CESIS
Electronic Working Paper Series

Paper No. 126

What Drives the Productive Efficiency of a Firm? The Importance of Industry, Location, R&D, and Size

Oleg Badunenko
(DIW Berlin)

Michael Fritsch
(Friedrich Schiller University Jena, Max Planck Institute of Economics Jena and DIW Berlin)

Andreas Stephan
(Jönköping International Business School, DIW Berlin, CESIS Stockholm)

April 2008

What Drives the Productive Efficiency of a Firm? The Importance of Industry, Location, R&D, and Size[*]

Oleg Badunenko
German Institute for Economic Research, DIW Berlin, 10108, Berlin, Germany
(e-mail: obadunenko@diw.de)

Michael Fritsch
Friedrich Schiller University Jena, Max Planck Institute of Economics Jena and German Institute for Economic Research (DIW Berlin), Germany (e-mail: m.fritsch@uni-jena.de)

Andreas Stephan
Jönköping International Business School, CESIS Stockholm and German Institute for Economic Research (DIW Berlin), Germany (e-mail: andreas.stephan@ihh.hj.se)

Abstract

This paper investigates the factors that explain the level and dynamics of manufacturing firm productive efficiency. In our empirical analysis, we use a unique sample of about 39,000 firms in 256 industries from the German Cost Structure Census over the years 1992-2005. We estimate the efficiencies of the firms and relate them to firm-specific and environmental factors. We find that (1) about half the model's explanatory power is due to industry effects, (2) firm size accounts for another 20 percent, and (3) location of headquarters explains approximately 15 percent. Interestingly, most other firm characteristics, such as R&D intensity, outsourcing activities, or the number of owners, have extremely little explanatory power. Surprisingly, our findings suggest that higher R&D intensity is associated with being less efficient, though higher R&D spending increases a firm's efficiency over time.

Keywords: Frontier analysis, determinants of efficiency, firm performance, industry effects, regional effects, firm size

JEL classification: D24, L10, L25

[*]This paper is based on a cooperation project with the German Federal Statistical Office in Wiesbaden, which enabled access to confidential firm-level micro data from the annual Cost Structure Census in Germany. We are grateful to Gerald Göbel, Ottmar Hennchen, and Roland Sturm at the German Federal Statistical Office for their support of this project. The paper benefited from comments by Johan Klaesson, Charlie Karlsson, Martin Andersson, and seminar participants at the Jönköping International Business School, the Royal Economic Society Conference (2005) in Nottingham, the meeting of the Verein für Socialpolitik (2005) in Bonn, the International Industrial Organization Conference (2006) in Atlanta, the Spring Meeting of Young Economists (2006) in Seville, and the Efficiency and Productivity Analysis Workshop (HEPAW, 2006) in Brussels.

1 Introduction

Empirical analyses show that firm productivity varies considerably even when the firms are operating in the same market (for an overview, see Bartelsman and Doms (2000)). While some firms operate at the technological frontier and earn high profits, others lag considerably behind and barely survive. There may be many reasons for these differences, including, among others, managerial restrictions, slow adaptation to changes in the market environment and/or technology, location, and frictions in the labor market. It is the intent of this paper to identify the determinants of such differences in the performance at the firm level. We analyze the level and the development of firm technical[1] efficiency, which is its relative productivity compared to the highest attainable level. Specifically, we are looking for answers to questions such as: What are the reasons for diverging efficiency of firms? Which factors explain why some firms are more efficient than others? How does firm efficiency evolve over time?

Empirical investigation into the determinants of efficiency dates back to the early 1990s. For instance, Lovell (1993) stated that identifying the factors that explain differences in efficiency is essential for improving the results of firms, but that, unfortunately, economic theory does not supply a theoretical model of determinants of efficiency. However, Caves and Barton (1990) and Caves (1992) suggested that several studies have developed a strategy for identifying the determinants of efficiency, which can be grouped into several categories: (i) factors external to the firm; (ii) factors internal to the firm; and (iii) ownership structures (e.g., public vs. private).

To find answers to the questions set out above, we take a look at each of these categories of determinants. In particular, we distinguish between firm-specific and environmental factors much in the spirit of Caves and Barton (1990). Environmental factors are not under direct control of the firm, at least not in the short run. We consider industry affiliation and firm location to be important environmental factors. Firm-specific factors, on the other hand, are characteristics that can be influenced by the firm in the short run. Among the firm-specific factors we analyze are firm size, R&D intensity, and degree of outsourcing.

Our study makes several important contributions to the literature on the determinants of efficiency. First, to the best of our knowledge, none of the previous analyses used such a rich dataset to simultaneously analyze the influence of numerous firm-specific and environmental factors on efficiency. Indeed, previous studies either focus on industry characteristics (e.g., Roudaut, 2006) or regional (e.g., Li and Hu, 2004), or size effects (e.g., Oczkowski and Sharma, 2005; Söderbom and Teal, 2004), aand thus provide only limited insight into the relative importance of a single influence. Second, we are not aware of any study using a representative sample of firms for the whole manufacturing sector of a national economy. Third, we apply the concept of partial R^2 in the second step of our analysis because doing so is a more appropriate method of describing the importance of factors than the commonly used t-values when the number of

[1] The terms productive and technical efficiency are used interchangeably throughout the paper.

observations is huge, as in our case. In contrast to t-values, partial R^2s enable us to compare the relative importance of continuous variables with the relative importance of categorical ones, such as industry or location.

