that the coefficients of the input factors come closer to those of the conventional fixed effects model. The predicted firm effects and the efficiency scores we obtain from the different specifications also meet our expectations. Compared to the basic specifications of the GLS and MLE model, the proposed extensions help to reduce the downward bias in efficiency scores. We do not find this effect on the efficiency scores in the TRE specifications, in which the random constant already captures the firm effects. These results are in line with the findings by Farsi et al. (2005, 2005a) and Abdulai and Tietje (2007). For the predicted firm effects, we find similar distributions across different specifications, and fairly strong positive correlations. These findings indicate that the proposed specifications to model heterogeneity can serve as a possible alternative to the "true" effects models.

However, the final decision as to which model is appropriate for an empirical application must be based on the researcher's knowledge of the analyzed sector, the dataset at hand and the assumptions one is willing to make. The first consideration would be whether the firms actually work under heterogeneous production conditions that should be controlled. In cases where the dataset at hand contains information on such conditions, the respective (proxy-) variables can be included in the production function, as in Good et al. (1993) and Coelli et al. (1999), for example. If there is reason to assume that the firms in the sample have adjusted their level of input usage according to their production conditions, the coefficients for the input variables will be affected by heterogeneity bias. However, we can make use of this fact and include the Mundlak control variables into the estimation equation as a model of unobserved (to the researcher) heterogeneity. The second consideration must consider whether the lack of any time-persistent inefficiency can be assumed. Supporting factors for this assumption could be i: a competitive operating environment that forces inefficient firms either to improve their efficiency rapidly or to drop out of the market; ii: changes in the operating and management conditions during the observed time period (e.g., policy and regulation) that require the firms to adapt; and iii: a sufficient number of observed time periods. If a researcher cannot be sure that these conditions are met, e.g., if a dataset contains only a couple of time periods, it is possible that some firms exhibit non-varying efficiency levels. In that case, an individual effect, as in the TRE model, would consequently capture parts of this inefficiency.

4. DECOMPOSING LABOR PRODUCTIVITY GROWTH – THE CASE OF SMALL AND MEDIUM-SIZED BREWERIES IN GERMANY

Abstract

In this paper, we provide a method to decompose aggregate industry labor productivity growth into seven distinct components: input deepening, technical change, technical efficiency, scale effect, between-firm reallocation and effects from exits and entry. The first four components measure the productivity growth that takes place within a firm. The latter three components capture industry dynamics. Applied to a sample of 118 small and medium sized breweries in Germany between 1996 and 2008, we found that within-firm effects and in particular technical change and the scale change effect clearly dominated the effects from industry restructuring.

Keywords: labor productivity decomposition, structural change

JEL Classification: D24, J24, L16, O12

This chapter is based on the article *Decomposing labor productivity growth – the case of small and medium-sized breweries in Germany* by Magnus Kellermann, Giannis Karagiannis, Klaus Salhofer and Stefan Kilian. The article is currently under revision at the *International Journal of Production Economics*. The author of this dissertation is the main author of the article.

4.1. Introduction

Today, the brewing industry worldwide is highly concentrated. The five largest firms (AB-InBev, SAB Miller, Heineken, Carlsberg and China Resource Brewery) account for two-thirds of the global profits in this sector and produce approximately half of the beers worldwide (NGG, 2013). As a global exception, the brewing industry in Germany is still dominated by relatively small firms. Only two of the five worldwide market leaders (AB-InBev as number two and Carlsberg as number nine) are listed among Germany's top ten breweries, and these firms account for approximately 15% of the German beer production (NGG, 2013). Moreover, the largest German brewery (Radeberger Gruppe KG) is in only the 23rd position worldwide (NGG, 2013). Nevertheless, the German brewing industry has faced considerable structural changes in the last two decades. Beer consumption was relatively stable in the 1970s and 1980s at approximately 146 liters per capita and was still at 141.9 liters per capita in 1991, but it constantly decreased to 106.4 liter per capita (-25%) in 2013 (Private Brauereien Bayern e.V., 2014). In addition to a decrease in quantity, there was also a considerable change in consumer preferences away from consuming Pils and Lager on a more frequent basis in pubs to occasional consumption of specialty beers at home. Although the net-exports increased by approximately 2.7 million hectoliters (hl) between 1998 and 2012, this was not enough to compensate for the decrease in domestic consumption (NGG, 2013). As a consequence, beer production decreased by 16.3% from 115.3 million hl in 1995 to 96.5 million hl in 2012 (Private Brauereien Bayern e.V., 2014). During the same time, the number of brewery employees decreased by almost half from 48,216 in 1995 to 26,915 in 2012 (NGG, 2009; 2013).

Despite the decreasing demand for beer, the German brewing sector is still characterized by a comparably low concentration. In fact, Germany is the country with the most breweries of the EU nations (Berkhout et al., 2013). Interestingly, in the last decade, the number of breweries has increased from 1,275 in 2002 to 1,349 in 2013 (Private Brauereien Bayern e.V., 2014). However, these aggregated numbers give a very incomplete picture of the developments. New establishments entered the market only in the group of very small breweries (producing up to 1,000 hl/year). Their number increased from 523 in 2006 to 668 in 2013 (Private Brauereien Bayern e.V., 2012; 2014). In all other groups, we observe a sharp decrease, e.g., the number of breweries with over 5,000 hl/year decreased by 33.5% between 1995 and 2013 (Deutscher Brauerbund e.V., 2009;

Private Brauereien Bayern e.V., 2014). These breweries are in a fierce competition for a decreasing demand.

The aim of this paper is to investigate the development of the labor productivity in this sector as a key factor for firms to increase their competitiveness. To do so, we combine two strands of the literature on (labor) productivity decomposition. One strand originates from empirical studies that use micro-data to describe the productivity growth dynamics of a sector. Several decomposition methods have been proposed to analyze the sources of aggregate productivity change via a within-firm effect and the reallocation effects between incumbent firms as well as entering and exiting firms (Baily et al., 1992; 1996; 2001; Griliches and Regev, 1995; Foster et al., 2001; Melitz and Polanec, 2012). The other strand combines index number theory with stochastic frontier analysis (Nishimizu and Page, 1982; Bauer, 1990) and decomposes firm-specific productivity growth into several components. Here, we show how to combine those two approaches to analyze the dynamics of aggregated industry labor productivity in great detail. In particular, based on this procedure, we can decompose industry labor productivity change into seven components: input deepening, technical change, technical efficiency, a scale change effect, between-firm reallocation and the effects from exits and entry. The first four of these components constitute the within-firm effect. Applying our method to a sample of 118 German breweries between 1996 and 2008 provides useful insights into the development of (labor) productivity and its driving forces.

4.2. Method

The labor productivity (LP) of a single firm i in time t in its logarithmic form is defined as $LP_{it} = \ln\left(\frac{y_{it}}{l_{it}}\right)$, where y_{it} is the quantity of output produced and l_{it} is the utilized amount of labor. Moreover, we define labor productivity of the whole industry I consisting of N firms (or a sample of firms N within an industry) at time t as the share-weighted average labor productivity $LP_t^I = \sum_{i=1}^{N} s_{it} LP_{it}$, where s_{it} represents a firm's activity share within the industry (Olley and Pakes, 1996). The change in labor productivity of a single firm and of the whole industry from period $t-1$ to t is given by $\Delta LP_{it} = LP_{it} - LP_{it-1}$ and $\Delta LP_t^I = LP_t^I - LP_{t-1}^I$, respectively.[79]

[79] Because labor productivity is in logarithms, ΔLP_{it} (ΔLP_t^I) is the percentage change or a discrete rate of change in a firm's (industry's) labor productivity.

To decompose the change in industry labor productivity (ΔLP_t^I) into its components we proceed as follows. In a first step, we differentiate between effects within firms, effects between firms and effects from firms that enter and exit the sample and/or the industry. In a second step, we further decompose the within-firms component into the effect of technical change, the scale effect, the change in technical efficiency effect and an input deepening effect. Our decomposition of the change in industry labor productivity (ΔLP_t^I) in the first step is closely related to the one proposed by Griliches and Regev (1995). Given the nature of our data, we further decompose the net entry term to distinguish the effect associated with firms that enter/exit the industry (and therefore the sample) from that of firms that drop in and out of the sample for other unknown reasons[80]. Therefore, in each period, our sample is divided into continuing firms (C), new firms that enter the industry (N_E), existing firms that enter the sample (N_S), firms that shut down (or are sold) (X_E) and firms that exit the sample for other reasons but continue to produce (X_S). Given this, the industry's labor productivity can be decomposed into

$$\begin{aligned}\Delta LP_t^I = &\sum_{i\in C} \widetilde{s_{it}}\,\Delta LP_{it} + \sum_{i\in C}\left(\widetilde{LP_{it}} - \widetilde{LP_t^I}\right)\Delta s_{it} + \sum_{i\in N_E} s_{it}\left(LP_{it} - \widetilde{LP_t^I}\right)\\ &+ \sum_{i\in N_S} s_{it}\left(LP_{it} - \widetilde{LP_t^I}\right) - \sum_{i\in X_E} s_{it-1}\left(LP_{it-1} - \widetilde{LP_t^I}\right) \qquad \forall\, t \neq 1 \qquad (4.1)\\ &- \sum_{i\in X_S} s_{it-1}\left(LP_{it-1} - \widetilde{LP_t^I}\right)\end{aligned}$$

where a tilde over a variable denotes the arithmetic mean of the variable in t and $t-1$ (i.e., $\widetilde{s_{it}} = \frac{1}{2}(s_{it} + s_{it-1})$), and a delta in front of a variable denotes its first-difference (i.e., $\Delta s_{it} = s_{it} - s_{it-1}$). The first term on the right hand side is the aggregated effect of the individual firms' weighted labor productivity change (within-firm component). Loosely, this is positive if firms improve their performance on average. The second term shows the effect of shifts in the shares between firms (between-firm component) weighted by the firm's deviation in its average productivity in t and $t-1$ from the industry's respective productivity. This is positive if the relative weight of high-productivity to low-productivity firms increases. The third (fifth) term is the effect of firms that enter (exit) the industry and therefore the sample. The effects on labor productivity of the whole industry are positive if better (worse) than average performing firms enter (exit) and negative

[80] Baily et al. (1992) noted the possibility that a firm that exits their sample may still operate. However, they do not pursue this issue but note that they do not regard this as a problem of magnitude for their sample (Baily et al. 1992, fn. 11). The firms in our dataset participate on a voluntary basis; hence, we must not neglect the issue of exits/entries to the sample that do not reflect the true behavior of the firm.

otherwise. Finally, the fourth (sixth) term gives the aggregated effects of firms that exit or enter the sample but not the industry. The same reasoning applies for the direction of the effects on industry productivity.

Various methods have been used in the literature to decompose the change in labor productivity into these components. Baily et al. (1992) were the first to differentiate between a within-firm and a between-firm component and also distinguished between surviving, entering and exiting firms. The main difference between their method and the method we use based on Griliches and Regev (1995) is that the latter introduces the average aggregate industry productivity level between the two periods $\widetilde{LP}_t^I$ as a reference point (Melitz and Polanec, 2012). This has the interpretive advantage that the contribution of entering and exiting firms (terms three to six in equation (2.1)) on the industry productivity change can be positive or negative, but the contribution of entry (exit) is always positive (negative) in Baily et al. (1992). Another popular decomposition is Foster et al. (2001). Although it also adds an additional component, a cross-firm effect, the main difference compared to Griliches and Regev (1995) is that Foster et al. use the industries' initial productivity level LP_{t-1}^I rather than the time average $\widetilde{LP}_t^I$ as a reference point. Recently, Melitz and Polanec (2012) introduced another decomposition: a dynamic version of the well-known static Olley and Pakes (1996) decomposition. Hence, they decompose the change in industry labor productivity into a change in the unweighted mean in the productivity, the covariance change between market share and productivity, and the contributions of entrants and exiting firms. In contrast to Griliches and Regev (1995) and Foster et al. (2001), Melitz and Polanec use surviving firms at time t $(t-1)$ as a benchmark to value the contribution of entering (exiting) firms. Although it is clear that any choice of reference group will influence the contribution of entrants and exiting firms, it remains debatable which approach is superior. Balk (2003) and Diewert and Fox (2010) argue that the decomposition of Griliches and Regev (1995) (as compared to that of Foster et al. (2001)) has the advantage of treating time in a symmetric fashion, which makes the within term in this decomposition a Divisia index of the continuing firms' productivity change (Foster et al., 2008).

In a second step, we further decompose the within-firm component of productivity growth (first right-hand-side term in equation (4.1)) by using a parametric frontier approach following Nishimizu and Page (1982) and Bauer (1990). To do so, we describe a firm's production technology with a well-behaved production function but also account for the possibility of technical inefficiency:

$$y = f(\boldsymbol{x}, t) TE(t) \tag{4.2}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_J)$ is a vector of J inputs, y is a scalar output, $TE(t)$ is the output-oriented measure of technical efficiency defined over the range $(0,1]$ and t is a time trend that accounts for technological change in the production function. Hence, if $TE(t) = 1$, the production of a firm is at the technically efficient level, which is described by the production frontier $y = f(\boldsymbol{x}, t)$. Taking logarithms of both sides of equation (4.2) and totally differentiating them with respect to time results in:

$$\Delta \ln y = \sum_{j=1}^{J} \varepsilon_j \, \Delta \ln x_j + \Delta T + \Delta \ln TE\,. \tag{4.3}$$

The delta in front of a variable denotes its difference over adjacent time periods, i.e., its rate of growth (e.g., $\Delta \ln y = \frac{d}{dt} \ln(y)$), $\varepsilon_j = \partial \ln f\,(\boldsymbol{x}, t)/\partial \ln x_j$ is the partial output elasticity of the jth input, $\Delta T = \partial \ln f\,(\boldsymbol{x}, t)/\partial t$ is the primal rate of technical change and $\Delta \ln TE = \partial \ln TE\,(t)/\partial t$ is the rate of change in technical efficiency.

By subtracting the growth of labor input $\Delta l = \frac{d}{dt} \ln(l)$ from both sides of (4.3) and by adding and subtracting aggregate input growth $\sum_{j=1}^{J} \frac{\varepsilon_j}{\varepsilon} \Delta x_j$ [81], we rearrange equation (4.3) to

$$\begin{gathered} \Delta LP = \Delta \ln y - \Delta \ln l = \\ \textstyle\sum_{j=1}^{J-1} \left(\frac{\varepsilon_j}{\varepsilon}\right) \left(\Delta \ln x_j - \Delta \ln l\right) + \Delta T + (\varepsilon - 1) \sum_{j=1}^{J} \left(\frac{\varepsilon_j}{\varepsilon}\right) \Delta \ln x_j + \Delta \ln TE. \end{gathered} \tag{4.4}$$

The input *labor* is defined to be the J-th input, i.e., $x_J = l$. Equation (4.4) decomposes firm level labor productivity growth into four elements. The first term on the right-hand side is the input deepening effect, i.e., it accounts for changes in factor intensities. Input deepening relates to factor substitution and indicates that labor productivity can increase if the other inputs grow faster than labor and eventually replace it in the production process. Technical change (second term) has a one-to-one contribution to labor productivity growth and positively affects it if it is progressive. The contribution of the scale effect (third term) is positive if the production technology exhibits increasing returns to scale ($\varepsilon > 1$) and the aggregate input usage expands or if $\varepsilon < 1$ and the input usage is reduced. In the case of constant returns to scale ($\varepsilon = 1$) or constant input quantities, the scale effect becomes

[81] Aggregate input growth is denoted by $\sum_{j=1}^{J} c_j \, \Delta \ln x_j$, where $c_j = (w_j x_j)/C$ is the cost share of the jth input, w_j is the respective input price and C are the total costs. Under the assumption of allocative efficiency and competitive input markets $c_j = \frac{\varepsilon_j}{\varepsilon}$, where $\varepsilon = \sum_{j=1}^{J} \varepsilon_j$ is the scale elasticity (Chan and Mountain, 1983).

zero. Finally, technical efficiency change (fourth term) indicates a catching up effect (Färe et al. 1994) that contributes positively to labor productivity growth as firms move closer to the production frontier. The last three terms correspond to Bauer's (1990) and Lovell's (1996) decompositions of total factor productivity growth (ΔTFP). Hence, one may rewrite equation (4.4) as $\Delta LP = \sum_{j=1}^{J-1}\left(\frac{\varepsilon_j}{\varepsilon}\right)\left(\Delta \ln x_j - \Delta \ln l\right) + \Delta TFP$, where $\Delta TFP = \Delta \ln y - \sum_{j=1}^{J}\left(\frac{\varepsilon_j}{\varepsilon}\right)\Delta \ln x_j$. This highlights the advantage of the present decomposition of labor productivity growth in equation (4.4). It preserves the intuitive concept of a partial productivity measure but still features the differences between the substitution effects and productivity growth due to technical progress, efficiency change and scale changes.

Once the parameters of the production frontier $f(\boldsymbol{x},t)TE(t)$ are econometrically estimated, we can calculate all four components without knowledge of the input prices and the assumption of constant returns to scale (Bauer, 1990). Several models for the econometric estimation of production (or cost) frontiers from panel data have been proposed and discussed in the literature (e.g., Greene, 2008). A stochastic frontier panel model can be formulated as

$$\ln y_{it} = \ln \boldsymbol{x}'\boldsymbol{\beta} + \alpha_i + u_{it} + e_{it}. \tag{4.5}$$

where $\boldsymbol{\beta}$ are parameters to be estimated, α_i are time-invariant firm-specific effects, u_{it} is a non-negative term that represents inefficiency and e_{it} is statistical noise.

The main distinguishing features of the various proposed models are the way inefficiency (u_{it}) is modeled, whether inefficiency is allowed to vary over time (u_{it} versus u_i) and the way firm heterogeneity α_i is taken into account.[82] Greene (2005, 2005a) addressed the issue of between-firm heterogeneity and proposed the "true" fixed-effects and "true" random-effects model, where α_i is a constant or an iid normal distributed random term, respectively. The "true" effects models present a great improvement in dealing with potential between-firm heterogeneity in the stochastic frontier framework. Nevertheless, the models have some complexities, and their implementation requires involved econometric estimation procedures.[83]

---

[82] Pitt and Lee (1981) and Schmidt and Sickles (1984) made early contributions to panel data models and assume time-invariant technical efficiency. Battese and Coelli (1992) and Cornwell et al. (1990) extend the models of Pitt and Lee (1981) and Schmidt and Sickles (1984) to allow for time-varying inefficiency. These earlier models did not specifically account for firm heterogeneity (α_i) within the model. Hence, the contamination of the measure of inefficiency with unobserved firm-specific heterogeneity is an issue discussed in the more recent literature (e.g., Greene 2005, Farsi et al. 2005).

[83] See Greene (2005, 2005a, 2008) and Wang and Ho (2010) for further details.

Here, we follow a different strategy by explicitly modelling firm heterogeneity in the spirit of Mundlak (1978). In particular, we model the stochastic production frontier as

$$\ln y_{it} = \sum_{j=1}^{J} \beta_j \ln x_{jit} + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} \beta_{jk} \ln x_{jit} \ln x_{jit} + \beta_t t + \frac{1}{2}\beta_{tt} tt + \sum_{j=1}^{J} \beta_{tj}\, t \ln x_{jit} + \alpha_i - u_{it} + e_{it}$$

with

$$\alpha_i = \sum_{j=1}^{J} \gamma_j \left(\overline{\ln x_{jit}}\right) + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} \left(\overline{\ln x_{jit}\ \ln x_{kit}}\right) + \sum_{j=1}^{J} \gamma_j \left(\overline{t \ln x_{jit}}\right), \tag{4.6}$$

$$u_{it} = u_i \exp\left(-\eta(t-T)\right) \rightarrow u_i \sim N^{+}(0, \sigma_u^2),$$

$$e_{it} \sim iidN(0, \sigma_e^2)$$

where all the βs, γs and η are parameters to be estimated. A bar over a variable denotes its cross-section mean, i.e., $\bar{x}_{it} = \frac{1}{T_i}\sum_{t=1}^{T_i} x_{it}\ \forall\, i$. To avoid imposing unnecessary a priori restrictions on the production technology, we use the flexible translog form with symmetry imposed as $\beta_{jk} = \beta_{kj}\ \forall\ j,k$. The firm-specific effect α_i is explicitly modelled based on the following reasoning: unobserved heterogeneity may be unobservable to only the econometrician but not the decision making unit. Thus, we can expect that the firms have adjusted their inputs according to their given production conditions. Hence, unobserved heterogeneity is assumed to be correlated with the observed levels of input usage. If this assumption holds, we can model α_i by adding the individual group means of inputs as auxiliary variables. In this way, we can account for the unobserved heterogeneity that is correlated with the firm's level of input usage.

A second but no less relevant virtue of this approach is that it mitigates the heterogeneity bias in the slope parameters. Mundlak (1978) showed that including the group means of the explanatory variables in a GLS random effects model yields the unbiased within estimator for the slope parameters. This result cannot be strictly applied to stochastic frontier models with an asymmetric composed error term. However, we can expect the heterogeneity bias to be minimal to the extent that the auxiliary variables capture the correlation between the unobserved effect and input quantities (Farsi et al.

2005). To allow for temporal variation in the one-sided inefficiency component u, we use the time-varying formulation of Battese and Coelli (1992). Finally, we have a symmetric noise component e_{it}.

After the econometric estimation of (4.6) we can calculate the four components of firm-level labor productivity growth. To calculate the input deepening effect $\left(ID_{it} = \sum_{j=1}^{J-1} \left(\frac{\varepsilon_{jit}}{\varepsilon_{it}}\right) (\Delta \ln x_{jit} - \Delta \ln l_{it})\right)$ and the scale effect $\left(SC_{it} = (\varepsilon_{it} - 1) \sum_{j=1}^{J} \left(\frac{\varepsilon_{jit}}{\varepsilon_{it}}\right) \Delta \ln x_{jit}\right)$ of firm i in time t, we need percentage changes in inputs ($\Delta \ln x_{jit}$) and the scale elasticity calculated as

$$\varepsilon_{it} = \sum_{j=1}^{J} \varepsilon_{jit} = \sum_{j=1}^{J} \frac{\partial \ln f(\boldsymbol{x}, t)}{\partial \ln x_{jit}} = \sum_{j=1}^{J} \left(\hat{\beta}_j + \sum_{k=1}^{J} \hat{\beta}_{jk} \ln x_{kit} + \hat{\beta}_{tj} t \right) \tag{4.7}$$

where a hat over a parameter indicates that it is an estimated value. The technology exhibits constant returns to scale ($\varepsilon_{it} = 1$) for $\sum_{j=1}^{J} \hat{\beta}_j = 1$; $\hat{\beta}_{jk} = 0 \; \forall \, j, k$ and $\hat{\beta}_{tj} = 0 \; \forall \, j$. Output elasticities ε_{jit} and associated scale elasticities vary across producers and time unless $\hat{\beta}_{jk} = 0 \; \forall \, j, k$ and $\hat{\beta}_{tj} = 0 \; \forall \, j$, respectively.

The primal rate of technical change for firm i in time t is calculated as

$$\Delta T_{it} = \frac{\partial \ln f(\boldsymbol{x}, t)}{\partial t} = \hat{\beta}_t + \hat{\beta}_{tt} t + \sum_{j=1}^{J} \hat{\beta}_{tj} \ln x_{jit} \tag{4.8}$$

Technical change varies across producers unless it is Hicks-neutral with respect to inputs ($\hat{\beta}_{tj} = 0 \; \forall \, j$) and across periods except $\hat{\beta}_{tt} = \hat{\beta}_{tj} = 0 \; \forall \, j$. The technical efficiency change of firm i in time t can be derived from

$$\Delta TE_{it} = \frac{\partial \ln \widehat{TE}(t)}{\partial t} = -\frac{\partial \hat{u}_{it}}{\partial t} = \hat{u}_i \hat{\eta} \, exp\big(-\hat{\eta}(t - T)\big) \tag{4.9}$$

This expression also varies across producers unless $\hat{u}_i = u \; \forall \, i$ and across periods with the same trend for all i unless $\hat{\eta} = 0$, but the latter case would imply a time-invariant technical efficiency. Finally, the within-firm component as measured by the first right-hand-side term in equation (4.1) is calculated as

$$\sum_{i \in C} \widetilde{s_{it}} \, \Delta LP_{it} = \sum_{i \in C} \widetilde{s_{it}} \, (ID_{it} + SC_{it} + TC_{it} + TE_{it}) \tag{4.10}$$

where s_{it} is a firm's share in the total wage expenditures.

4.3. Data and Empirical Implementation

We use an unbalanced panel of 118 German breweries that were participating in a voluntary benchmarking program conducted on behalf of the German brewers' association over a period of 13 years from 1996 to 2008. On average, each brewery was observed for approximately 7 years, which resulted in 827 observations. We excluded microbreweries that produce less than 5,000 hl/year and very large breweries that produce more than 400,000 hl/year [84] from the sample because it can be expected that these breweries use different production technologies. Most breweries produce less than 100,000 hl/year (70.3%) on average. Hence, the breweries in the sample are small and midsized businesses with an average of 48 employees and revenues of 7.8 million €. Nevertheless they represent the core of the German brewing sector. Most of the observed breweries are located in Bavaria (57%) and Baden-Württemberg (19%) in southern Germany.

Table 4-1 summarizes the descriptive statistics for the input and output variables. We aggregate the inputs into three categories: *materials* including expenses for malt and barley, hops, water, energy and purchased goods and services; *labor* measured by the total wages[85] paid; and *capital* given by the end of year value of all machinery, equipment and buildings. Using appropriate price indices from the German Federal Statistical Office (Destatis), all the monetary values were deflated to base year 2005 values. Output is measured by total revenues deflated by a firm-specific price index. This allows us to take any price dispersion between the breweries and price changes over time into account and create a quantity-type measure of output[86] and productivity. Compared to the use of a common industry-based price index as a deflator, this approach is beneficial in two ways. First, we avoid an omitted variable bias in the econometric estimation of the production technology. Klette and Griliches (1996) note that, in most cases, omitted price dispersion will be negatively correlated with input quantities and introduces a downward bias in the

[84] The original sample contained a few observations of very large breweries that produce up to 2.2 million hl/year.

[85] We use data on the wages instead of the mere number of employees because we are missing information on the actual work hours, the educational status and tenure of employees in the firms. Hence, we follow Fox and Smeets (2011), who show that the wage bill is a good approximation of quality adjusted labor input among others in the Danish food and beverages industry.

[86] Our dataset contains information on the physical production and the respective revenues from various categories of beverages. These categories include beer, beer-mix beverages, and non-alcoholic beverages, which are all distinguished by whether they are packaged in bottles or kegs and by beer produced in license brewing. From the reported revenues and the physical output, we calculate category-specific prices that are then aggregated to a firm-specific price index using the categories revenue shares as weights. This index is also normalized using the year 2005 as the base, i.e., the average price index across all firms in the year 2005 is equal to 100. Eslava et al. (2004), Mairesse and Jaumandreu (2005) and Ornaghi (2006) also use firm-specific prices to deflate revenues to generate a quantity-type measure of output.

estimated scale elasticities. Second, we ensure that we measure physical productivity growth that is free of demand-side price effects[87]. Abbott (1990) showed that revenue-based productivity growth equals physical productivity plus a price change component. In addition, Foster et al. (2008) show how firms' output prices are positively correlated with firm-specific demand factors and negatively correlated with physical total factor productivity.

Table 4-1: Summary statistics of input and output variables

	Mean	Max.	Min.	Std. Dev.
Material (1000 €)	2846.7	13776.7	210.0	2730.1
Labor (1000 €)	1833.8	6530.7	99.8	1376.8
Capital (1000 €)	3577.5	26523.3	210.4	3523.8
Output (1000 €)	7849.2	57703.5	584.0	6359.9

As summarized in table 4-2, we perform several specification tests on our empirical model in equation (4.6). The hypotheses that the Cobb-Douglas production function, which is a special case of the translog functional form, is a sufficient specification of the production technology is rejected at the 1% level. We also reject the hypotheses that all breweries are technically efficient and that they operate on the production frontier. This result favors the stochastic frontier model over the conventional average production function approach. The hypothesis of zero and Hicks neutral technical change is rejected at the 1% level. Hence, the technical change component has a significant effect on output growth. We reject the hypotheses of constant returns to scale and time-invariant technical efficiency at the 1% and the 5% level, respectively. These results indicate that all three components of the within TFP growth contribute to growth in labor productivity and should be included in (4.2). Based on a Hausman[88] test, we can reject the null hypothesis that the individual effects α_i are not correlated with the explanatory variables at the 1% level. Moreover, the null hypothesis that all auxiliary group mean variables of the Mundlak adjustment are jointly equal to zero is rejected at the 1% level. We take the

[87] We do not observe firm-specific prices on the input side of production. Hence, our measure of "physical" productivity may still contain price effects on the input side, i.e., firms that face higher factor prices will appear to utilize a relatively higher level of inputs and thus to be less productive. As Foster et al. (2008) note, using quantity output, productivity reflects firms' "idiosyncratic cost components, both technological fundamentals and factor prices." See also Ornahghi (2006) for a discussion on the effects of input price differences.

[88] The test statistic is based on the comparison of the estimates of conventional fixed and random effects models.

results of the last two tests as an indication that the input variables are correlated with individual effects i.e., unobserved firm heterogeneity. All the tests together confirm our model specification in equation (4.6).

Table 4-2: Model specification tests

Hypotheses	**LR-statistic**	**Critical value** (α = 0.05 / 0.01)
Cobb-Douglas ($H_0: \beta_{jk} = \gamma_{jk} = 0, \forall\, j,k$)	56.25	${\chi^2}_{12}$ = 21.03 / 26.22
Technical efficiency ($H_0: \sigma_u = 0$)	1014.72	${\chi^2}_{1}$ = 2.71 / 5.41 [a]
Zero technical change $H_0: \beta_t = \beta_{tt} = \beta_{tj} = \gamma_{tj} = 0, \forall\, j$	55.19	${\chi^2}_{8}$ = 15.51 / 20.09
Hicks neutral technical change $H_0: \beta_{tj} = \gamma_{tj} = 0, \forall\, j$	32.31	${\chi^2}_{6}$ = 12.59 / 16.81
Constant returns to scale ($H_0: \sum \beta_j = 1; \sum \beta_{1j} = \sum \beta_{2j} = \sum \beta_{3j} = \sum \beta_{tj} = 0\ \forall\, j$)	15.71	${\chi^2}_{5}$ = 11.07 / 15.09
Time-invariant technical efficiency ($H_0: \eta = 0$)	5.78	${\chi^2}_{1}$ = 3.84 / 6.63
Individual effects ($H_0: \alpha_i \perp x_{kit}$)	32,32	${\chi^2}_{14}$ = 23.69 / 29.68
Individual effects $H_0: \gamma_j = \gamma_{kj} = \gamma_{tj} = 0\ \ \forall\, k,j$	118,94	${\chi^2}_{9}$ = 16.91 / 21.67

[a] Kodde and Palm (1986)

4.4. Results and Discussion

The estimated parameters for the production frontier and the composed error term are reported in table 4-3. The coefficients of the first-order parameters are positive and significantly different from zero. The coefficients of the trend variables are positive but not significantly different from zero. However, the significant positive and negative coefficient of the variables *material* and *labor* interacted with the time trend, indicate material-using and labor-saving technical change, respectively. We check whether the theoretical requirements for a well-behaved production function implied by economic theory, namely the monotonicity and quasi-concavity, are met at all the data points. Flexible functional forms such as the translog functional form, in contrast to the Cobb-Douglas, do not meet these requirements globally (Lau, 1978; Diewert and Wales, 1987). Hence, the function's properties must be imposed or checked a posteriori to avoid serious implications for the interpretation of the obtained parameters and efficiency scores (Sauer et al., 2006; Henningsen and Henning, 2009). In a production function, monotonicity requires positive marginal products for all inputs. Because both y and $\boldsymbol{x}$ contain only strictly positive numbers, it is sufficient to check the sign of the output elasticities (ε_{jit}) at all data points. We find no violations of monotonicity.

Table 4-3: Parameter estimates of the production frontier

Parameter	Coefficient	S.E.
β_0	0.2354	0.0342
β_1 $(Material)$	0.3262	0.0283
β_2 $(Labor)$	0.5233	0.0320
β_3 $(Capital)$	0.0477	0.0131
β_{11}	0.0777	0.0587
β_{22}	0.1162	0.0913
β_{33}	0.0180	0.0172
β_{12}	-0.1068	0.0668
β_{13}	-0.0089	0.0323
β_{23}	-0.0134	0.0283
β_t $(Trend)$	0.0076	0.0047
β_{tt} $(Trend^2)$	0.0005	0.0005
β_{t1} $(t * Material)$	0.0076	0.0033
β_{t2} $(t * Labor)$	-0.0111	0.0035
β_{t3} $(t * Capital)$	0.0003	0.0016
γ_1	0.9392	0.1932
γ_2	-1.0591	0.1757
γ_3	0.1190	0.1190
γ_{11}	0.3268	0.2852
γ_{22}	1.0329	0.4099
γ_{33}	0.2557	0.1362
γ_{12}	-0.4502	0.3404
γ_{13}	0.0750	0.1898
γ_{23}	-0.4650	0.1461
γ_{t1}	-0.1134	0.0251
γ_{t2}	0.1132	0.0234
γ_{t3}	0.0143	0.0162
$\lambda = \sigma_u/\sigma_v$	5.4288	0.0219
σ_u	0.4164	0.0089
η	-0.0174	0.0048
$Log\ LF$	726.53	

[a,b,c] statistical significance on 1%, 5 %, 10% level

To check for quasi-concavity, we find that the condition of a negative semi-definite bordered Hessian matrix of the first- and second-order derivatives is met in more than 98% of the data points. Hence, we conclude that the estimated translog production frontier is well-behaved and satisfies the regularity conditions of monotonicity and quasi-concavity very well.

Based on the input data and the estimated coefficients of the production frontier, we calculate the output elasticities and returns to scale. The reported elasticities in table 4-4 indicate that the inputs material and labor contribute the most to the production of beer on average and that the impact of additional capital on production is rather low. We observe decreasing returns to scale at the sample mean.

Table 4-4: Average output elasticities and returns to scale

Average output elasticities		
Material	0.385	(0.022)[a]
Labor	0.460	(0.025)[a]
Capital	0.050	(0.009)[a]
Returns to scale	0.894	(0.022)[a]

[a] Standard errors computed using Krinsky and Robb (1986)

In table 4-5, we present the decomposition of labor productivity growth for three periods of four years, respectively. This is done to avoid year-to-year fluctuations. Between 2004 and 2008, we measure an aggregated sector-wide labor productivity change of 4.71%. However, part of the measured change comes from firms that enter (-0.23%) or exit (0.69) the sample but not the industry. Abstracting from this issue, the aggregated labor productivity change of our sample of 118 firms is 4.07%. The biggest share of this change (3.86%) is due to productivity increases within the firm. Within the firm, mainly technical change (4.63%) and an increase in scale efficiency (1.40%) are important, but the decrease in the firms' technical efficiency had a significant negative effect (-1.15%). Neither deepening of material nor capital play a major role. Additionally, changes within the firms' aggregated labor productivity can increase from industry dynamics. We distinguish between two effects: a shift of shares from less to more productive firms (0. 77%) and industry exits of firms that are less productive than the average firm (1.06%). If we take all three time periods into account, the following tendencies seem important: i.) The within-firm effect is almost twice as strong as the effects from industry dynamics (between-firms and industry-exiting effect). ii.) The within-firm effect is mainly driven by technical change. Hence, the main strategy of these mid-sized firms to increase productivity is to maintain their market shares rather than to increase their market shares through acquisition. iii.) The technical efficiency of firms decreases over time.

Table 4-5: Decomposition of aggregate labor productivity growth

Component	96/00	00/04	04/08
Within firms	1.44	2.35	3.86
Deepening	0.21	0.91	0.01
Material	0.17	0.99	-0.25
Capital	0.03	-0.09	0.26
TFPC	1.24	1.44	3.85
Technical change	1.64	2.86	4.63
Tech. eff. change	-1.15	-1.57	-2.18
Scale eff. change	0.74	0.15	1.40
Industry dynamics	0.97	1.19	1.83
Between firm	0.44	1.36	0.77
Exit (industry)	0.53	-0.17	1.06
Exit (sample)	0.72	3.17	0.69
Entry (sample)	1.55	1.81	-0.23
Residual[a]	-0.35	0.01	-1.44
Overall aggregate	4.34	8.53	4.71
Overall aggregate without sample exits and entries	2.07	3.55	4.25

[a] The residual category is necessary because of differences between the directly calculated within-firm effect in equation (1) (first right-hand side term) and the one measured from the estimated production frontier (equation (10)).

This gives us the following picture. Although the production frontier as formed by the best firms is shifted upwards, not all firms are able to follow this development. Hence, the performance of the firms in the sector diverges. This seems typical for shrinking sectors with some firms following an active strategy and investing in new technology and others staying passive and producing as long as possible with the existing technology. It is important to note here that most mid-sized breweries in Germany are family owned and that many have a long tradition of brewing. As opposed to listed companies, such family businesses are often driven by continuity rather than rates of returns. iv.) Within the firm, the scale change effect is less important than technical change but clearly positive. In the light of decreasing returns to scale, firms reduce their input usage and improve their productivity by adjusting the scale of their operations. v.) Except for material in the second period, the deepening effect is rather minor. This points at a rather limiting production technology with minor substitution effects between production factors. v.) The reallocation effect also contributes to aggregated labor productivity; however, changes within firms are more important. vi) The effect of firms that enter and exit the sample but not the sector is quite large. This shows the sensitivity of our results to changes in our sample. vii.) Overall labor productivity increases over time because both the within-firm effect and the effect of industry restructuring increase.

4.5. Conclusions

Although the global beer market is dominated by five big, international players, the number of breweries in Germany, which is the largest beer producing country in the EU, is still large and with a relatively low concentration. Nevertheless, significantly decreasing domestic demand and only slightly increasing net-exports for more than two decades has caused structural changes. In particular, although the number of microbreweries and especially those with a production of less than 1000 hl/year has steadily increased and the number of very large firms with more than 1 million hl/year has remained fairly stable, the number of medium sized breweries (5,000 – 100,000 hl/year) has sharply decreased. According to NGG (2013), 40 producers with 320 production plants were put out of business between 2006 and 2012.

In this paper, we examine how labor productivity in this segment of the sector (5,000 – 400,000 hl/year) developed based on a sample of 118 breweries between 1996 and 2008. We provide a method to decompose industry labor productivity into seven components: input deepening, technical change, technical efficiency, scale effect, reallocation effect and the effects from exits and entry. Although the first four of these components occur within firms, the last three can be attributed to industry dynamics. Our main finding is that the within effect is much more important than changes within the sector. Moreover, although most of the increase in labor productivity comes from technical change, technical efficiency decreases over time. In addition to technical change, some increase in labor productivity comes from adjusting the size of the firm to the decreased demand (rationalization). All these findings together fit well into a sector that is dominated by mid-sized, family owned businesses with a long tradition. On average, these firms stay rather passive and either try to defend their market shares by becoming more productive through investments in technology or continue producing with the old technology as long as possible.

5. DAIRY FARMING ON PERMANENT GRASSLAND: CAN THEY KEEP UP?

Abstract

Based on an extensive data set for southern Germany, we compare the productive performance of dairy farms that operate solely on permanent grassland and dairy farms using fodder crops from arable land. We allow for heterogeneous production technologies and identify more intensive and extensive production systems for both types of farms, whereby we base our notion of intensive vs. extensive dairy production on differences in stocking density and milk yield per cow and year. To be able to compare the productivity levels and productivity developments of the various groups of farms, we develop a group- and chain-linked multilateral productivity index. We also analyze how technical change, technical efficiency change and a scale change effect contribute to productivity growth between the years 2000 and 2008. Our results reveal that permanent grassland farms can generally keep up with fodder-crop farms, even in an intensive production setting. However, extensively operating farms, especially those on permanent grassland, significantly lag behind in productivity and productivity change and run the risk of losing ground.

Keywords: dairy farm, permanent grassland, total factor productivity, stochastic production frontier

JEL classifications: D24, Q12

This chapter is based on the article *Dairy farming on permanent grassland: can they keep up?* by Magnus Kellermann and Klaus Salhofer, published in the *Journal of Dairy Science* (Vol. 97, pp. 6196–6210; DOI: 10.3168/jds.2013-7825). The author of this dissertation is the main author of the article.

5.1. Introduction

In addition to being an important basis for agricultural production, grasslands provide a variety of essential environmental and social benefits. For instance, grasslands act as a carbon sink (Soussana et al., 2007) and generally ensure a high level of biodiversity because they provide habitats for flora and fauna (Pflimlin and Poux, 2005). The preservation of ground and surface water quality and the provision of an attractive environment for recreational activities and tourism are additional benefits (e.g., Prochnow et al., 2009; Peeters, 2009; Sanderson et al., 2012). Hence, the preservation of permanent grassland is an important topic in the agricultural conservation policy of the United States and the European Union. The US Grassland Reserve Program was established as part of the 2002 Farm Bill and is one example for these efforts (USDA, 2013). In the EU, several agri-environmental programs contain grassland protection elements. In addition, plans exist to strengthen legislation that prevents the conversion of grassland to arable land as part of the *greening* strategy of the 2013 reform of the Common Agricultural Policy (European Commission, 2013). The productive potential of permanent grassland can be exploited only by ruminants and, with some limitations, biogas plants. Hence, dairy farming plays the key role in agricultural production in many grassland regions. In the European heartland, the regions with agricultural production based solely on permanent grassland are generally found in elevated and mountainous areas, e.g., in the surroundings of the Alps and the Massif Central. Dairy farms in these areas often face some natural disadvantages. Most notably, the cultivation of fodder crops, such as corn silage, is not feasible because of comparably high precipitation, lower average annual temperature and a shorter vegetation period (Hein, 2002; Meisser and Wyss, 1998). The relatively low energy yield per hectare of permanent grassland compared to corn silage illustrates these circumstances effectively. In 2010, the numbers vary between 42 – 67 GJ NE_L per ha for grass silage and 87 – 110 GJ NE_L per ha for corn silage for dairy farms in Bavaria (LfL, 2012). Moreover, Thaysen et al. (2010) show that the disadvantage of grassland regarding the energy content of the forage increases over time. Analyzing data from northern Germany between 1985 and 2008, they find an average annual increase in NE_L yield of approximately 1 GJ per ha and year for corn silage compared to only approximately 0.45 GJ per ha and year for grass silage. Nevertheless, grassland dairy farmers have to compete with farmers growing fodder crops on arable land because in most cases they are acting in the same markets. First, the distances to the more favorable areas are minor. Referring to the zones of ruminant rearing systems in Europe identified by Pflimlin et al. (2005), we find areas labeled as “arable land and livestock regions” and “forage crops regions with temporary grassland plus corn” in close proximity to the permanent grassland regions in parts of Austria, Switzerland, southern Germany and the eastern part of France. Second, the produced milk is not promoted to generate higher farm prices in many cases, for example as “mountain- or hay-milk.” However, certain approaches in the marketing of these products can be observed, e.g., the promotion of high-quality cheese in the context of a Protected Designation of Origin. Given the ongoing market liberalization in the dairy sector and the latest farm-price fluctuations, serious concerns exist (e.g., Hopkins, 2011) as to whether dairy farms that operate solely on permanent grassland can compete with farms that use arable land to produce fodder crops.

The objective of this paper is to measure the levels and growth rates of total factor productivity (TFP) of dairy farms in Bavaria and examine whether grassland dairy farms are able to keep up with their fodder-crop counterparts in terms of productive performance. If dairy farming in permanent grassland areas is getting less productive compared to areas with arable land, either agricultural production will be abandoned in these regions or payments directed towards these areas, e.g., less-favored area payments, have to increase over time. MacDonald et al. (2000) discuss some of the undesirable effects agricultural abandonment in mountainous regions can have on environmental parameters, e.g., reductions in biodiversity and landscape quality.

In general, when comparing the productivity of various groups of farms (e.g., organic vs. conventional, intensive vs. extensive, irrigated vs. rain-fed, country A vs. country B) it is important to have information on both the difference in absolute productivity levels and the differences in productivity growth. Only the combination of these components can give a full picture of the present and future performance of one group compared to another. Nevertheless, many studies on the performance of groups of dairy farms concentrate on differences in the TFP growth rates and its decomposition; examples are Brümmer et al. (2002) for the dairy sector in various EU countries, Newman and Matthews (2006) for specialist versus “other” dairy farms and, more recently, Ma et al. (2012) for dairy farms of various size classes. We follow this strand of the literature and calculate total factor productivity growth. Using the generalized Malmquist productivity index described by Orea (2002), we decompose productivity growth into technical change, technical efficiency change and a scale change effect. However, this procedure is not enough to fully answer our research question. Two groups can have equal growth rates, and yet one of them may be much less productive. Likewise, 2 groups can be equally productive at a point in time and still drift apart over time because of very different growth rates. To get a full picture of what is going on in this sector, we need to map the TFP levels of both groups over time. To do this, we provide a group- and chain-linked multilateral productivity index based on the indices first introduced by Caves et al. (1982a) and refined by Good et al. (1997). This index makes it possible to analyze the variation of TFP within and between various groups in a sample and how this variation develops over time. We note that throughout this article the term productivity refers to total factor productivity, i.e., an index of all aggregated outputs over all aggregated inputs.

5.2. Materials and Methods

5.2.1 Empirical model

To model the multi-input, multi-output technology of agricultural production, we use a parametric output-oriented distance function $D^O\left(\boldsymbol{x},\boldsymbol{y},t\right)$, where $\boldsymbol{x}$ refers to a nonnegative vector of inputs used to produce a nonnegative vector of outputs $\boldsymbol{y}$ in time period t. See Färe and Primont (1995) for the theoretical derivation of the distance function and its properties. We choose output orientation because we assume that the farms in our sample are less flexible in the adjustment of their inputs than their outputs. Labor input, which predominantly consists of family workforce, is one example for a rather inflexible input. Breustedt et al. (2011) also note the low flexibility of the inputs of labor and land in Bavarian dairy farming. Contrariwise, although the aggregated amount of milk is limited by the quota system, the very well established quota trading system in Germany assures unrestricted output at the single farm level. Hence, we argue that farmers decide on a set of short-term inflexible inputs for a given year and aim to obtain the maximum output from those inputs. Our assumptions are in line with Brümmer et al. (2002), Newman and Matthews (2007) and Emvalomatis (2012) who also choose output orientated distance functions as representations of production technologies for dairy farms in various EU countries. We use a flexible translog functional form to limit a priori restrictions on the relationships among inputs and outputs (Morrison Paul et al., 2000; Karagiannis et al., 2004).

Hence,

$$
\begin{aligned}
&\ln D_{it}^{O}\left(\boldsymbol{y},\boldsymbol{x},t\right)= \\
&\alpha_0+\sum_{m=1}^{M}\alpha_m \ln y_{mit}+\sum_{k=1}^{K}\beta_k \ln x_{kit}+\frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M}\alpha_{mn}\ln y_{mit}\ln y_{nit}+ \\
&\frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{K}\beta_{kj}\ln x_{kit}\ln x_{jit}+\sum_{m=1}^{M}\sum_{k=1}^{K}\delta_{mk}\ln y_{mit}\ln x_{kit}+ \\
&\tau_1 t+\frac{1}{2}\tau_2 t^2+\sum_{m=1}^{M}\zeta_{mt}\,t\,\ln y_{mit}+\sum_{k=1}^{K}\theta_{kt}\,t\ln x_{kit}
\end{aligned}
\tag{5.1}
$$

In equation (5.1), the subscripts $i=1,2,...,\mathrm{N}$ and $t=1,2,...,T$ denote individual farms and time periods, respectively; the subscripts $k,\, j=1,2,\ldots,K$ indicate various types of inputs, and $m,n=1,2,\ldots,M$ indicate various types of outputs; $\alpha,\,\beta,\,\delta,\,\tau,\,\zeta,\,\theta$ are parameters to be estimated.

The parameters of this function must satisfy the symmetry restrictions $\alpha_{mn}=\alpha_{nm}$ and $\beta_{kj}=\beta_{jk}$. We follow Lovell et al. (1994) and Coelli and Perelman (2000) and impose homogeneity of degree 1 in output quantities ($\sum_{m=1}^{M}\alpha_m=1$ and $\sum_{m=1}^{M}\alpha_{mn}=\sum_{k=1}^{K}\delta_{mk}=\sum_{m=1}^{M}\zeta_{mt}=0$) by normalizing the function by 1 of the output quantities:

$$
\ln\left(\frac{D_{it}^{O}\left(\boldsymbol{y},\boldsymbol{x},t\right)}{y_{Mit}}\right)=TL\left(\boldsymbol{y}^{*},\boldsymbol{x},t\right)\ \text{with}\ y_{mit}^{*}={y_{mit}}\big/{y_{Mit}}
\tag{5.2}
$$

where TL indicates the translog functional form, and $TL\left(\boldsymbol{y}^{*},\boldsymbol{x},t\right)$ is the right hand side of equation (5.1) after dividing all the output quantities by y_M. Because the dependent variable $\ln D_{it}^{O}$ is unobservable, we have to rearrange the distance function for estimation. We add a random error term v_{it}, and, given that $\ln D_{it}^{O}\le 0$, we replace $\ln D_{it}^{O}$ with $-u_{it}$ such that

$$
-\ln y_{Mit}=TL\left(\boldsymbol{y}^{*},\boldsymbol{x},t\right)+u_{it}+v_{it}
\tag{5.3}
$$

Equation (5.3) is a “stochastic frontier model” (Kumbhakar and Lovell, 2000) and can be estimated by maximum-likelihood methods given that v_{it} is a normally distributed random variable $N\left(0,\sigma_v^2\right)$ that reflects statistical noise and other stochastic shocks, and u_{it} is a non-negative random error term $N^{+}\left(0,\sigma_{u}^{2}\right)$ that represents inefficiency (Aigner et al., 1977).

Because the aim of the paper is to compare the efficiency and productivity of permanent grassland farms and fodder-crop farms, we split our sample into these 2 groups. We know each farm’s share of permanent grassland in its total utilized land. This information is used to distinguish between the 2 types of farms. However, we have only limited information on potentially heterogeneous production technologies within those 2 groups. To allow for the possibility of unobserved differences in the production technologies, we apply a latent class stochastic frontier model for each of the 2 groups as described in Orea and Kumbhakar (2004), Greene (2005) and Alvarez and del Corral (2010).

The reader might ask why we did not use the latent class model on the complete dataset and let the econometric model assign the farms to classes with different (unobserved) production technologies. Alvarez et al. (2012) find that latent class models are better suited to identify heterogeneous technologies than is an a priori split of the sample. However, our research question aims to identify productivity differences between 2 groups of farms, which are not latent to us; rather, we have a clear conception for their separation based on the farms’ shares of permanent grassland. Still, technological homogeneity can be a strong assumption even within those more closely defined subsamples. Hence, we use the latent class model to relax this restriction and allow for unobserved technological heterogeneity.

In the latent class framework, equation (5.3) can be rewritten as

$$-\ln y_{Mit} = TL\left(\boldsymbol{y}^{*},\boldsymbol{x},t\right)|_{g} + u_{it}\,|_{g} + v_{it}\,|_{g} \tag{5.4}$$

where the vertical bar simply indicates that we estimate a specific set of coefficients for each class $g = 1,\ldots,G$, and the overall functional relationship remains the same for all the classes. Hence, the heterogeneity in the production technology is captured by a class-specific parameter vector. The true class membership of each farm is unknown to us. It is assumed that a latent relationship between the observations in the sample exists, which translates into G different classes. Following Greene (2005), under the aforementioned distributional assumptions on u_{it} and v_{it}, the contribution to the conditional likelihood function for each farm i is the product of the likelihood functions in each period:

$$LF_{ig} = \prod_{t=1}^{T} LF_{itg} \tag{5.5}$$

where LF_{itg} is the likelihood function for each observation in each class. To get the unconditional likelihood function for farm i, a weighted average of all its likelihood functions over the G classes is calculated using the G prior probabilities of class membership P_g as weights:

$$LF_i = \sum_{g=1}^{G} P_g LF_{ig} \tag{5.6}$$

The prior probabilities must be specified to have a value between 0 and 1($0 \leq P_g \leq 1$) and must sum up to 1 ($\sum_{g=1}^{G} P_g = 1$, $\forall i$). A simple parameterization of the prior probabilities with 1 class (G) as reference that fulfills these requirements is in the form of a logit model:

$$P_g = \frac{\exp\left(\rho_g\right)}{\sum_{g=1}^{G} \exp\left(\rho_g\right)} \quad , \quad \rho_G = 0 \tag{5.7}$$

where ρ_g are parameters to be estimated. We impose the constraint on the last ρ_G because we only need $G-1$ parameters to specify the G probabilities. Consequently, the last probability is 1 minus the sum of the first $G-1$ (Greene, 2003 p. 440). Using this formulation, the prior probability of class membership is constant for all the individuals. Based on the concomitant variable model by Dayton and MacReady (1988), we incorporate variables on characteristics of each individual i. By this means, the individual specific prior probabilities of class membership can be parameterized by a multinomial logit model:

$$P_{ig} = \frac{\exp\left(\boldsymbol{\rho}_g \boldsymbol{z}_i\right)}{\sum_{g=1}^{G} \exp\left(\boldsymbol{\rho}_g \boldsymbol{z}_i\right)} \quad , \quad \boldsymbol{\rho}_G = 0 \tag{5.8}$$

The vector $\boldsymbol{z}_i$ contains time invariant farm specific characteristics, and $\boldsymbol{\rho}_g$ is the respective vector of parameters to be estimated. The parameter in $\boldsymbol{\rho}_g$ denotes the impact of the farm characteristic on the prior probability of belonging to class g . Hence, a positive coefficient in $\boldsymbol{\rho}_g$ implies that a higher value of the respective variable increases the prior probability that a farm i belongs to class g (Wedel and DeSarbo, 2002). If concomitant variables are incorporated in the model to derive individual specific estimates of the prior probabilities, equation (5.6) changes accordingly and P_g is replaced by P_{ig}. We derive the log-likelihood-function used to estimate the parameters of the production frontier, the composed error term and the prior class probabilities from the sum of the individual log LF:

$$\log LF = \sum_{i=1}^{N} \ln LF_i = \sum_{i=1}^{N} \ln \sum_{g=1}^{G} P_{ig} \prod_{t=1}^{T} LF_{itg} \tag{5.9}$$

Following Greene (2005), the log-likelihood-function (5.9) is maximized with respect to the parameter set $\left[\mathbf{\Theta}_{\mathrm{g}}\right]$, where $\mathbf{\Theta}$ contains all the parameters of the stochastic distance function $\left[\boldsymbol{\alpha}_g, \boldsymbol{\beta}_g, \boldsymbol{\delta}_g, \boldsymbol{\tau}_g, \boldsymbol{\zeta}_g, \boldsymbol{\theta}_g, \sigma_g^2\right]$ and the prior class probabilities $\left[\boldsymbol{\rho}_{\mathrm{g}}\right]$. The estimated parameters are then used to estimate the conditional posterior probabilities of class membership from

$$P(g|i) = \frac{LF_{ig}\, P_{ig}}{\sum_{g=1}^{G} LF_{ig}\, P_{ig}} \quad . \tag{5.10}$$

As noted by Orea and Kumbhakar (2004), we can deduce from expression (5.10) that the posterior class probabilities depend not only on the estimated $\boldsymbol{\rho}$ parameters for the concomitant variables in the logit model (5.8) but on all the parameters contained in the set $\left[\mathbf{\Theta}_{\mathrm{g}}\right]$. Hence, if information about possible class characteristics is unavailable, the latent class model still clusters the sample using the goodness of fit of each estimated frontier.

Although we have panel data available, we estimate the posterior class probabilities as $P(g|i)$ instead of $P(g|it)$. This means that the posterior probabilities are the same for each observation of a farm and that the farms are not allowed to switch between the classes over time contrary to Alvarez and del Corral (2010). We argue that dairy farmers do not decide whether they produce using a more intensive or extensive production technology on a year-to-year basis. Rather, this decision has to be considered as the result of a medium- to long-term strategy because each production technology has particular requirements that are not always easy to adjust every year, e.g., special knowledge by the farmer or specific dairy breeds.

We follow Orea and Kumbhakar (2004) and use the Akaike Information Criterion (AIC) as an indication for the appropriate number of classes. For the econometric implementation of the described model, we use LIMDEP 10.0 (Greene, 2012; Econometric Software Inc., Plainview, NY).

5.2.2 *Production technology, technical efficiency and productivity*

Based on the stochastic frontier model in equation (5.4), we can examine various characteristics of the production technology and derive the technical efficiency and productivity growth of the farms in our sample. The elasticities of the distance functions with respect to inputs and outputs characterize the multi-output production technology in a way similar to the more common parameters of a Cobb-Douglas production function. The first-order elasticities with respect to inputs

$$\varepsilon_{kit} = \frac{\partial \ln D^O}{\partial \ln x_k} = \beta_k + \sum_{j=1}^{K} \beta_{kj} \ln x_{jit} + \sum_{m=1}^{M} \delta_{mk} \ln y_{mit} + \theta_{kt}\, t \tag{5.11}$$

are therefore interpreted as the percentage change in the overall output due to a 1% increase in the respective inputs (Morrison Paul et al., 2000). The sum of these input elasticities is defined as the scale elasticity (Färe and Primont, 1995):

$$\varepsilon_{it} = \sum_{k=1}^{K} \varepsilon_{kit} \tag{5.12}$$

The first-order elasticities with respect to output are calculated as

$$\varepsilon_{mit} = \frac{\partial \ln D^O}{\partial \ln y_m} = \alpha_m + \sum_{n=1}^{M} \alpha_{mn} \ln y_{nit} + \sum_{k=1}^{K} \delta_{mk} \ln x_{kit} + \zeta_{mt}\, t \tag{5.13}$$

and reflect the profit maximizing revenue share of the m-th output for a given level of inputs (Brümmer et al., 2002). Grosskopf et al. (1995) provide a discussion of the various measures that can be derived from the estimated elasticities on the output side and their economic interpretation.

In contrast to the standard stochastic frontier model, in which 1 homogeneous production technology is assumed, the latent class model can establish several frontiers. In accordance with the estimated a posteriori probability from (5.10), the farms in the sample are assigned to the identified classes, and their individual technical efficiency score is measured against the respective class technology frontier. Consequently, the technical efficiency scores for individual farms are estimated using the common formula of Jondrow et al. (1982):

$$TE_{it} = \exp\left(-E\left(u_{it} \mid u_{it} + v_{it}\right)\right). \tag{5.14}$$

Hence, technical efficiency scores are a measure of the productive homogeneity of the respective class and cannot be used to compare the efficiency of farms across classes. Orea and Kumbhakar (2004) also discuss an alternative approach for the estimation of efficiency scores in a latent class model. It aims to take the uncertainty about class membership into account and calculates a weighted average of the farms' technical efficiency scores as against all possible frontiers:

$$TE_{it}^{*} = \sum_{g=1}^{G} P\left(g|i\right) TE_{it|g} . \tag{5.15}$$

Malmquist index of TFP growth

To examine productivity growth, we use the well-established generalized Malmquist index of total factor productivity as suggested by Orea (2002). Details on the derivation of the index and applications to the agricultural sector can be found in Coelli et al. (2005), Newman and Matthews (2006) and Key et al. (2008). The generalized Malmquist index of productivity change between 2 time periods t and $s = t + 1$ can be written as

$$\begin{aligned} \ln G_{it,is} &= \frac{1}{2}\left[\frac{\partial \ln D_{is}^{O}}{\partial t} + \frac{\partial \ln D_{it}^{O}}{\partial t}\right] + \left[\ln D_{is}^{O} - \ln D_{it}^{O}\right] \\ &+ \frac{1}{2}\sum_{k=1}^{K}\left[\left(\varepsilon_{is} - 1\right)\left(\varepsilon_{kis} / \varepsilon_{is}\right) + \left(\varepsilon_{it} - 1\right)\left(\varepsilon_{kit} / \varepsilon_{it}\right)\right]\ln\left(\frac{x_{kis}}{x_{kit}}\right) \end{aligned} \tag{5.16}$$

The index measures how the productivity of an individual farm changes over time and allows for the decomposition of productivity growth in technical change, technical efficiency change and a scale change effect. Given the estimated parameters of the stochastic distance function, the calculation of the components of productivity change is rather straightforward. Technical change (TC) – the first term on the right hand side of equation (5.16) – includes the partial derivatives of the distance function with respect to time for the periods t and $s = t + 1$, where $\frac{\partial \ln D_{it}^{O}}{\partial t} = \tau_1 + \tau_2 t + \sum_{m=1}^{M} \zeta_{mt} \ln y_{mit} + \sum_{k=1}^{K} \theta_{kt} \ln x_{kit}$.

The second term in equation (5.16) measures the change in technical efficiency (TEC) by the difference in the value of the output distance function from one period to the next. The third term on the right hand side of equation (5.16) contains the scale change effect (SC), which is based on the scale elasticity ε_{it} and changes in input use. We observe a positive (negative) contribution to productivity change if $\varepsilon_{it} > 1$ and input usage is expanded (reduced) or if $\varepsilon_{it} < 1$ and input usage is reduced (expanded). In the case of constant returns to scale ($\varepsilon_{it}=1$) or if input quantities do not change ($\ln\left(x_{kis}/x_{kit}\right)=0$), SC becomes 0.

Group- and chain-linked multilateral productivity index

As noted in the introduction, our analysis goes beyond the examination of the mere growth rates of productivity. To compare the productivity levels of fodder-crop and grassland farms relative to each other, we develop a group- and chain-linked multilateral productivity index. The initial form of this index, as introduced by Caves et al. (1982a), provides a measure of the productivity level of all the production units in a cross-section relative to one hypothetical average reference unit. We adapt this index to compare the productivity level of groups that consist of subordinate (micro-level) units as described in Delgado et al. (2002). This group-linked multilateral productivity index $\ln P_i$ can be written as

$$\begin{aligned}\ln P_i = &\left[\left(\frac{1}{2}\sum_{m=1}^{M}\left(R_{mi}^h+\overline{R_m^h}\right)\left(\ln y_{mi}^h-\overline{\ln y_m^h}\right)\right)+\left(\frac{1}{2}\sum_{m=1}^{M}\left(\overline{R_m^h}+\overline{\overline{R_m}}\right)\left(\overline{\ln y_m^h}-\overline{\overline{\ln y_m}}\right)\right)\right]\\ &-\left[\left(\frac{1}{2}\sum_{k=1}^{K}\left(C_{ki}^h+\overline{C_k^h}\right)\left(\ln x_{ki}^h-\overline{\ln x_k^h}\right)\right)+\left(\frac{1}{2}\sum_{k=1}^{K}\left(\overline{C_k^h}+\overline{\overline{C_k}}\right)\left(\overline{\ln x_k^h}-\overline{\overline{\ln x_k}}\right)\right)\right]\end{aligned} \tag{5.17}$$

In equation (5.17) the first term in the first square brackets is a multilateral output index that gives the relative output of the i-th unit in each group $h=1,\ldots,H$, against the output of a hypothetical group average unit. In our application, the superscript $h=1,\ldots,H$ includes all the G latent classes we identify among the permanent grassland and fodder-crop farms. Hence, R_{mi}^h denotes the unit-specific revenue share of the m-th output (y_m), and superscript h signifies the group membership of unit i. A single bar over a variable denotes its arithmetic mean for the respective group h (e.g., $\overline{R_m^h}=\frac{1}{N^h}\sum_{i=1}^{N^h}R_{mi}^h\ \forall h$). The second term in the first square brackets then compares the output of the hypothetical group average units themselves against a hypothetical overall average unit. A double bar over a variable indicates the arithmetic mean over the H groups (e.g., $\overline{\overline{R_m}}=\frac{1}{H}\sum_{h=1}^{H}\overline{R_m^h}$). The second square bracket performs the respective calculations on the input side.

Applying the first order conditions of cost minimization, we obtain the required input (cost) shares C_k under the assumption of allocative efficiency from the input-side elasticities of the estimated distance functions. Following Chan and Mountain (1983), C_k corresponds to the elasticities corrected by the respective returns to scale measure:

$$C_k = \frac{\partial \ln D^O}{\partial \ln x_k} / \sum_{k=1}^{K} \frac{\partial \ln D^O}{\partial \ln x_k} = \varepsilon_k / \varepsilon \qquad (5.18)$$

On the output side, we make use of the fact that the revenue shares are observed in the data. This way we can abstain from the assumption of an allocative efficient output mix and use the observed revenue shares for aggregation.

The resulting productivity index in (5.17) is purely cross-sectional. However, it allows interlinking various cross-sections or groups and provides a decomposition of the productivity variation within and between the various groups. The totaled productivity measure we obtain for all the units in all the groups consequently relates exclusively to the overall hypothetical reference unit. This way the "path of comparison" (Good et al., 1997 p. 11) is still well defined, and transitivity is maintained. The use of the arithmetic mean in equation (5.17) corresponds to the notion of democratic weights[89] for the construction of the hypothetical average reference units, i.e. every unit is given an equal weight in the construction of the group average unit and every group is given an equal weight in the construction of the overall average reference unit. However, different weighting schemes exit.[90]

Equation (5.17) yields equivalent results as the initial formulation by Caves et al. (1982a) if each groups share on the total number of observations is used as weights for the construction of the overall reference unit. A double bar over a variable then indicates the weighted arithmetic mean, e.g. $\overline{\overline{R_m}} = \sum_{g=1}^{G} \omega^g \overline{R_m^g}$ where $\omega^g = \frac{N^g}{N}$. Thus, groups with a larger (smaller) number of observations will have a higher (lower) impact on the overall reference unit. Provided that the number of observations for the different groups reflects the true distribution and importance of the groups in the industry (sector) under consideration this may be a desirable property. However, this may or may not be true when micro-level data is used. Considering the case that the number of observations in the groups is arbitrary, the overall reference unit will be biased towards the group with most observations. In three obvious cases the "observation share-" and the "democratic-"

[89] The term "democratic weights" was first used by Prais (1959) in an index number context.

[90] Diewert (1986) discusses various weighting schemes and their properties in the context of inter-country comparisons. He notes that no clear-cut answer exists on which formula is the best in all occasions. Depending on which properties are considered to be the most important in an empirical application, the "right" weights have to be chosen on a case-by-case basis.

weights yield the same results: if we observe the same number of observations in all groups, if all groups are on average equally productive and if productivity and group size are uncorrelated. For our dataset we do not have sufficient information to state that the number of yearly observations represents the true importance of the different groups in the sector. For this reason, but also because we want to focus on a general comparison of the productivity of the different identified groups in our dataset we use democratic weights[91].

We now further extend this group-linked multilateral productivity index for use with panel data in the way Good et al. (1997) describe for the comparison of single firms over time. Here, we calculate the group-linked multilateral productivity index for each period $t=1,\ldots,T$ and then chain-link the overall hypothetical reference units over time. The group- and chain-linked productivity index can then be written as

$$\begin{aligned}
\ln P_{it} = & \left[\left(\frac{1}{2}\sum_{m=1}^{M}\left(R_{mit}^{h}+\overline{R_{mt}^{h}}\right)\left(\ln y_{mit}^{h}-\overline{\ln y_{mt}^{h}}\right)\right)\right. \\
& +\left(\frac{1}{2}\sum_{m=1}^{M}\left(\overline{R_{mt}^{h}}+\overline{\overline{R_{mt}}}\right)\left(\overline{\ln y_{mt}^{h}}-\overline{\overline{\ln y_{mt}}}\right)\right) \\
& \left.+\left(\frac{1}{2}\sum_{s=2}^{T}\sum_{m=1}^{M}\left(\overline{\overline{R_{ms}}}+\overline{\overline{R_{ms-1}}}\right)\left(\overline{\overline{\ln y_{ms}}}-\overline{\overline{\ln y_{ms-1}}}\right)\right)\right] \\
& -\left[\left(\frac{1}{2}\sum_{k=1}^{K}\left(C_{kit}^{h}+\overline{C_{kt}^{h}}\right)\left(\ln x_{kit}^{h}-\overline{\ln x_{kt}^{h}}\right)\right)\right. \\
& +\left(\frac{1}{2}\sum_{k=1}^{K}\left(\overline{C_{kt}^{h}}+\overline{\overline{C_{kt}}}\right)\left(\overline{\ln x_{kt}^{h}}-\overline{\overline{\ln x_{kt}}}\right)\right) \\
& \left.+\left(\frac{1}{2}\sum_{s=2}^{T}\sum_{k=1}^{K}\left(\overline{\overline{C_{ks}}}+\overline{\overline{C_{ks-1}}}\right)\left(\overline{\overline{\ln x_{ks}}}-\overline{\overline{\ln x_{ks-1}}}\right)\right)\right]
\end{aligned} \tag{5.19}$$

This index then provides a measure of the level of productivity for each unit i in each group h and each time period t relative to the overall hypothetical reference unit in the base year $t=1$. Furthermore, we obtain information on the variability of TFP within and between the H groups and how this variability develops over time. In this application, we are mainly interested in productivity gaps between the identified groups of dairy farms. The multilateral Törnqvist index does not allow for the measurement of the effects of non-constant returns to scale or technical inefficiency to explain productivity dispersion. We

[91] One may also challenge our decision to use democratic weights for the construction of the groups' hypothetical reference units. However, we abstain from adding another layer of complexity by using revenue or quantity weights mainly because the relatively high number of observations and the fairly symmetric distributions of farm size in each group mitigate a possible "tiny country" effect (see e.g. Diewert 1986).

note that these restrictions are necessary for the calculation of the relative productivity measures across distinct production technologies.

5.2.3 Data

We employ an unbalanced panel dataset taken from farm bookkeeping records, which serves as a basis for the European Commission's Farm Accountancy Data Network. The farm level dataset includes 9,482 observations of 1,142 dairy farms observed over the years 2000 to 2008 in the Federal state of Bavaria in south Germany. All the farms are specialized dairy farms with more than 66% of the farms' total revenues coming from dairy production. The observations are evenly spread over the period under consideration with 8.3 observations per farm on average. Overall, 958 farms are identified as fodder-crop farms, and 184 farms are permanent grassland farms. The farms considered to be grassland farms are operating with 100% permanent grassland during the entire period. Though fodder-crop farms may operate partly on permanent grassland, it never exceeds 90% of their total farm land. The average share of permanent grassland in the fodder-crop group is 45%. All the monetary figures from accounting data are deflated using relevant price indices from the German Federal Bureau of Statistics.

Table 5-1: Summary statistics of main variables for fodder-crop and grassland farms

	Mean	Max.	Min.	Std. Dev.
	Fodder-crop (958 farms; 7,999 observations)			
Labor (fte)[1]	1.57	3.94	0.40	0.46
Land (ha)	46.2	318.3	9.4	26.0
Intermediate inputs (€)	54,171	247,089	6,145	29,462
Capital (€)	216,589	106,3470	14,767	129,886
Milk output (€)	75,950	310,611	9,122	39,307
Other output (€)	31,852	239,780	2,276	17,863
Cattle LU/ha forage land[2]	2.36	5.64	0.85	0.58
Yearly milk yield kg/ cow	6,079	9,162	2,253	1,001
	Grassland (184 farms; 1,483 observations)			
Labor (fte)[1]	1.49	2.97	0.35	0.35
Land (ha)	31.2	81.4	12.3	11.9
Intermediate inputs (€)	28,862	101,581	5,349	12,792
Capital (€)	152,885	398,067	21,395	81,251
Output (€)	68,623	204,891	21,301	24,072
Cattle LU/ha forage land[2]	1.54	3.06	0.70	0.35
Yearly milk yield kg/ cow	5,947	8,258	3,365	915

[1]fte = full-time equivalent

[2]LU = livestock unit

For the fodder-crop farms, we aggregate the outputs into 2 categories, *milk* and *other output*. *Milk* is equal to revenues from raw milk and milk products. This allows accounting for quality differences because the price that the individual farmer receives from the processor varies depending on the fat and protein content in the milk (see Reinhard et al., 1999 and Emvalomatis, 2012). The variable *other output* contains revenues from beef, crops and other commodities. We also tried a multi-output specification for the grassland farms. However, the estimated distance function encountered substantial monotonicity problems. This is not surprising, given that grassland farms are highly specialized. Their average share of dairy production on the total output is 92.6% with 80.2% as a minimum. Therefore, as common practice in the case of highly specialized dairy farms (e.g., Abdulai and Tietje, 2007), we estimate a production frontier that is the single-output special case of an output distance function. For both groups of farms, we aggregate the utilized inputs into 4 categories (*labor*, *land*, *intermediate inputs* and *capital*). Table 5-1 summarizes the descriptive statistics for the main variables of the stochastic frontier model.*Labor* includes family and hired labor in full-time equivalents (fte). *Land* measures the total cultivated land in hectares. Thus, differences in land quality are omitted. We tackle this issue by introducing regional dummies for various agricultural production areas, as defined by the Bavarian Agricultural Research Institute (LfL, 2009). *Intermediate inputs* include all the expenses for forage and crop production (e.g., seed, fertilizer, pesticides, contractors), for animal production (e.g., veterinary, concentrates) and for water, energy, fuel and other expenses linked to production. *Capital* includes the end-of-year value of livestock, buildings, technical facilities and machinery related to agricultural production. Similar input specifications for dairy farms can be found in Kumbhakar and Heshmati (1995), Brümmer et al. (2002) and Newman and Matthews (2006). In addition to the variables of the distance function, we use the farm average value of the variables *cattle livestock unit per ha of forage land* and *yearly milk yield per cow* as concomitant variables in the model for the prior probabilities of the latent classes. These variables most commonly characterize intensive or extensive dairy production systems (Müller-Lindenlauf et al., 2010; Nehring et al., 2011).

5.3. Results

5.3.1 Econometric estimates of the production technology

Estimation results based on 7,999 (1,483) observations for the group of fodder-crop (grassland) farms are presented in tables 5-2 and 5-3. In the case of fodder-crop farms, 87% of the estimated parameters are statistically significantly different from zero at the 5% level (52% for the grassland farms).

Table 5-2: Estimated parameters for the 2 classes of fodder-crop dairy farms ($n = 7,999$)

	Latent Class 1 (intensive)			Latent Class 2 (extensive)		
Parameter	Coefficient	SE		Coefficient	SE	
α_0 (Constant)	0.0188	(0.0225)		-0.0421	(0.0107)	a
β_1 (Labor)	0.1247	(0.0176)	a	0.1746	(0.0200)	a
β_2 (Land)	0.1388	(0.0159)	a	0.1371	(0.0170)	a
β_3 (Interm. Inputs)	0.5046	(0.0140)	a	0.5430	(0.0141)	a
β_4 (Capital)	0.1447	(0.0106)	a	0.1252	(0.0122)	a
β_{11}	-0.1340	(0.0424)	a	-0.1068	(0.0439)	b
β_{22}	-0.1651	(0.0298)	a	-0.2367	(0.0340)	a
β_{33}	0.0742	(0.0246)	a	0.1249	(0.0257)	a
β_{44}	-0.0983	(0.0152)	a	-0.0980	(0.0155)	a
β_{12}	0.1186	(0.0269)	a	0.0584	(0.0295)	b
β_{13}	-0.0300	(0.0255)		0.0492	(0.0246)	b
β_{14}	-0.0164	(0.0196)		-0.1150	(0.0220)	a
β_{23}	-0.0425	(0.0218)	c	-0.0220	(0.0253)	
β_{24}	0.0800	(0.0140)	a	0.1756	(0.0166)	a
β_{34}	0.0385	(0.0159)	b	-0.0993	(0.0144)	a
α_1 (Milk)	-0.7018	(0.0104)	a	-0.7401	(0.0114)	a
α_2 (Other)[1]	*-0.2982*	-		*-0.2599*	-	
α_{11}	-0.1090	(0.0153)	a	-0.0975	(0.0145)	a
α_{22}	*-0.1090*	-		*-0.0975*	-	
α_{12}	*0.1090*	-		*0.0975*	-	
δ_{11}	0.0132	(0.0191)		-0.0363	(0.0228)	
δ_{21}	0.0401	(0.0151)	a	0.0619	(0.0170)	a
δ_{31}	0.0000	(0.0151)		-0.0614	(0.0108)	a
δ_{41}	0.0067	(0.0100)		0.0273	(0.0123)	b
δ_{12}	*-0.0132*	-		*0.0363*	-	
δ_{22}	*-0.0401*	-		*-0.0619*	-	
δ_{32}	*0.0000*	-		*0.0614*	-	
δ_{42}	*-0.0067*	-		*-0.0273*	-	
τ_1 (Trend)	0.0108	(0.0033)	a	-0.0041	(0.0037)	
τ_2	0.0004	(0.0006)		0.0030	(0.0007)	a
θ_{1t}	-0.0025	(0.0031)		-0.0065	(0.0035)	c
θ_{2t}	0.0039	(0.0023)	c	0.0089	(0.0025)	a
θ_{3t}	0.0080	(0.0024)	a	-0.0037	(0.0024)	
θ_{4t}	-0.0048	(0.0018)	a	-0.0021	(0.0019)	
ζ_{1t}	-0.0011	(0.0017)		0.0041	(0.0019)	b
ζ_{2t}	*0.0011*	-		*-0.0041*	-	
$area\_1$	0.2146	(0.0318)	a	0.1864	(0.0168)	a
$area\_2$	0.1978	(0.0108)	a	0.1965	(0.0102)	a
$area\_3$	0.1890	(0.0092)	a	0.1819	(0.0114)	a

$area\_4$	0.1574	(0.0083)	a	0.1263	(0.0112)	a
$area\_5$	0.1835	(0.0182)	a	0.1639	(0.0129)	a
$area\_6$	0.0382	(0.0203)	c	0.1042	(0.0259)	a
$area\_7$	0.0588	(0.0071)	a	0.0530	(0.0082)	a
$area\_8$	0.0831	(0.0092)	a	0.0281	(0.0107)	a
$area\_9$	0.0495	(0.0070)	a	-0.0049	(0.0080)	
$\lambda = \sigma_u / \sigma_v$	0.1208	(0.0072)	a	0.1986	(0.0035)	a
$\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$	0.4615	(0.2569)	c	2.6154	(0.1505)	a
Log-likelihood:	5059.2					

The estimated coefficients of the output distance function are multiplied with -1 to facilitate the comparison with the production frontier.

[1]Italic parameters calculated using the homogeneity restriction

[a,b,c] statistical significance on the 1, 5 and 10% level

Table 5-3: Estimated parameters for the 2 classes of grassland dairy farms ($n = 1,483$)

	Latent Class 1 (intensive)			Latent Class 2 (extensive)		
Parameter	Coefficient	SE		Coefficient	SE	
α_0 (Constant)	0.0916	(0.0209)	a	-0.1973	(0.0396)	a
β_1 (Labor)	0.1591	(0.0340)	a	0.2702	(0.0660)	a
β_2 (Land)	0.2923	(0.0298)	a	0.1949	(0.0488)	a
β_3 (Interm. Inputs)	0.3906	(0.0249)	a	0.3650	(0.0414)	a
β_4 (Capital)	0.0911	(0.0204)	a	0.1292	(0.0315)	a
β_{11}	-0.0256	(0.0861)		-0.1933	(0.1731)	
β_{22}	-0.1656	(0.0851)	c	-0.1409	(0.1164)	
β_{33}	0.0455	(0.0471)		0.0802	(0.0568)	
β_{44}	-0.0945	(0.0332)	a	-0.1791	(0.0385)	a
β_{12}	0.0398	(0.0637)		0.1088	(0.1150)	
β_{13}	0.0427	(0.0521)		-0.1429	(0.0952)	
β_{14}	-0.0298	(0.0366)		0.0281	(0.0539)	
β_{23}	-0.1644	(0.0453)	a	-0.0429	(0.0660)	
β_{24}	0.1933	(0.0391)	a	0.1586	(0.0446)	a
β_{34}	0.0499	(0.0272)	c	-0.0462	(0.0314)	
τ_1 (Trend)	0.0274	(0.0063)	a	0.0092	(0.0095)	
τ_2	-0.0032	(0.0012)	a	-0.0010	(0.0018)	
θ_{1t}	-0.0071	(0.0062)		-0.0208	(0.0105)	b
θ_{2t}	0.0005	(0.0052)		0.0272	(0.0080)	a
θ_{3t}	0.0073	(0.0042)	c	0.0019	(0.0064)	
θ_{4t}	-0.0028	(0.0034)		-0.0176	(0.0046)	a
$area\_1$	-0.0978	(0.0150)	a	0.0541	(0.0311)	c
$area\_2$	-0.0246	(0.0140)	c	0.1781	(0.0324)	a
$\lambda = \sigma_u / \sigma_v$	1.4739	(0.2037)	a	1.9560	(0.3281)	a
$\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$	0.1395	(0.0063)	a	0.1686	(0.0091)	a
Log-likelihood:	1099.6					

[a,b,c] statistical significance on the 1, 5 and 10% level

For the group of fodder-crop farms, the AIC indicates that a model with 2 classes (AIC -1.245) is preferred over a model with 1 class (-0.740). Attempts to estimate a model with 3 or more classes failed to achieve convergence. For the group of grassland farms, the AIC also supports a model with 2 classes against a model with 1 class (AIC of -1.020 and -1.412, respectively). However, for this group, a model with 3 classes and a slightly lower AIC can be estimated, but with very poor results. In particular, the additional class is very small and the distance function coefficients are insignificant with extreme standard errors. In accordance with Greene (2005), who reports a similar case, we prefer a model with 2 classes in both groups. The coefficients for the concomitant variables in the latent class model are shown in table 5-4. The positive signs indicate that a higher value of the variables *milk yield per cow and year* and *cattle livestock unit per ha of forage land* increases the probability of a farm to belong to class 1 in both groups. For the size of the effect and a more detailed interpretation, the marginal effects of the variables on the prior probability have to be calculated. However, the signs of the marginal effects depend on the signs of the coefficients. Because we are not interested in the size of the marginal effect, we check only the estimated coefficients. Using a likelihood ratio test, we examine their statistical significance. For both groups, we find that the variables provide additional information to classify the sample (table 5-4).

Table 5-4: Parameter of the latent class prior probability function

		Fodder-crop			Grassland		
		intensive		extensive	intensive		extensive
		Coef.	SE		Coef.	SE	
Constant	(ρ_0)	-7.563	(0.677)[a]	-	-7.090	(1.662)[a]	-
Yearly milk yield kg/cow	(ρ_1)	1.183	(0.178)[a]	-	2.801	(0.611)[a]	-
Cattle LU/ha forage land[1]	(ρ_2)	0.080	(0.009)[a]	-	0.059	(0.020)[a]	-
LR Test		LR statistic			LR statistic		
$H_0 : \rho_1 = \rho_2 = 0$		186.7 [a]			35.3 [a]		

[a,b,c] statistical significance on the 1, 5 or 10% level

[1]LU = livestock unit

However, it is important to note that the assignment of farms to the classes is only slightly affected by the specification of the prior probability function. Including the concomitant variables (*cattle livestock unit per ha of forage land and yearly milk yield per cow*) in the probability function (see equation (5.8)), we find that only 27 out of the 958 fodder-crop farms (6 out of the 184 grassland farms) were assigned to a different class as compared to the simpler model with no concomitant variables included (equation (5.7)). Hence, we argue that the major effect for sorting the farms into the classes can be attributed to differences in the distance functions' coefficients, representing distinct production technologies. Taking the results from table 5-4 in combination with the key properties of the various classes into account (table 5-5), we conclude that the degree of intensification is at least 1 major distinguishing characteristic of the 2 identified classes in each group. Hence, we refer to the classes as intensive and extensive, but we are fully aware that other interpretations of these results might exist. As depicted in table 5-5, in both groups (henceforth referred to as PGL-I and PGL-E for intensive and extensive permanent grassland and FC-I and FC-E for the respective fodder-crop farms), the so called intensive farms produce relatively more milk on average and utilize an almost equal work force and area under cultivation, respectively.

Table 5-5: Characteristics of identified production systems (class averages)

	Fodder-crop		Grassland	
	intensive	extensive	intensive	extensive
Farms	491	467	116	68
Observations	4,100	3,899	951	532
Labor (fte)[1]	1.60	1.54	1.49	1.48
Land (ha)	46.9	45.5	31.4	30.9
Intermediate inputs (€)	58,602	49,510	31,116	24,833
Cows	41	32	30	22
Milk production (1000 kg/yr)[2]	265.1	185.3	180.9	125.7
Concentrate/Cow (€)	312.3	277.2	259.7	274.1
Cattle LU/ha forage land[3]	2.54	2.18	1.64	1.37
Milk yield (kg/cow yr)	6,432	5,708	6,087	5,693
Vet. cost/Cow (€)	90.3	99.6	76.0	96.8
av. growth rate milk yield (%/yr)	1.80	1.87	1.61	1.41
av. growth rate milk prod. (%/yr)	2.90	2.39	2.61	1.17
No agricult. education (%)	6.7	8.8	9.8	14.1
Basic agricult. education (%)	61.3	68.0	69.5	69.4
Higher agricult. education (%)	32.0	23.2	20.7	16.5
Farmers age (yr)	46.7	48.2	46.9	47.3

[1]fte = full-time equivalent
[2]yr = year
[3]LU = livestock unit

These farms have a higher stocking rate and larger dairy herds, use more intermediate inputs, achieve higher milk yields per cow and have expanded their production faster than the extensive farms. For the PGL farms, the difference in the latter is most striking. The PGL-E farms achieve a yearly growth rate of approximately 1.17%, but the PGL-I farms increased their milk output more than 2.6% per year during the observed period. The FC-I (FC-E) farms expand their milk production at a rate of 2.9% (2.39%) per year. Hence, the group of PGL-E farms lags behind the other 3 groups regarding the expansion of production. FC farms show a higher stocking rate compared to PGL farms. This finding is assigned to the higher forage yields from arable land compared to permanent grassland. Surprisingly and despite the higher milk yields per cow in the intensive classes in both groups, we find higher costs for veterinary services per cow in the extensive classes. Likewise, the costs of concentrate per cow are larger for PGL-E farms compared to intensive ones. Both observations may be explained by better management skills in the intensive classes. This argument is supported by observed educational differences. The highest percentage of farm managers with no agricultural education is found in the PGL-E class (14.1%), and the lowest is found in the FC-I (6.7%) class. Contrariwise, about twice as many farmers (32%) in the FC-I class have received higher agricultural education as those in the PGL-E class (16.5%). For the FC-E (23.2%) and the PGL-I (20.7%) class, the respective figures lie somewhere in between.

To examine the estimated representations of the production technologies for the 4 identified classes, we focus on the first-order production and scale elasticities and the estimated efficiency scores. We calculate the first-order elasticities with respect to inputs and outputs at each data point and present their average values for each class in table 5-6.

Table 5-6: Average elasticities for the 4 identified classes[1]

	Fodder-crop[2]		Grassland	
	intensive	extensive	intensive	extensive
Labor	0.104 (0.047)[a]	0.155 (0.066)[b]	0.122 (0.027)[c]	0.203 (0.080)[d]
Land	0.172 (0.077)[a]	0.177 (0.112)[a]	0.284 (0.099)[c]	0.313 (0.096)[d]
Intermediate inputs	0.543 (0.044)[a]	0.523 (0.062)[b]	0.421 (0.048)[c]	0.370 (0.054)[d]
Capital	0.121 (0.048)[a]	0.136 (0.077)[b]	0.080 (0.067)[c]	0.089 (0.094)[c]
Scale	0.940 (0.037)[a]	0.991 (0.091)[b]	0.907 (0.054)[c]	0.976 (0.083)[d]
Milk output	-0.721 (0.055)[a]	-0.715 (0.052)[b]	-	-
Other output	-0.279 (0.055)[a]	-0.285 (0.052)[b]	-	-

[a-d]Means within a row with different superscripts differ significantly on the 1% level based on a Welch test
[1]Standard deviation in parentheses.
[2]The coefficients of the output distance function are multiplied with -1 to facilitate the comparison with the production frontier. For that reason we find positive (negative) signs for the elasticities on the input- (output-) side.

In all the classes, the average elasticities show the expected signs. As a general pattern, *intermediate inputs* have the highest average output elasticity followed by *land*, *labor* and *capital*. For the FC-I class, the average elasticity of the *intermediate inputs* with respect to output equals 0.543. That is, a 1% increase of this input results in a 0.543% increase of aggregate output. Comparing the elasticities between the 2 groups of farms, we find output to be relatively more responsive to *land* and *labor* for the PGL farms. On the other hand, *intermediate inputs* and *capital* have a higher impact on output for the FC farms. These findings show the relative importance of the input *land* in particular for the PGL dairy farms. This reflects the more pronounced necessity of pure grassland-based farms for additional land to increase their output. On the contrary, FC farms have more alternative means of output expansion because arable land is more versatile. The elasticities on the output side reflect the high share of milk output in the total production of both fodder-crop classes. We check whether the estimated parametric representations of the production technologies satisfy the monotonicity condition at each data point. We find only few violations, primarily for the input capital (table 5-7). However, the percent of violations is still low, and we expect no substantial consistency problems regarding monotonicity.

Table 5-7: Monotonicity violations in percent

	Fodder-crop		Grassland	
	intensive	extensive	intensive	extensive
Labor	1.6	0.8	0.0	0.2
Land	2.0	5.7	0.1	0.0
Intermediate inputs	0.0	0.0	0.0	0.0
Capital	0.4	3.6	12.1	15.9
Milk output	0.0	0.0	-	-
Other output	0.0	0.0	-	-

Returns to scale are slightly decreasing for all the classes at their sample mean. Considering the relatively modest size of the dairy herds in our sample this result is surprising at first sight. However, it is not an uncommon finding in similar research: see for example Emvalomatis (2012) for dairy farms in Germany; Abdulai and Tietje (2007) for northern German dairy farms; Barnes (2008) for dairy farms in Scotland. Moreover Alvarez and Arias (2003) show, that diseconomies of size can occur if farms increase their size but not managerial abilities. We checked the distributions of the scale elasticities and found that they are more widely distributed in the extensive classes, which show a considerable number of observations with decreasing and increasing returns to scale. The distributions for the intensive classes are narrower with only a small number of observations with returns to scale greater than 1.

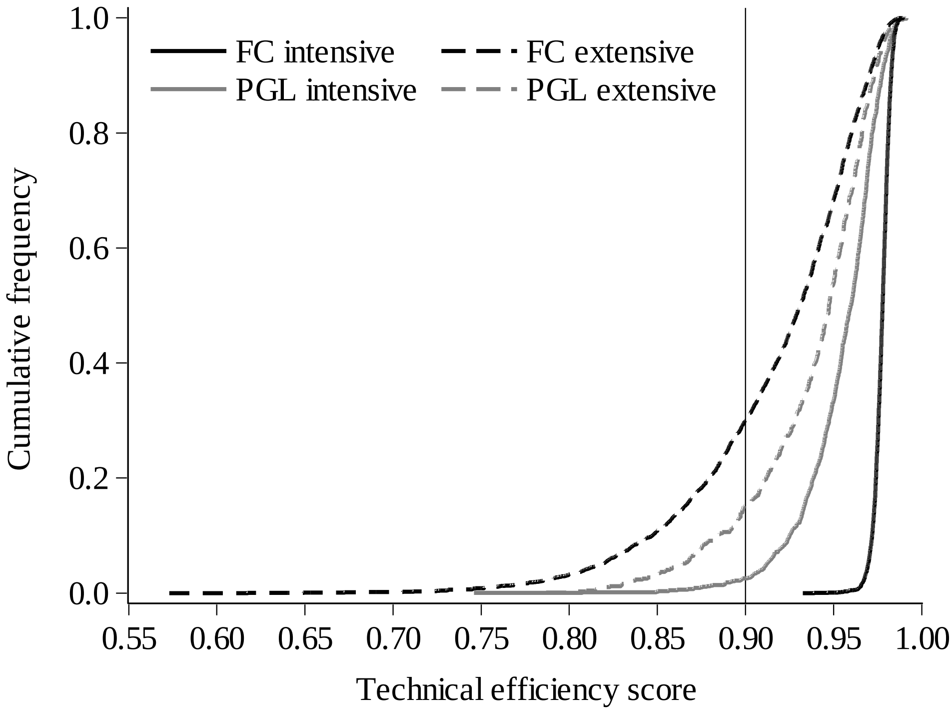
Regarding the coefficients associated with time (all θs and τs in tables 5-2 and 5-3), we find non-neutral technical change in all 4 identified technologies. With only 1 exception (parameter θ_{3t} in the FC-E class), technical change is labor and capital saving and land and intermediate inputs using (Morrison Paul et al., 2000). However, the coefficients vary in value and significance across all the classes. For instance, technical change for the PGL-I class is almost exclusively disembodied, but we do not find significant disembodied technical change for the extensive class. This implies that the contribution of technical change to total factor productivity for the PGL-E class stems mainly from changes in the marginal productivities of single inputs over time. We test whether the presence of non-neutral technical change can be rejected for the class of PGL-I farms. The Wald statistic for the hypothesis $H_0: \theta_{1t} = \theta_{2t} = \theta_{3t} = \theta_{4t} = 0$ is equal to $\chi^2(4) = 3.913 < \chi^2_{0.95} = 9.487$. Hence, the coefficients for non-neutral technical change are not significantly different from zero. Because we do not face restrictions in degrees of freedom, we still keep the respective variables in the model to preserve a model specification consistent with the other classes.

5.3.2 Technical efficiency

As noted above, the technical efficiency scores we report are measured against the most likely frontier for each farm. Hence, this approach does not take the uncertainty of class membership into account and is in the line of thought by Greene (2007a), who favors the parameter associated with the most likely class as the best estimator for a "unit specific" parameter in a latent class model. However, the uncertainty in the class assignment of the farms is very low. This is reflected by average a posteriori probabilities of class membership of 0.983 for the FC farms and 0.982 for PGL farms. Consequently, we find almost identical results, if the technical efficiency scores are calculated as a weighted average using equation (5.15).

The average technical efficiency scores for the intensive and extensive FC classes are 0.977 and 0.917, respectively. The average values for the PGL classes are 0.954 (intensive) and 0.937 (extensive). Hence, the farms in the intensive classes are on average more efficient and, as a more homogeneous class, closer to their own frontier for both groups. The cumulative distributions of efficiency scores in figure 5-1 show that most of the farms are highly efficient in all the classes. However, especially in the extensive classes, inefficient farms exist, which drag down the class's average efficiency. For the FC-E and the PGL-E classes, we find that 31.1% and 15.4% of the farms exhibit efficiency scores below 0.9, respectively. In contrast, only 2.6% of the PGL-I and none of the FC-I farms show efficiency scores below 0.9. In figure 5-1, technical efficiency of 0.9 is indicated by the vertical grid line. Cabrera et al. (2010) also identify a positive relationship between intensification and technical efficiency for dairy farms by using an inefficiency effects model. Alvarez and del Corral (2010) report similar results, and they ascribe these findings to the assumption that intensive systems might be easier to manage and farmers are less prone to make mistakes. The shape of the distributions in figure 1 supports this interpretation.

Figure 5-1: Empirical cumulative distributions of technical efficiency scores.



Differences in the managerial quality could be another interpretation. This is supported by the aforementioned differences in veterinary costs and the farmers' agricultural education. Putting the estimated efficiency scores for the 4 groups in context with the rich literature on dairy farm efficiency, we find the class-average scores to be relatively high (cf. Bravo-Ureta et al., 2007). We attribute this finding mainly to the methodological design of the study, which mitigates the confusion of technological heterogeneity with technical inefficiency (see also Alvarez and del Corral, 2010).

5.3.3 Productivity analysis

The results for the calculation of the Malmquist index of TFP change and its decomposition are summarized in table 5-8. Technical change accounts for the major part of TFP change in all the classes. Except for the extensive PGL farms, the effect of technical efficiency change is only very small. We find negative scale change effects for both the fodder-crop and the PGL-E farms and a positive effect for the PGL-I farms. Again, except for the FC-E farms, the impact of the scale effect on TFP growth is negligible.

Table 5-8: Malmquist index of total factor productivity (TFP) change for the 4 identified classes

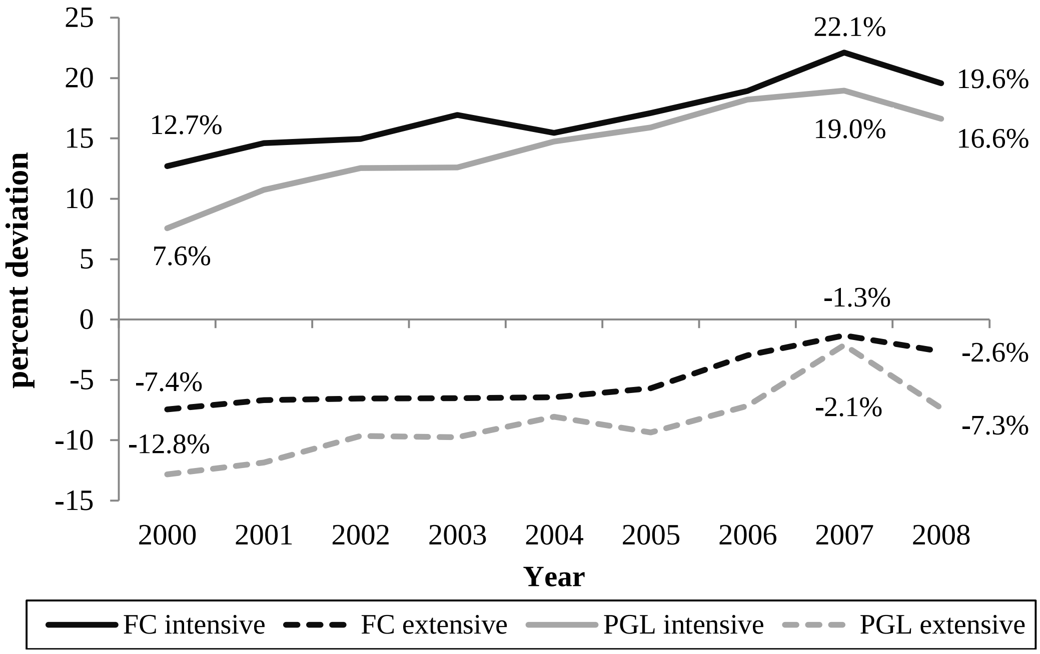
	Fodder-crop				Grassland			
	intensive		extensive		intensive		extensive	
	ΔTFP[1]	cumul.[2]	Δ TFP	cumul.	Δ TFP	cumul.	Δ TFP	cumul.
		0.00		0.00		0.00		0.00
2000/01	1.06	1.06	0.62	0.62	2.60	2.60	0.78	0.78
2001/02	0.90	1.97	0.34	0.97	1.81	4.40	1.21	1.99
2002/03	1.16	3.13	0.46	1.42	1.21	5.61	0.45	2.44
2003/04	0.99	4.12	0.67	2.10	1.59	7.20	0.98	3.42
2004/05	1.42	5.54	1.33	3.43	1.09	8.29	-0.24	3.19
2005/06	1.41	6.95	1.73	5.16	1.05	9.34	0.83	4.01
2006/07	1.58	8.53	1.74	6.90	0.55	9.89	1.26	5.28
2007/08	1.37	9.90	0.83	7.73	-0.48	9.41	-1.06	4.22
	Average annual change rate (%)							
$TFPC$	1.24		0.97		1.18		0.53	
TC	1.27		1.10		1.18		0.66	
TEC	-0.01		-0.02		-0.03		-0.13	
SC	-0.03		-0.12		0.03		-0.01	

[1]ΔTFP = class average percentage TFP change between the indicated years
[2]cumul. = cumulative percentage TFP change

Table 5-8 also shows year-to-year TFP changes for the 4 classes and the cumulative growth from 2000 to 2008. We observe a steady TFP growth for the FC-I farms with slightly increased rates after 2004. The class of FC-E farms exhibits rather low TFP growth in the first years, but the growth rates increase from 2004 on and catch up to those of the FC-I class. The PGL-I class shows the highest growth rates in the first years, but the rates decrease over time; we even find a decrease in productivity from year 2007 to 2008. The average TFP change rates for the PGL-E farms are unsteady. In the FC group the intensive (extensive) farms reach an average annual rate of 1.24% (0.97%). The PGL farms improve their productivity on average at an annual change rate of 1.18% for the intensive class but only 0.53% for the extensive class. These rates result in a cumulative productivity growth during the observed period of 9.9% and 7.7%, respectively, for the intensive and extensive FC classes and 9.4% and 4.2% for the respective PGL classes.

Turning now to the results of the multilateral Törnqvist index, we check for differences in the productivity levels of the 4 identified classes. Figure 5-2 summarizes the main results. Each line represents the percentage deviation of the 4 groups' hypothetical reference farm from the productivity of the overall reference farm in the base year and its evolution over time. For example, the FC-I class reaches an average productivity level of 22.1% in the year 2007. This means that the average farm in the FC-I class produces 22.1% more output using a given amount of inputs in 2007 than the overall average reference farm in the year 2000.

Figure 5-2: Percentage deviation of the class reference unit's total factor productivity (TFP) from the overall average TFP in the base year 2000



Starting in the base year 2000, we find the FC-I class to be 12.7% more productive than the overall reference farm. The PGL-I class also lies well above the overall reference (7.6%). In contrast, the extensive classes are consequently underperforming (-12.8% and -7.4%). The ranking of the 4 groups does not change during the observed period. However, the differences between the intensive and extensive classes increase over time. Although the difference between the intensive and extensive FC (PGL) farms is 20.1% (20.4%) in 2000, it is 22.2% (23.9%) in 2008.

We note that a direct comparison of the results of the Malmquist index of TFP growth in table 5-8 and the multilateral Törnqvist index in figure 5-2 should be exercised only cautiously. First, the indices have different reference points. The Malmquist index measures individual productivity growth rates from one period to the next, but the multilateral Törnqvist index measures the productivity of each farm relative to the average farm in the first year. Second, the Malmquist index allows for non-constant returns to scale and inefficiency. For the calculation of the multilateral Törnqvist index, we assume constant returns to scale and full efficiency. However, considering our results that inefficiency and non-constant returns to scale have only a minor effect, we are not particularly concerned about these assumptions. Third, the Malmquist index is based on econometrically estimated distance functions. Hence, statistical noise in the data is taken into account in the error term v_{it} and eliminated from the calculations. The estimations are performed for all the years of the panel in one step with a time trend included to allow for non-neutral technical change. This leads to a much smoother productivity change measure than the more flexible Törnqvist index, where we assume the data to be free of any errors in measurement, reporting or specification and calculate discrete measures for every year (Coelli et al., 2005). The latter aspect explains why the productivity pathways obtained from the two different indices begin to differ substantially in the years 2007 and 2008, especially for the FC classes. The years in question can be considered to be extreme events for the dairy sector due to the erratic price movements that took place on the input and the output side.

5.4. Discussion

The key results of the study can be summarized as follows. With respect to their technology's own frontier, most farmers in all the classes are highly technically efficient. However, we find substantial differences by comparing the productivity of the 4 classes. The intensive classes in both groups of farms are more productive and are also able to increase their productivity to a greater extent over the observed period. Technical progress is by far the most important component of TFP growth for all the classes. Our calculations of the multilateral Törnqvist index reveal that both the FC classes are more productive than their PGL counterparts. However, our distinction between intensive and extensive classes shows that there are highly productive grassland farms that can keep up with the fodder-crop farms and that those farms are predominantly found in the more intensive class of farms. The more substantial productivity gap exists between the intensive and extensive classes, and this gap widens because of the higher productivity growth rates in the intensive classes. To some extent, this finding questions payments particularly targeted to grassland farms.

There is some evidence that being in the group of intensive or extensive farms is at least to some extent not an intentional decision made by the farmer but is also related to his/her management skills. In both of the intensive groups, farmers are on average better educated and have to spend less on veterinary services. The management skills are also reflected in the technical efficiency scores. The farmers in the intensive classes produce very close to their frontier. This indicates hardly any errors in management. The abolishment of the milk quota system in 2015 might change the situation to some extent. Production controls usually limit technical change and one could argue that this helped permanent grassland farmers to keep up with their advantaged counterparts. Given that the PGL-E class has the lowest productivity and productivity growth rates, this suggests a shift toward more intensive production technologies in permanent grassland regions. This could occur either because extensive farms are more prone to leave the market or because farmers manage to steadily increase the intensity of their production. However, regarding the aforementioned additional benefits (biodiversity, carbon storage, landscape) of cultivated permanent grassland, we have to stress that these benefits predominantly apply to a more extensive form of cultivation (Poetsch, 2007). Müller-Lindenlauf et al. (2010), who identify the same 4 groups of dairy farms (intensive vs. extensive; grassland vs. fodder-crop), note that intensive farm types tend to be advantageous in global categories, such as climate impact and land demand and the low input farm types have advantages in local categories, such as the emission of ammonia, animal welfare and milk quality. Therefore, on the one hand, stimulation of further intensification could help to improve productivity in the dairy sector, especially in permanent grassland regions. On the other hand, exactly where we identify the greatest potential for productivity gains, further intensification could be counterproductive for the provision of local external benefits.

Our analysis stresses the importance of looking at both the levels of productivity and productivity changes when comparing various groups. To do so, we provide a group- and chain-linked multilateral productivity index. We consider this index to be particularly useful for empirical applications that combine inter- and intra-group comparisons over time. Based on firm-level data, it can provide a convenient measure of the productivity variation among firms in a group and between the groups over time. A natural application would be the comparisons of industries in various countries or regions. Other examples are conventional versus organic production and large versus small operations. However, in this context, we have to point toward the weighting adjustments that could become necessary to obtain adequate results and to take the potential differences between aggregated productivity, average productivity and the productivity of the average production unit into account.

The research of Kumbhakar et al. (2009), Mayen et al. (2010) and Moreira and Bravo-Ureta (2010) is closely related to, but to some extent different from our framework. Mayen et al. (2010) use a propensity score matching approach to compare the efficiency and productivity of organic and conventional dairy farms in the US. Hence, they directly compare the efficiency and productivity of pairs of conventional and organic farms, which are completely identical except for their technology under ideal circumstances. This has 2 advantages: first, it accounts for a potential self-selection bias from endogenous technology choice; second, the measured differences are directly and solely attributable to the difference in technology. However, the problem we face here is somewhat different. In our case, technology is not endogenous but rather is given by the exogenous location of the farm. Hence, a self-selection problem is not present. Moreover, because our question is whether grassland farms can compete, we are not interested in the differences in productivity attributable only to technology but rather in the overall difference. For example, grassland farms in mountainous areas may be smaller than their counterparts because access to land is limited given the topology. Observed differences in efficiency and productivity might originate from differences in technology or differences in size. The additional scope for diversification might be an additional advantage for the group of fodder-crop farms. Kumbhakar et al. (2009) compare the performance of organic and conventional dairy farms in Finland. They account for a potential self-selection bias by jointly estimating production technologies and technology choice decisions. Subsequently, they calculate predicted outputs for each farm utilizing the observed input mix and estimated parameters of the alternative production technologies. They interpret the gap between each farm’s predicted outputs from various production technologies given observed inputs as the difference in productivity between the respective technologies. Moreira and Bravo-Ureta (2010) analyze the technical efficiency of dairy farms in 3 South American countries, whereby they use the metafrontier concept described in O’Donnell et al. (2008) to compare efficiency scores across various technologies. This involves the estimation of a specific frontier for each technology group and a metafrontier that envelops all the group frontiers. The efficiency of each farm is then measured against its own group frontier, and the group frontiers are subsequently compared to the metafrontier. The approaches used by Kumbhakar et al. (2009) and Moreira and Bravo-Ureta (2010) measure the potential productivity of a farm if it had access to a different technology but keeps its initial input mix fixed. However, it is likely that switching to an alternative production technology also implies changes in the input mix.

6. DISCUSSION

In this thesis, I tackle methodological and empirical questions in regard to the measurement of firm's productivity and efficiency. To do so, I use a frontier approach, which allows for a firm's shortfall from its potential best practice production technology. Empirically, this best practice production technology is estimated econometrically in a parametric, stochastic framework based on firm-level data. As a main contribution, I particularly pay attention to various manifestations of observed and unobserved forms of firm heterogeneity that can hamper the correct identification of production technologies, productivity and efficiency scores.

In the first part of the thesis, I briefly review the origins and the theoretical background as well as parts of the huge literature in the field of productivity and efficiency analysis. I also introduce the key properties of the applied methodologies. In addition, I discuss some of the issues evoked through unaccounted heterogeneity which motivated my work and comment shortly on the apparent ambiguity that complicates a clear-cut definition of heterogeneity and isolating it from measures of productivity and efficiency.

In the second chapter, entitled *“Total factor productivity decomposition and unobserved heterogeneity in stochastic frontier models”*, I examine how the decomposition of total factor productivity growth is affected by the choice of the econometric specification, which is used to estimate the underlying production technology. I concentrate on nine stochastic frontier models, of which eight have been widely applied in empirical TFP growth studies; in addition, I introduce a new model by combining the GLS random effects stochastic frontier model, which allows for firm-specific time varying efficiency (Cornwell et al., 1990) with the Mundlak adjustment to account for unobserved heterogeneity, as proposed by Farsi et al. (2005a). I discuss the properties of all nine models and pay particular attention to their potential to take unobserved heterogeneity into account. For the empirical application, I use an unbalanced panel dataset of 974 dairy farms and an output-oriented translog distance function to represent the production technology. To obtain comparable results of the nine different models, I keep the data, the specification of variables and the functional form idential across all econometric specifications. The values and rankings of the estimated efficiency scores are found to be quite sensitive to model specification. This result was expected; it reflects the different abilities of the models to take unobserved heterogeneity into account and is in line with findings by Farsi et al. (2005; 2005a), Greene (2005), Abdulai and Tietje (2007) and Filippini et al. (2008). In contrast to the prevailing works in the stochastic frontier literature, I elaborate not only on the efficiency scores but also on the estimated representation of the production technology itself. Then, based on the corresponding production elasticities, the returns to scale measures and inefficiency estimates, I measure productivity growth and decompose it into the three components that are most commonly found in empirical applications: technical change, changes in technical efficiency and the scale change effect. The results of the decomposition show that the relative contribution of the components to TFP growth is quite sensitive to the choice of the econometric model. The average growth rates of technical change vary in a range between 1.64% and 1.19%. For technical efficiency change (scale change effect), average growth rates lie in a range between -0.06% and -0.51% (-0.02% and -0.13%).

The fact that different econometric models, which impose different assumptions on the data and the data-generating process, lead to different results is obvious. The problem arises if we cannot choose reliably among the models. Many of the stochastic frontier models are not nested; hence, standard statistical testing cannot reveal the “one” correct model for any dataset. I apply a range of statistical tests to reduce the number of applicable models and offer further criteria from which to choose between the non-nested models. These are based on the requirements implied by microeconomic theory, namely monotonicity and quasi-concavity in inputs and concavity in outputs, as well as the properties of the dataset and the sector under consideration. Despite the fact that these criteria can be simply applied by empirical researchers, further research should focus on the application of specification test procedures for non-nested models in the stochastic frontier context (see Lai and Huang (2010) for a recent example).

In the third chapter *“Accounting for endogenous effects in stochastic frontier models”*, we elaborate an alternative modeling approach for observed and unobserved heterogeneity. We further develop the line of thought by Farsi et al. (2005) and provide a general framework that allows taking heterogeneity in stochastic frontier models into account. Following Mundlak (1978), we use the group means of the input variables to construct a model of the unobserved heterogeneity. This improves econometricians’ ability to account for heterogeneity that is unobserved to them but not to producers, who adjust their input decisions conditional on their production conditions. Extensions of this heterogeneity model include the incorporation of environmental variables and the notion that not all group-mean variables are necessarily correlated with firm heterogeneity. We adapt this extended modeling of heterogeneity to three different stochastic frontier models: the so-called "true" random-effects model (Greene, 2005) with stochastically time-varying efficiency; a MLE random-effects model (Battese and Coelli, 1992) with deterministically time-varying technical efficiency; and the GLS random effects model (Schmidt and Sickles, 1984) with time-invariant technical efficiency. The results of our empirical application support the proposed specifications. The heterogeneity bias in the technology parameters is reduced in the sense that the coefficients of the input factors come closer to those of the conventional fixed-effects panel model. The predicted firm effects and the efficiency scores we obtain from the different specifications also meet our expectations. Compared to the basic specifications of the GLS and MLE model, the proposed extensions help to reduce the downward-bias in the efficiency scores. We don't find this effect on the efficiency scores in the TRE specifications in which the random constant already captures firm effects. These results are in line with findings by Farsi et al. (2005, 2005a) and Abdulai and Tietje (2007). For the predicted firm effects, we find similar distributions across the different specifications, as well as fairly strong positive correlations between them. These findings indicate that the proposed specifications to model heterogeneity can serve as an alternative to the "true" effects models.

In the fourth chapter *"Decomposing labor productivity growth: the case of small and medium-sized breweries in Germany"* we investigate the evolution of labor productivity in the German brewing industry. As a global exception, the brewing industry in Germany is still dominated by relatively small firms and the market concentration is comparably low. Nevertheless, the German brewing industry has faced considerable structural changes in the last two decades, mainly due to a substantial reduction in the domestic beer consumption. We focus on small- and medium-sized breweries, many of which are regional, family owned and operated businesses with a long-standing tradition.

In this chapter we add to the literature by tackling the micro-macro linkage of productivity growth, one of the "areas of future work" as put forward by Bartelsman and Doms (2000, p. 592). In particular we combine two strands of the literature on (labor) productivity decomposition. One strand originates from empirical studies that use micro-data to describe the productivity growth dynamics of a sector as in Baily et al. (1992; 2001), Griliches and Regev (1995) and Foster et al. (2001). The other strand combines index number theory with stochastic frontier analysis (e.g., Nishimizu and Page, 1982; Bauer, 1990) and decomposes firm-specific productivity growth into several components.

We show how the two approaches can be combined to analyze the dynamics of aggregated industry labor productivity in great detail. Based on this procedure, we can decompose industry labor productivity change into seven components: input deepening, technical change, technical efficiency, a scale change effect, between-firm reallocation and the effects from exits and entry. Given the nature of our data, we further decompose the net entry term to distinguish the effect associated with firms that enter/exit the industry (and therefore the sample) from that of firms that drop in and out of the sample for other unknown reasons. In order to account for unobserved heterogeneity in the econometric estimation of the production technology we apply an augmented stochastic frontier model, as proposed in chapter three. We also factor in output price heterogeneity between the firms in the sample by constructing a firm specific price index based on the physical output and the revenues of each brewery.

Our results show that productivity growth within the breweries has a much higher impact on the sectoral productivity, compared to the effects from industry dynamics. From 1996 to 2008 labor productivity in our sample increased by 7.65% due to improvements happening within the firms. The reallocation of workforce towards more productive firms and the shutdown of underperforming firms increased labor productivity by 3.99%. These findings are in line with Baily et al. (1996) and Griliches and Regev (1995) how also find a high share of aggregate labor productivity growth (0.79 and 0.83, respectively) due to improvements within the firms (see table 8.1 in Foster et al. (2001) for further examples). Disentangling the within firm effect, technical change demonstrates to be the main driver of firms' productivity growth. Technical efficiency, however, decreases over time. Hence, although the production frontier is shifted upwards through technical change, not all firms are able to follow this development. Accordingly, the performance of the firms in the sector diverges and the average technical efficiency decreases. All these findings together fit well into a sector that mainly consists of mid-sized, family owned businesses with a long tradition. These family businesses appear to be "squeezed in" between the very large breweries which aim towards cost-leadership and/or invest in substantial marketing activities and the flexible and innovative micro-breweries, which produce for regional niche-markets. On average the mid-sized businesses stay rather passive and either try to defend their market shares by becoming more productive through investments in technology or continue producing with the old technology as long as possible.

Although we cover several issues in the design and implementation of our study, it has some limitations. Generally decreasing demand for beer in combination with sunk capital, which can hardly be adjusted (brewing kettles, filling lines, cold storage), challenge the elementary assumption that firms utilize their production factors at full capacity (see e.g., Jorgenson and Griliches, 1967). We were unable to obtain firm-specific measures of the total capacity of the breweries due to the anonymity of the firms. However, including the total demand for beer could have served as an ad-hoc approximation. Baldwin et al. (2013) in a similar application build on a suggestion by Berndt and Fuss (1986) to account for variations in capital utilization through the share of capital income in the total value added. This strategy remains to be explored. A second limitation concerns the functional specification of time-varying inefficiency in our econometric model (Battese and Coelli, 1992). Despite its widespread application in numerous empirical studies on productivity change, it can only depict monotonous patterns of efficiency change. Low levels of efficiency are connected to high rates of efficiency change, subject to a single parameter η which is common to all firms. The comparative analysis of different specifications in chapter 2 shows that the decomposition of productivity growth can be quite sensitive to the choice of the econometric model (see also Karagiannis and Tzouvelekas, 2009). Third, we use the complete sample for the econometric estimation of the technology parameters, which are then used to perform the decomposition of productivity growth. However, only the “continuing” firms contribute to the calculation of the within firm effect between two time periods. This mismatch could lead to a bias in the technology parameters, if the group of continuing firms uses a separate production technology. A potential solution to this issue could be the concept of a bilateral production frontier as proposed by Karagiannis et al. (2011). Finally, we use a firm specific price index to construct an output measure which takes price dispersion on the output side into account. On the input side, however, we use price indices from the German Federal Statistical Office (Destatis) to account for price effects over time. In case of the input labor, we use the wage bill to factor in quality differences in the workforce. Apart from quality differences, breweries located near urban centers will probably have to pay higher wages for employees with equivalent qualifications compared to breweries in structurally weak regions. One could also argue that large firms have more bargaining power compared to small firms and therefore buy intermediate inputs at a cheaper rate. In both cases, firms with lower (higher) costs due to input price dispersion may appear to be more (less) productive.

In the fifth chapter *“Dairy farming on permanent grassland: can they keep up”* we perform an analysis of the productive performance of Bavarian dairy farms and focus thereby in particular on differences in the production technologies. We distinguish two types of dairy farms; one type operates solely on permanent grassland whereas the other uses fodder crops from arable land. In both groups we allow for further, unobserved heterogeneity in the production technology by applying a latent class stochastic frontier model. As a result we identify more intensive and extensive production systems for both types of farms, whereby we base our notion of intensive vs. extensive dairy production on differences in stocking density and milk yield per cow and year. To assess the productive performance of the different groups of farms, we analyze the levels and growth rates of total factor productivity. We use the well-established generalized Malmquist index as suggested by Orea (2002) to analyze productivity growth; to compare the productivity levels of the different groups of farms we develop a group- and chain-linked multilateral productivity index. Our index is based on the multilateral TFP index introduced by Caves et al. (1982a). We adapt their index, however, by combining approaches by Good et al. (1997) and Delgado et al. (2002) to analyze the productivity level of groups that consist of subordinate (micro-level) units over time.

Our results show that most farms are highly technically efficient with respect to their own technology frontier, but we find substantial differences in their productivity levels. The intensively producing farms in both, the grassland and the fodder-crop group, are more productive and are also able to increase their productivity to a greater extent over the observed period. Starting in the base year 2000, we find the class of intensive fodder-crop farms to be on average 12.7% more productive than the overall reference farm. The intensive grassland farms also lies well above the overall reference (7.6%). In contrast, the extensive classes are on average consequently underperforming (-12.8% for the grassland and -7.4% for the fodder-crop farms). The ranking of the 4 groups does not change during the observed period, but we find that the differences between the intensive and extensive classes increase over time.

Despite our efforts to whittle down the effects of unobserved heterogeneity in this empirical study, some remaining differences in the production conditions cannot be ruled out. In order to account for heterogeneous land quality and topographical effects we include regional dummies for various agricultural production areas. However, even these well-defined areas can only be considered as a rough approximation. A promising solution would be the use of a detailed land quality index, which was in principle included in our dataset, but revealed some unexplainable inconsistencies. Another issue, is the use of two distinct productivity indices. A direct comparison of the results of the Malmquist index of

TFP growth and the multilateral Törnqvist index should be exercised only cautiously as the indices have different reference points. While the Malmquist index measures individual productivity growth rates from one period to the next, the multilateral Törnqvist index measures the productivity of each firm relative to the average firm in the first year. In addition, the Malmquist index allows for non-constant returns to scale and inefficiency. For the calculation of the multilateral Törnqvist index, however, we assume constant returns to scale and full efficiency. Considering our results that inefficiency and non-constant returns to scale have only a minor effect, we are not particularly concerned about those additional assumptions. Third, the Malmquist index is based on econometrically estimated distance functions. Hence, statistical noise in the data is taken into account and eliminated from the calculations. The estimations are performed for all the years of the panel in one step with a time trend included to allow for non-neutral technical change. This leads to a much smoother productivity change measure than the more flexible Törnqvist index, where we assume the data to be free of any errors in measurement, reporting or specification, and calculate discrete measures for every year.

The thesis highlights the significance of firm heterogeneity in micro-level datasets; in particular I emphasize the necessity to take observed and unobserved differences in the firms and their technologies into account to obtain unbiased measures of efficiency and productivity (growth). I discuss various methodological and empirical questions in this respect and hope to contribute to a clear presentation of the challenges and some strategies to tackle them. An important field of future research, which deserves further attention, is the incorporation of approaches to account for capacity utilization and imperfect competition in the parametric frontier framework. Some effort has been made by the author in this respect, attempting to identify the effects of markups on total factor productivity growth and to distinguish it from other components, e.g., technical change, efficiency change and scale effects (Karagiannis et al., 2013).

7. REFERENCES

Abbott, T.A. 1991. Producer price dispersion, real output and the analysis of production. *Journal of Productivity Analysis* 2:179–195.

Abdulai, A., and H. Tietje. 2007. Estimating technical efficiency under unobserved heterogeneity with stochastic frontier models: application to northern German dairy farms. *European Review of Agricultural Economics* 34(3):393–416.

Abramovitz, M. 1956. Resource and output trends in the United States since 1870. *American Economic Review* 46(2):5–23.

Abrate, G., F. Erbetta, and G. Fraquelli. 2011. Public utility planning and cost efficiency in a decentralized regulation context: the case of the Italian integrated water service. *Journal of Productivity Analysis* 35:227–242.

Aiello, F., C. Mastromarco, and A. Zago. 2011. Be productive or face decline. On the sources and determinants of output growth in Italian manufacturing firms. *Empirical Economics* 41(3):787–815.

Aigner, D.J., and S.F. Chu. 1968. On estimating the industry production function. *American Economic Review* 58:568–598.

Aigner, D.J., C.A.K. Lovell, and P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *Journal Econometrics* 6(1):21–37.

Alchian, A.A., and R.A. Kessel. 1962. Competition, monopoly, and the pursuit of money. In: *Aspects of labor economics*. NBER out-of-print volume. Princeton, N.J.: Princeton University Press.

Alvarez, A., C. Amsler, L. Orea, and P. Schmidt. 2006. Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. *Journal of Productivity Analysis* 25:201–212.

Alvarez, A., and C. Arias. 2003. Diseconomies of size with fixed managerial ability in dairy farms. *American Journal of Agricultural Economics* 85:136–44.

Alvarez, A., and J. del Corral. 2010. Identifying different technologies using a latent class model: Extensive versus intensive dairy farms. *European Review of Agricultural Economics* 37:231–250.

Alvarez, A., L.W. Tauer, and J. del Corral. 2012. Modeling unobserved heterogeneity in New York dairy farms: one-stage versus two-stage models. *Agricultural and Resource Economics Review* 41:1–11.

Arrow, K.T., H.B. Chenery, B.S. Minhas, and R.M. Solow. 1961. Capital-labor substitution and economics efficiency. *The Review of Economics and Statistics* 43(3):225–250.

Atkinson, A.B., and J.E. Stiglitz. 1969. A new view of technological change. *The Economic Journal* 79(315):573–578.

Baily, M.N., C. Hulten, D. Campbell. 1992. Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity: Microeconomics* 4:187–267.

Baily, M.N., E.J. Bartelsman, and J. Haltiwanger. 1996. Downsizing and productivity growth: myth or reality? *Small Business Economics* 8:259–278.

Baily, M.N., E.J. Bartelsman, and J. Haltiwanger. 2001. Labor productivity: structural change and cyclical dynamics. *The Review of Economics and Statistics* 83(3):420–433.

Baldwin, J.R., W. Gu, and B. Yan. 2013. Export growth, capacity utilization, and productivity growth: evidence from the Canadian manufacturing plants. *The Review of Income and Wealth* 59(4):665–688.

Balk, B. 2001. Scale efficiency and productivity change. *Journal of Productivity Analysis* 15:159–183.

Balk, B. 2003. The residual: on monitoring and benchmarking firms, industries, and economies with respect to productivity. *Journal of Productivity Analysis* 20:5–47.

Baltagi, B. 1995. Econometric analysis of panel data. New York: Wiley and Sons.

Baltagi, B., and J.M. Griffin. 1988. A general index of technical change. *Journal of Political Economy* 96(1):20–41.

Baltagi, B., and O. Li. 1990. A Lagrange multiplier test or the error components model with incomplete panels. *Econometric Reviews* 9(1):103–107.

Barnes, A. 2008. Technical efficiency estimates of Scottish agriculture: A note. *Journal of Agricultural Economics* 59:370–376.

Bartelsman, E.J., and M. Doms. 2000. Understanding productivity: lessons from longitudinal microdata. *Journal of Economics Literature* 38(3):569–594.

Bartelsman, E.J., and P. Dhrymes. 1998. Productivity dynamics: U.S. manufacturing, plants 1972–1986. *Journal of Productivity Analysis* 9(1):5–34.

Bartelsman, E.J., J. Haltiwanger, S. Scarpetta. 2009. Measuring and analyzing cross-country differences in firm dynamics. In: T. Dunne, J.B. Jensen, and M.J. Roberts. eds., *Producer dynamics: new evidence from micro data*. Chicago: University of Chicago Press.

Bartelsman, E.J., J. Haltiwanger, S. Scarpetta. 2013. Cross-country differences in productivity: the role of allocation and selection. *American Economic Review* 103(1):305–334.

Basu, S., and J.G. Fernald. 2002. Aggregate productivity and aggregate technology. *European Economics Review* 46:963–991.

Battese, G.E., and T.J. Coelli. 1992. Frontier production functions, technical efficiency and panel data. With application to paddy farmers in India. *Journal of Productivity Analysis* 3(1-2):153–169.

Battese, G.E., and T.J. Coelli. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20(2):325–32.

Battese, G.E., and G. Corra. 1977. Estimation of a production frontier model: with application to the pastoral zone of eastern Australia. *Australian Journal of Agricultural Economics* 21(3):169–179.

Bauer, P.W. 1990. Decomposing TFP growth in the presence of cost inefficiency, nonconstant returns to scale, and technological progress. *Journal of Productivity Analysis* 1(4):287–99.

Bayarsaihan, T., and T.J. Coelli. 2003. Productivity growth in pre-1990 Mongolian agriculture: spiraling disaster or emerging success? *Agricultural Economics* 28(2):121–137.

Berkhout, B., L. Bertling, Y. Bleeker, W. de Wit, G. Kruis, R. Stokkel, and R. Theuws. 2013. The contribution made by beer to the European economy: full report. The Brewers of Europe, Amsterdam.

Berndt, E.R., and M.A. Fuss. 1986. Productivity measurement with adjustment for variation in capacity utilization and other forms of temporal equilibrium. *Journal of Econometrics* 33(1):7–29.

Berndt, E.R., and M.S. Khaled. 1979. Parametric productivity measurement and choice among flexible functional forms. *Journal of Political Economy* 87(6):1220–1245.

Blundell, R., and S. Bond. 2000. GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews* 19(3):321–340.

Bravo-Ureta, B.E., D. Solis, V. Moreira, J. Maripani, A. Thiam and T. Rivas. 2007. Technical efficiency in farming: a meta-regression analysis. *Journal of Productivity Analysis* 27:57–72.

Breustedt, G., U. Latacz-Lohmann, T. Tiedemann. 2011. Organic or conventional? Optimal dairy farming technology under the EU milk quota system and organic subsidies. *Food Policy*. 36:223–229.

Brown, M., and J. Popkin. 1962. A measure of technological change and returns to scale. *Review of Economics and Statistics* 44(4):402–411.

Brümmer, B., T. Glauben, and G.J. Thijssen. 2002. Decomposition of productivity growth using distance functions: the case of dairy farms in three European countries. *American Journal of Agricultural Economics* 84(3):628–644.

Cabrera, V.E., D. Solis, and J. del Corral. 2010. Determinants of technical efficiency among dairy farms in Wisconsin. *Journal of Dairy Science* 93:387–393.

Caves, D.W., L.R. Christensen, and W.E. Diewert. 1982. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50(6):1393–1414.

Caves, D.W., L.R. Christensen, and W.E. Diewert. 1982a. Multilateral comparisons of output, input, and productivity using superlative index numbers. *Economic Journal* 92:73–86.

Chamberlain, G., 1982. Multivariate regression models for panel data. *Journal of Econometrics* 18:5–46.

Chambers, R.G. 1988. Applied production analysis: a dual approach. New York: Cambridge University Press.

Chan, M.W.L, and D.C. Mountain. 1983. Economies of scale and the Törnqvist discrete measure of productivity growth. *Review of Economics and Statistics* 65:663–667.

Charnes, A., W.W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision-making units. *European Journal of Operations Research* 2(6):429–444.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau. 1973. Transcendental logarithmic production frontiers. *The Review of Economics and Statistics* 55(1):28–45.

Coelli, T.J. 2000. On the econometric estimation of the distance function representation of a production technology. Discussion Paper 2000/42, Center for Operations Research and Econometrics, Universite Catholique de Louvain.

Coelli, T.J., and S. Perelman. 1999. A comparison of parametric and non-parametric distance functions: with application to European railways. *European Journal of Operational Research* 117(2):326–339.

Coelli, T.J., and S. Perelman. 2000. Technical efficiency of European railways: a distance function approach. *Applied Economics* 32:1967–1976.

Coelli, T.J., S. Perelman, and E. Romano. 1999. Accounting for environmental influences in stochastic frontier models: with application to international airlines, *Journal of Productivity Analysis* 11(3):251–273.

Coelli, T.J., D.S. Prasada Rao, C.J. O'Donnell, and G.E. Battese. 2005. An introduction to efficiency and productivity analysis. 2 ed. New York: Springer.

Coelli, T.J., S. Rahman, and C. Thirtle. 2003. A stochastic frontier approach to total factor productivity measurement in Bangladesh crop agriculture, 1961 - 92. *Journal of International Development* 15(3):321–333

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

Cornwell, C., P. Schmidt, and R.C. Sickles. 1990. Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics* 46(1-2):185–200.

Crépon, B., E. Duguet, and J. Mairesse. 1998. Research, innovation, and productivity: an econometric analysis at the firm level. *Economics of Innovation and New Technology* 7(2):115–158.

Cuesta, R.A. 2000. A production model with firm-specific temporal variation in technical efficiency: with application to Spanish dairy farms. *Journal of Productivity Analysis* 13(2):139–158.

Dayton, C.M., and G.B. Macready. 1988. Concomitant variable latent class models. *Journal of the American Statistical Association* 83:173-178.

De Alessi, L. 1983. Property rights, transaction costs, and X-efficiency : an essay in economic theory. *American Economic Review* 73(1):64–81.

Debreu, G. 1951. The coefficient of resource utilization. *Econometrica* 19(3):273–292.

De Loecker, J. 2011. Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity. *Econometrica* 79:1407–1451.

De Loecker, J., and J. Konings. 2006. Job reallocation and productivity growth in a post-socialist economy: evidence from Slovenian manufacturing. *European Journal of Political Economy* 22:388–408.

Delgado, M.A., J.C. Farinas, and S. Ruano. 2002. Firm productivity and export markets: a non-parametric approach. *Journal of International Economics* 57:397–422.

Demidova, S., and R.C. Andrés. 2009. Trade policy under firm-level heterogeneity in a small economy. *Journal of International Economics* 78(1):100–112.

Denny, M., M. Fuss, and L. Waverman. 1981. The measurement and interpretation of total factor productivity in regulated industries, with an application to Canadian telecommunications. In: T.G. Cowing and R.E. Stevenson, eds., *Productivity Measurement in Regulated Industries*. New York: Academic Press.

Deprins, D., and L. Simar. 1989. Estimating technical inefficiencies with corrections for environmental conditions with an application to railway companies. *Annals of Public and Cooperative Economics* 60(1):81–102.

Destatis (Statistisches Bundesamt Deutschland). 2012. Sachgebiete und Statistiken, Preise. (http://www-genesis.destatis.de/genesis/online/online)

Deutscher Brauerbund e.V., 2009. Die deutsche Brauwirtschaft in Zahlen 2008. Self-published, Berlin.

Diewert, W.E. 1976. Exact and superlative index numbers. *Journal of Econometrics* 4:115–145.

Diewert, W.E., and K.J. Fox. 2010. On measuring the contribution of entering and exiting firms to aggregate productivity growth. In: W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox, and A.J. Nakamura. eds., *Price and productivity measurement: volume 6 - index number theory*, Trafford Press.

Diewert, W.E., and T.J. Wales. 1987. Flexible functional forms and global curvature conditions. *Econometrica* 55(1):43–68.

Dhrymes, P. 1991. The structure of production technology: productivity and aggregation effects. Columbia University Econ. Dept. working paper 551.

Emvalomatis, G. 2012. Productivity growth in German dairy farming using a flexible modelling approach. *Journal of Agricultural Economics* 63(1):83–101.

Eslava, M., J.C. Haltiwanger, A. Kugler, and M. Kugler. 2004. The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from Colombia. *Journal of Development Economics* 75:333–371.

European Commission. 2013. CAP Reform – an explanation of the main elements. MEMO/13/937.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. 1994. Productivity growth, technical progress and efficiency changes in industrialized countries. *American Economic Review* 84(1):66–83.

Färe, R., and D. Primont. 1995. Multi-output production and duality: theory and applications. Kluwer Academic Publishers, Boston.

Fan, S. 1991. Effects of technological change and institutional reform on the productivity growth in Chinese agriculture. *American Journal of Agricultural Economics* 73(2):266–275.

Farrell, M.J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A, General* 120(3):253–282.

Farsi, M., M. Filippini, and W.H. Greene. 2005. Efficiency measurement in network industries. Application to the Swiss railway companies. *Journal of Regulatory Economics* 28(1):69–90.

Farsi, M., M. Filippini, and M. Kuenzle. 2005a. Unobserved heterogeneity in stochastic cost frontier models. An application to Swiss nursing homes. *Applied Economics* 37(18): 2127–2141.

Fecher, F., P. Pestieau. 1993. Efficiency and competition in OECD financial services. In H.O. Fried, C.A.K. Lovell, and S.S. Schmidt eds., *The Measurement of productive efficiency: techniques and applications*. Oxford: Oxford University Press.

Filippini, M., N. Hrovatin, and J. Zoric. 2008. Cost efficiency of Slovenian water distribution utilities: an application of stochastic frontier methods. *Journal of Productivity Analysis* 29:169–182.

Filippini, M., N. Hrovatin, and J. Zoric. 2010. Productivity growth and price regulation of Slovenian water distribution utilities. *Zbornik radova Ekonomskog Fakulteta u Rijeci* 28(1):89–112.

Foster, L., J.C. Haltiwanger, and C.J. Krizan. 2001. Aggregate productivity growth: lessons from microeconomic evidence. In: C.R. Hulten, E.R. Dean, and M.J. Harper. eds., *New developments in productivity analysis*, Chicago: University of Chicago Press.

Foster, L., J.C. Haltiwanger, and C. Syverson. 2008. Reallocation, firm turnover and effciency: selection on productivity or profitability? *American Economic Review* 98(1):394–425.

Førsund, F.R., C.A.K. Lovell, and P. Schmidt. 1980. A survey of frontier production functions and of their relationship to efficiency measurement. *Journal of Econometrics* 13(1):5-25.

Fox, J. 2012. Problems with (dis)aggregating productivity, and another productivity paradox. *Journal of Productivity Analysis* 37:249–259.

Fox, J., and V. Smeets. 2011. Does input quality drive measured differences in firm productivity? *International Economics Review* 52:961–989.

Fried, H.O., C.A.K. Lovell, and S.S. Schmidt. 2008. Efficiency and productivity. In: H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. eds., *The measurement of productive efficiency and productivity growth*. Oxford: Oxford University Press.

Fuentes, H.J., E. Grifell-Tatjé, and S. Perelman. 2001. A parametric distance function approach for Malmquist index estimation. *Journal of Productivity Analysis* 15:79–94.

Fuss, M., D. McFadden, and Y. Mundlak. 1978. A survey of functional forms in the economic analysis of production. In: M. Fuss, and D. McFadden. eds., *Production economics: a dual approach to theory and applications. Vol. 1: The theory of production*. Amsterdam: North-Holland Publishing Company.

Giannakas, K., K.C. Tran, and V. Tzouvelekas. 2003. On the choice of functional form in stochastic frontier modeling. *Empirical Economics* 28:75–100.

Gollop, F., and D. Jorgenson. 1980. U.S. productivity growth by industry, 1947-73. In: J.W. Kendrick, and B.N. Vaccara. eds., *New development in productivity analysis.* Chicago: University of Chicago Press.

Good, D.H., M.I. Nadiri, L.H. Roller, and R.C. Sickles. 1993 Efficiency and productivity growth comparisons of European and U.S. air carriers: a first look at the data. *Journal of Productivity Analysis* 4(1): 115–125.

Good, D.H., M.I. Nadiri, and R.C. Sickles. 1997. Index number and factor demand approaches to the estimation of productivity. In: H. Pesaran and P. Schmidt. eds., *Handbook of Applied Econometrics: Microeconometrics*. *Volume II*. Oxford: Blackwell Publishers.

Goto, M., and T. Sueyoshi. 2009. Productivity growth and deregulation of Japanese electricity distribution. *Energy Policy* 37(8):3130–3138.

Greene, W.H. 1993. The econometric approach to efficiency analysis. In H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. eds., *The measurement of productive efficiency: techniques and applications*. Oxford: Oxford University Press.

Greene, W.H. 2003. *Econometric Analysis*, 5^{th} ed. Upper Saddle River, NJ: Prentice Hall.

Greene, W.H. 2004. Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics* 13:959–980.

Greene, W.H. 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* 126(2):269–303.

Greene, W.H. 2005a. Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis* 23:7–23.

Greene, W.H. 2007. LIMDEP: Version 9.0. Plainview, NY: Econometric Software.

Greene, W.H. 2007a. LIMDEP Econometric modelling guide, Vol. 1. Econometric Software, New York.

Greene, W.H. 2008. The econometric approach to efficiency analysis. In: H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. eds., *The Measurement of productive efficiency and productivity growth*. Oxford: Oxford University Press.

Greene, W.H. 2012. LIMDEP: Version 10.0. Econometric Software, New York.

Griliches, Z. 1957. Specification bias in estimates of production functions. *Journal of Farm Economics* 39(1):8–20.

Griliches, Z. 1964. Research expenditures, education, and the aggregate agricultural production function. *American Economic Review* 54(6):961–974

Griliches, Z. 1996. The discovery of the residual: a historical note. *Journal of Economic Literature* 24:1324–1330.

Griliches, Z., and J. Hausman. 1986. Errors in variables in panel data. *Journal of Econometrics* 31:93–118.

Griliches, Z., and J. Mairesse. 1991. R&D and productivity growth: comparing Japanese and U.S. manufacturing firms. In: C.R. Hulten. ed., *Productivity growth in Japan and the United States.* Chicago: University of Chicago Press.

Griliches, Z., and J. Mairesse. 1998. Production functions: the search for identification. In: Z. Griliches, ed., *Practicing econometrics: essays in method and application*. Cheltenham: Edward Elgar.

Griliches, Z., and H. Regev. 1995. Firm productivity in Israeli industry: 1979-1988. *Journal of Econometrics* 65:175-203.

Grosskopf, S. 1993. Efficiency and productivity. In: H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. eds., *The measurement of productive efficiency: techniques and applications*. Oxford: Oxford University Press.

Hadley, D. 2006. Patterns in technical efficiency and technical change at the farm-level in England and Wales. *Journal of Agricultural Economics* 57(1):81–100.

Hall, R.E. 1988. The relation between price and marginal cost in U.S. industry. *Journal of Political Economy* 96(5):921–947.

Hall, R.E. 1990. Invariance properties of Solow's productivity residual. In: P. Diamond. ed., *Growth/productivity/unemployment: essays to celebrate Bob Solow's birthday*. Cambridge: MIT Press.

Hall, M., and C. Winsten. 1959. The ambiguous notion of efficiency. *The Economic Journal* 69(273):71–86.

Hallam, D., and F. Machado. 1996. Efficiency analysis with panel data: a study of Portuguese dairy farms. *European Review of Agricultural Economics* 23(1):79–93.

Haney, A.B., and Pollitt, M. 2009. Efficiency analysis of energy networks: an international survey of regulators. Energy Policy 37(12):5814–5830.

Hausman, J.A., and E.E. Taylor. 1981. Panel data and unobservable individual effects. *Econometrica* 49(6):1377-1398.

Hay, D.A., and G.S. Liu. 1997. The efficiency of firms: what difference does competition make? *Economic Journal* 107:597–617.

Hayami, Y., and V.W. Ruttan. 1970. Agricultural productivity differences among countries. *American Economic Review* 60(5):895–911.

Hein, W. 2002. Silomaisanbau in klimatischen Grenzlagen. Pages 21–26 in Silomais und Feldfutterbau im Alpenraum. Proc. Silomaisfachtagung 2002. Agricultural Research and Education Center (AREC) Raumberg-Gumpenstein.

Henningsen, A., and C. Henning. 2009. Imposing regional monotonicity on translog stochastic production frontiers with a simple three-step procedure. *Journal of Productivity Analysis* 32(3):217–229.

Hoch, I. 1962. Estimation of production function parameters combining time-series and cross-section data. *Econometrica* 30(1):34–53.

Hopkins, A. 2011. Mountainous farming in Europe. Pages 3–12 in Grassland farming and land management systems in mountainous regions. Proceedings of the 16th Symposium of the European Grassland Federation. Pötsch, E.M., B. Krautzer, and A. Hopkins. ed. Agricultural Research and Education Center (AREC) Raumberg-Gumpenstein.

Hossain, M.A., and N.D. Karunaratne. 2004: Trade liberalization and technical efficiency: evidence from Bangladesh manufacturing industries. *Journal of Development Studies* 40(3):87–114.

Huang, C.J., and T.J. Liu. 1994. Estimation of a non-neutral stochastic frontier production function. *Journal of Productivity Analysis* 5(2):171–180.

Hulten, C.R. 2001. Total factor productivity. A short biography. In: C.R. Hulten, E.R. Dean, and M.J. Harper. eds., *New development in productivity analysis.* Chicago: University of Chicago Press.

Jin, S., H. Ma, J. Huang, R. Hu, and S. Rozelle. 2010. Productivity, efficiency and technical change: measuring the performance of China's transforming agriculture. *Journal of Productivity Analysis* 33(3):191–207.

Jondrow, J., C.A.K. Lovell, I.S. Materov, and P. Schmidt. 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19(2-3):233–238.

Jorgenson, D.W., and Z. Griliches. 1967. The explanation of productivity change. *Review of Economics Studies* 34(3):249–283.

Jovanovic, B. 1982. Selection and evolution of industry. *Econometrica* 50:649–70.

Karagiannis, G., M. Kellermann, K. Salhofer. 2013. Output price deflators and the sources of productivity growth. Paper presented at the XIII European Workshop on Efficiency and Productivity Analysis, June 16–20, Aalto University, Helsinki, Finland.

Karagiannis, G., P. Midmore, and V. Tzouvelekas. 2004. Parametric decomposition of output growth using a stochastic input distance function. *American Journal of Agricultural Economics* 86(4):1044–1057.

Karagiannis, G., K. Salhofer, and F. Sinabell. 2011. Explaining productivity differences between conventional and organic fooder crop farms in Austria. mimeo.

Karagiannis, G., and V. Tzouvelekas. 2005. Explaining output growth with a heteroscedastic non-neutral production frontier: the case of sheep farms in Greece. *European Review of Agricultural Economics* 32(1):51–74.

Karagiannis, G., and V. Tzouvelekas. 2010. Parametric measurement of time-varying technical inefficiency: results from competing models. *Agricultural Economic Review* 10(1):50–79.

Katayama, H., S. Lu, J.R. Tybout. 2009. Firm-level productivity studies: illusions and a solution. *International Journal of Industrial Organization* 27:403–413.

Kellermann, M., K. Salhofer, W. Wintzer, and C. Stockinger. 2011. Der Zusammenhang zwischen technischer Effizienz und wirtschaftlichem Erfolg: eine Analyse für bayerische Milchviehbetriebe. *German Journal of Agricultural Economics* 60(4):230–242.

Key, N., and W. McBride. 2007. The changing economics of U.S. Hog production. ERR-52. U.S. Dept. of Agric., Economic Research Service.

Key, N., W. McBride, and R. Mosheim. 2008. Decomposition of total factor productivity change in the U.S. hog industry. *Journal of Agricultural and Applied Economics* 40(1):137–149.

Kim, S., and G. Han. 2001. A decomposition of total factor productivity growth in Korean manufacturing industries. A stochastic frontier approach. *Journal of Productivity Analysis* 16(3):269–281.

Klette, T., and Z. Griliches. 1996. The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics* 11(4):343–361.

Knight, F.H. 1933. The economic organization. New York: Harper and Row.

Koopmans, T.C. 1951. An analysis of production as an efficient combination of activities. In: T.C. Koopmans, ed., *Activity analysis of production and allocation*. Cowles Commission for Research in Economics 13. New York: John Wiley and Sons.

Kumbhakar, S.C., S.N. Ghosh, and J.T. MacGuckin. 1991. A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms. *Journal of Business and Economic Statistics* 9(3):279–286.

Kumbhakar, S.C., and A. Heshmati. 1995. Efficiency measurement in Swedish dairy farms: an application of rotating panel data, 1976–1988. *American Journal of Agricultural Economics* 77:660–674.

Kumbhakar, S.C., and L. Hjalmarsson. 1993. Technical efficiency and technical progress in Swedish dairy farms. In: H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. eds., *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford: Oxford University Press.

Kumbhakar, S.C., and G. Lien. 2010. Impact of subsidies on farm productivity and efficiency. In: V.E. Ball, R. Fanfani, L. Gutierrez. eds., The economic impact of public support to agriculture. New York: Springer.

Kumbhakar, S.C., G. Lien, and J.B. Hardaker. 2014. Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis* 41(2):321–337.

Kumbhakar, S.C., and C.A.K. Lovell. 2000. *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Kumbhakar, S.C., R. Ortega-Argilés, L. Potters, M. Vivarelli, and P. Voigt. 2012. Corporate R&D and firm efficiency: evidence from Europe's top R&D investors. *Journal of Productivity Analysis* 37(2):125–140

Kumbhakar, S.C., E.G. Tsionas, and T. Sipiläinen. 2009. Joint estimation of technology choice and technical efficiency: An application to organic and conventional dairy farming. *Journal of Productivity Analysis* 31:151–161.

Lai, H., and C.J. Huang. 2010. Likelihood ratio tests for model selection of stochastic frontier models. *Journal of Productivity Analysis* 34(1):3–13.

Last, A.K., and H. Wetzel. 2010. The efficiency of German public theaters: a stochastic frontier analysis approach. *Journal of Cultural Economics* 34(1):89–110.

Lau, L.J. 1978. Testing and imposing monotonicity, convexity and quasi-convexity constraints. In: M. Fuss, and D. McFadden. eds., *Production economics: a dual approach to theory and applications. Vol. 1: The theory of production*. Amsterdam: North-Holland Publishing Company.

Leibenstein, H. 1966. Allocative efficiency vs. X-efficiency. *American Economic Review* 56(3):392–415.

Leibenstein, H. 1976. Beyond economic man. Cambridge, MA: Harvard University Press.

Levinsohn, J., and A. Petrin. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70:317–341.

Liu, Y., and R. Myers. 2009. Model selection in stochastic frontier analysis with an application to maize production in Kenya. *Journal of Productivity Analysis* 31:33–46.

Lovell, C.A.K. 1993. Production frontiers and productive efficiency. In: H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. eds., *The measurement of productive efficiency: techniques and applications*. Oxford: Oxford University Press.

Lovell, C.A.K., S. Richardson, P. Travers, and L.L. Wood. 1994. Resources and functionings: a new view of inequality in Australia. In: W. Eichhorn. ed. *Models and measurement of welfare and inequality*. Berlin: Springer.

Lovell, C.A.K. 1996. Applying efficiency measurement techniques to the measurement of productivity change. *Journal of Productivity Analysis* 7(2-3):329–340.

LfL (Bayerische Landesanstalt für Landwirtschaft). 2009. Buchführungsergebnisse des Wirtschaftsjahres 2007/2008. Freising.

LfL (Bayerische Landesanstalt für Landwirtschaft). 2012. Milchreport Bayern 2011 – Ergebnisse der Betriebszweigabrechnung Milchproduktion 2010/11. Freising.

Ma, H., L. Oxley, A. Rae, C. Fan, J. Huang, and S. Rozelle. 2012. The evolution of productivity performance on China's dairy farms in the new millennium. *Journal of Dairy Science* 95:7074–7085

MacDonald, D., J.R. Crabtree, G. Wiesinger, T. Dax, N. Stamou, P. Fleury, J. Gutierrez Lazpita, and A. Gibon. 2000. Agricultural abandonment in mountain areas of Europe: Environmental consequences and policy response. *Journal of Environmental Management* 59:47–69.

Maddala, G.S. 1987. Recent developments in the econometrics of panel data analysis. *Transportation Research* 21:303–326.

Mairesse, J., and J. Jaumandreu. 2005. Panel-data estimates of the production function and the revenue function: what difference does it make? *Scandinavian Journal of Economics* 107(4):651–672.

Marschak, J., and W. Andrews. 1944. Random simultaneous equations and the theory of production. *Econometrica* 12:143–205.

Mayen, C.D., J.V. Balagtas, and C.E. Alexander. 2010. Technology adoption and technical efficiency: organic and conventional dairy farms in the United States. *American Journal of Agricultural Economics* 92:181–195.

Meisser, M., and U. Wyss. 1998. Wettereinfluss auf Wachstum und Reifung von Silomais. In: *Agrarforschung* 5:317–320.

Melitz, M. 2003. The impact of trade on intra-industry reallocation and aggregate industry productivity. *Econometrica* 71(6):1695–1725.

Melitz, M.J., and S. Polanec. 2012. Dynamic Olley-Pakes productivity decomposition with entry and exit. Working Paper 18182, National Bureau of Economic Research, Cambridge.

Moreira, V.H., and B.E. Bravo-Ureta. 2010. Technical efficiency and technological gap ratios for dairy farms in three southern cone countries: a stochastic meta-frontier model. *Journal of Productivity Analysis* 33:33–45.

Morrison, C.J. 1992. Unraveling the productivity growth slowdown in the United States, Canada and Japan: the effects of subequilibrium, scale economies and markups. *Review of Economics and Statistics* 74(3):381–393.

Morrison Paul, C.J., W.E. Johnston, and G.A. Frengley. 2000. Efficiency in New Zealand sheep and beef farming: the impacts of regulatory reform. *Review of Economics and Statistics* 82(2):325–337.

Müller-Lindenlauf, M., C. Deittert, and U. Köpke. 2010. Assessment of environmental effects, animal welfare and milk quality among organic dairy farms. *Livestock Science* 128:140–148

Mundlak, Y. 1961. Empirical production function free of management bias. *Journal of Farm Economics* 43:44–56.

Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica* 46(1):69–85.

Mundlak, Y. 1996. Production function estimation: reviving the primal. *Econometrica* 64(2):431–438.

Mundlak, Y., and I. Hoch. 1965. Consequences of alternative specifications of Cobb-Douglas production functions. *Econometrica* 33(4):814–828.

Murillo-Zamorano, L. 2004. Economics efficiency and frontier techniques. *Journal of Economic Surveys* 18(1):33–77.

Nelson, R.R. 1968. A “diffusion” model of international productivity differences in manufacturing industry. *American Economics Review* 58(5):1219–1248.

Newman, C., and A. Matthews. 2006. The productivity performance of Irish dairy farms 1984-2000: a multiple output distance function approach. *Journal of Productivity Analysis* 26(2):191–205.

Newman, C., and A. Matthews. 2007. Evaluating the productivity performance of agricultural enterprises in Ireland using a multiple output distance function approach. *Journal of Agricultural Economics* 58:128–151.

NGG (Gewerkschaft Nahrung-Genuss-Gaststätten). 2009. Branchenbericht 2009 der Brauwirtschaft. Self-published.

NGG (Gewerkschaft Nahrung-Genuss-Gaststätten). 2013. Branchenbericht 2013 der Brauwirtschaft. Self-published.

Nicoletti, G., and S. Scarpetta. 2003. Regulation, productivity, and growth: OECD evidence. Working Paper. World Bank Policy Research.

Nishimizu, M., and J.M. Page. 1982. Total factor productivity growth, technological progress and technical efficiency change: dimensions of productivity change in Yugoslavia, 1965–78. *The Economic Journal* 92(Dec):920–936.

Nishimura, K.G., T. Nakajima, and K. Kiyota. 2005. Does the natural selection mechanism still work in severe recessions? Examination of the Japanese economy in the 1990s. *Journal of Economic Behavior & Organization* 58:53–78.

OECD. 2001. Measuring productivity – OECD manual: measurement of aggregate and industry-level productivity growth. Paris.

O'Donnell, C.J., and T.J. Coelli. 2005. A Bayesian approach to imposing curvature on distance functions. *Journal of Econometrics* 126(2):493–523.

O'Donnell, C.J., D.S.P. Rao, and G. Battese. 2008. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics* 34:231–255.

Olley, S., and A. Pakes. 1996. The dynamics of productivity in the telecommunications industry. *Econometrica* 64(6):1263–1298

Orea, L. 2002. Parametric decomposition of a generalized Malmquist productivity index. *Journal of Productivity Analysis* 18:5–22.

Orea, L., and S.C. Kumbhakar. 2004. Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics* 29(1):169–183.

Ornaghi, C. 2006. Assessing the effects of measurement errors on the estimation of production functions. *Journal of Applied Economics* 21:879–891.

Ohta, M. 1975. A note on the duality between production and cost functions: rate of returns to scale an rate of technical progress. *Economic Studies Quaterly* 25:63–65.

Oulton, N. 1998. Competition and labor productivity amongst UK companies. *Oxford Economic Papers* 50:23–38.

Pavcnik, N. 2002. Trade liberalization, exit and productivity improvements: evidence from Chilean plants. *Review of Economic Studies* 69(1):245–276.

Peeters, A. 2009. Importance, evolution, environmental impact and future challenges of grasslands and grassland-based systems in Europe. *Grassland Science* 55:113–125.

Perelman, S., and P. Pestieau. 1988. Technical performance in public enterprise: a comparative study of railways and postal services. *European Economic Review* 32:432–441.

Pflimlin, A., B. Buczinski, and C. Perrot. 2005. Proposition de zonage pour préserver la diversité des systèmes d'élevage et des territoires européens. *Fourrages* 182:311–329.

Pflimlin, A., and X. Poux. 2005. Biodiversity of grasslands and livestock systems in Europe – Redefining the political issues. In: R. Lillak, R. Viiralt, A. Linke, and V. Geherman. eds., *Integrating efficient grassland farming and biodiversity*. Proc. of the 13th International Occasional Symposium of the European Grassland Federation. Tartu, Estonia.

Pitt, M., and L.F. Lee. 1981. The measurement and sources of technical inefficiency in the indonesian weaving industry. *Journal of Development Economics* 9(1):43–64.

Poetsch, E.M. 2007. LIFS and livestock production – grassland and dairy farming in Austria. In: K. Biala, J.-M. Terres, P. Pointerau and M.L. Paracchini. eds., *Low Input Farming Systems: an Opportunity to Develop Sustainable Agriculture*. Proc. of the JRC Summer University, Ranco.

Polachek, S., and B. Yoon. 1996. Panel estimates of a two-tiered earnings frontier. *Journal of Applied Econometrics* 11(2):169–178.

Private Brauereien Bayern e.V. 2012. Die deutsche Brauwirtschaft: Stand Februar 2014. Self-published, Munich.

Private Brauereien Bayern e.V. 2014. Die deutsche Brauwirtschaft: Stand Februar 2014. Self-published, Munich.

Prochnow, A., M. Heiermann, M. Plöchl, B. Linke, C. Idler, T. Amon, and P.J. Hobbs. 2009. Bioenergy from permanent grassland – A review: 1. Biogas. *Bioresource Technology* 100:4931-4944.

Prucha, I.R., and M.I. Nadiri. 1996. Endogenous capital utilization and productivity measurement in dynamic factor demand models. Theory and an application to the U.S. electrical machinery industry. *Journal of Econometrics* 71:343–379.

QMS, Inc. 2007. EViews. Irvine, CA.

Rae, A.N., H. Ma, J. Huang, and S. Rozelle. 2006. Livestock in China: commodity-specific total factor productivity decomposition using new panel data. *American Journal of Agricultural Economics* 88(3)680–95.

Rasmussen, S. 2010. Scale efficiency in Danish agriculture. An input distance-function approach. *European Review of Agricultural Economics* 37(3):335–367.

Reinhard, S., C.A.K. Lovell, and G. Thijssen. 1999. Econometric estimation of technical and environmental efficiency: an application to Dutch dairy farms. *American Journal of Agricultural Economics* 81:44–60.

Richter, M.K. 1966. Invariance axioms and economic indexes. *Econometrica* 34:739–755.

Rizov, M., J. Pokrivcak, and P. Ciaian. 2013. CAP subsidies and productivity of EU farms. *Journal of Agricultural Economics* 64:537–557.

Roberts, M.J., and D. Supina. 2000. Output price and markup dispersion in micro data: the role of producer heterogeneity and noise. In: M.R. Baye. ed. *Advances in applied microeconomics, Vol. 9, Industrial organization.* Greenwich: JAI Press.

Saal, D.S., D. Parker, and T. Weyman-Jones. 2007. Determining the contribution of technical change, efficiency change and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985 - 2000. *Journal of Productivity Analysis* 28(1-2):127–139.

Sanderson, M.A., L.W. Jolley, and J.P. Dobrowolski. 2012. Pastureland and hayland in the USA: land resources, conservation practices, and ecosystem services. In: C.J. Nelson. ed. *Conservation outcomes from pastureland and hayland practices: assessment, recommendations, and knowledge gaps.* Lawrence, KS: Allen Press.

Sauer, J., K. Frohberg, and H. Hockmann. 2006. Stochastic efficiency measurement: the curse of theoretical consistency. *Journal of Applied Economics* 9(1):139–165.

Sauer, J., and C.J. Morrison-Paul. 2013. The empirical identification of heterogeneous technologies and technical change. *Applied Economics* 45(11):1461–1479.

Schmidt, P. 1988. Estimation of a fixed-effect Cobb-Douglas system using panel data. *Journal of Econometrics* 37(3):361–380.

Schmidt, P., and R.C. Sickles. 1984. Production frontiers and panel data. *Journal of Business & Economic Statistics* 2(4):367–74.

Schultz, T.W. 1956. Reflections of agricultural production, output and supply. *Journal of Farm Economics* 38(3):748–762.

Self, S.G., and K.Y. Liang. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* 82(398):605–610.

Shephard, R.W. 1953. Cost and production functions. Princeton, NJ: Princeton University Press.

Sherlund, S.M., C.B. Barrett, and A.A. Adesima. 2002. Smallholder technical efficiency controlling for environmental production conditions. *Journal of Development Economics* 69(1):85–101.

Solow, R.M. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics* 39(3):312–20.

Soussana, J.F., V. Allard, K. Pilegaard, P. Ambus, C. Amman, C. Campbell, E. Ceschia, J. Clifton-Brown, S. Czobel, R. Domingues, C. Flechard, J. Fuhrer, A. Hensen, L. Horvath, M. Jones, G. Kasper, C. Martin, Z. Nagy, A. Neftel, A. Raschi, S. Baronti, R.M. Rees, U. Skiba, P. Stefani, G. Manca, M. Sutton, Z. Tuba, and R. Valentini. 2007. Full accounting of the greenhouse gas (CO2, N20, CH4) budget of nine European grassland sites. *Agriculture, Ecosystems and Environment* 121:121-134

Stigler, G.J. 1967. The xistence of X-efficiency. *American Economic Review* 66(1):213–216.

Syverson, C. 2004. Market structure and productivity: a concrete example. *Journal of Political Economy* 112(6):1181–1222.

Thanassoulis, E. 2001. Introduction to the theory an application of Data Envelopment Analysis: a foundation text with integrated software. Boston: Kluwer Academic Publishers.

Thaysen, J., H. Dreyer, H.P. Sierts, and J. Thomsen. 2010. Yields of forage crops in Schleswig-Holstein 1985-2008 under farm-scale and trial conditions. In: H. Schnyder. ed. *Grassland in a Changing World.* Proc. of the 23rd General Meeting of the European Grassland Federation, Kiel.

Timmer, C.P. 1971. Using a probabilistic frontier function to measure technical efficiency. *Journal of Political Economy* 79:776–794.

Tinbergen, J. 1942. Zur Theorie der langfristigen Wirtschaftsentwicklung. *Weltwirtschaftliches Archiv* 55(1):511–549.

Tovar, B., F.J. Ramos-Real, and E.F. de Almeida. 2011. Firm size and productivity. Evidence from the electricity distribution industry in Brazil. *Energy Policy* 39 (2): 826–33.

Train, K. 2009. Discrete choice models with simulation, 2nd Ed. New York: Cambridge University Press.

Tsionas, M. 2002. Stochastic frontier models with random coefficients. *Journal of Applied Econometrics* 17(2):127–147.

Tzouvelekas, V., C.J. Pantzios, and C. Fotopoulos. 2001. Technical efficiency of alternative farming systems: the case of Greek organic and conventional olive-growing farms. *Food Policy* 26(6):549–569.

USDA (U.S. Department of Agriculture). 2013. Grassland Reserve Program. Last accessed Dec. 03, 2013. www.fsa.usda.gov/FSA/webapp?area=home&subject=copr&topic=grp

Van Beveren, I. 2012. Total factor productivity estimation: a practical review. *Journal of Economic Surveys* 26(1):98–128.

Van Biesebroeck, J. 2008. Aggregating and decomposing productivity. *Review of Business and Economics* 53(2):122–146.

Wang, H.J., and C.W. Ho. 2010. Estimating fixed-effect panel stochastic frontier models by model transformation. *Journal of Econometrics* 157(2):286–296.

Wang, H.J., and P. Schmidt. 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis* 18:129-144.

Wedel, M., and W.S. DeSarbo. 2002. Mixture regression models. In: J. Hagenaars, and A.L. McCutcheon. eds., *Applied latent class analysis*. New York: Cambridge University Press.

Wetzel, H. 2009. Productivity growth in European railways: technological progress, efficiency change and scale effects. In: Working Paper Series in Economics, Leuphana Universität Lüneburg.

Wooldridge, J.M. 2009. On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters* 104(3):112–114.

Wu, Y. 1995. Productivity growth, technological progress, and technical efficiency change in China: a three-sector analysis. *Journal of Comparative Economics* 21(2):207–229.

Yao, S., Z. Liu, and Z. Zhang. 2001. Spatial differences of grain production efficiency in China, 1987 - 1992. *Economics of Planning* 34(1-2):139–157.

Zellner, A., J. Kmenta, and J. Dreze. 1966. Specification and estimation of Cobb-Douglas production function models. *Econometrica* 34(4):784–795.

Zhu, X., and A. Oude Lansink. 2010. Impact of CAP subsidies on technical efficiency of crop farms in Germany, the Netherlands and Sweden. *Journal of Agricultural Economics* 61(3):545–564.