Our econometric analysis is based on data from the Cost Structure Census of the German Federal Statistical Office. This is a unique and representative micro-panel dataset containing approximately 39,000 firms and covering 40 percent of all manufacturing firms in Germany over the period from 1992 to 2005. We estimate efficiencies as firm-specific fixed effects, as proposed by Schmidt and Sickles (1984). The major advantage of this approach, compared to other stochastic frontier frameworks, is that it does not require any *a priori* assumption regarding the distribution of efficiency across firms. Such distributional assumptions are often quite restrictive and sometimes unsupported by the data.

The analysis yields some important results. (1) Industry affiliation is the most important factor for explaining efficiency at the firm level, contributing almost half of the model's explanatory power for the level, and even more so for the development, of efficiency. (2) Firm size and headquarter location contribute approximately 20 and 15 percent, respectively. (3) Other factors such as R&D, organization of production, and relative size (production share in domestic industry) have only negligible explanatory power, which is surprising given that these factors have been emphasized as important in previous studies (e.g., Ornaghi, 2006). This paper has mainly an explorative character; fundamental explanations of the influence mechanisms behind the various factors lies beyond its scope. Nevertheless, we provide novel insights into the importance of certain factors for explaining productive efficiency and its development.

The paper is structured as follows. Section 2 discusses hypotheses regarding the determinants of efficiency, which are tested in the empirical analysis. Section 3 describes the methodology for assessing productive efficiency, gives specifics on the data used to estimate the production function and efficiency scores, and discusses the obtained results. Section 4 reports the analysis of the determinants of productive efficiency, sets out our reasons for using the partial R^2 concept, and describes the variables of the Cost Structure Census dataset used in the second step of the analysis. Section 5 deals with the analysis of the dynamics of efficiency at the firm level during the period 1992-2005. Section 6 provides a summary of empirical findings and concluding remarks.

2 Productive efficiency of manufacturing firms

The classical microeconomic textbook treats all manufacturing firms as homogeneous producing units and, therefore, assume that all firms operate at the same level of efficiency. However, empirical studies frequently show that in the real world some firms are more efficient than others (e.g., Caves, 1989). Productive efficiency characterizes the firm's ability to derive the maximum output from a certain bundle of inputs with given technology. The concept of efficiency was introduced by Farrell (1957), who used the concept proposed by Koopmans (1951) and the radial type of efficiency measure considered by Debreu (1951). In this paper, we test five hypotheses on the determinants of efficiency differences across manufacturing firms in Germany during 1992-2005.

Hypothesis 1 *Industry affiliation explains a large proportion of the differences in productive efficiencies across firms.*

Industry affiliation refers to the main business activity of a firm. In the literature, it is often assumed that industry affiliation can be used as a proxy for the relevant product market (e.g., Schmalensee, 1985; Wernerfelt and Montgomery, 1988). If industry affiliation is related to the product market, it should indicate the degree of competition a firm faces. Therefore, in industries with intense competition, we hypothesize that average efficiency will be higher, as inefficient firms are forced by competitive pressure to leave the market. The firm's industry affiliation can also be interpreted as describing the unobserved characteristics of the production technology employed and of the product markets where the firms operate. Additionally, according to Klepper (1997) and Klepper and Simons (2005), the efficiency of an industry depends on its stage in the industry lifecycle.

Hypothesis 2 *Firm location is important in explaining firms' productive efficiencies.*

A firm's location influences its efficiency in several ways. For example, Beeson and Husted (1989) found that in the United States, a considerable part of the variation of efficiency can be attributed to regional differences of the labor force characteristics, levels of urbanization, and industry structure. Second, the firm's location may affect its innovation activities, with consequences for its production process and efficiency (for an overview, see Cooke, Heidenreich and Braczyk, 2004). Furthermore, the effect of locational conditions on efficiency is partly embedded in knowledge spillovers (Krugman, 1991; Antonelli, 2003). Third, spatial proximity to other establishments, as occurs in an agglomeration or a cluster, may be conducive to economic performance for a number of reasons, including, for example, rich and diversified input markets (Baptista and Swann, 1998; Porter, 1998, 2003).

Hypothesis 3 *Efficiency is positively related to firm size.*

From a theoretical viewpoint, the relationship between firm size and efficiency is not clear-cut (Audretsch, 1999). On the one hand, larger firms have better market penetration and are better able to exploit economies of scale and scope. Larger firms also have more money and are able to employ better managers (Kumar, 2003). On the other hand, it is more difficult to keep all departments coordinated and operating efficiently in a large firm (Leibenstein, 1966). In contrast, the employees of smaller firms may be more motivated by competitive-based incentive schemes rather than financial ones (Agell, 2004), thus possibly making them more efficient (profitable) to the firm. These hypotheses have been extensively tested in the literature. For instance, Gumbau-Albert and Maudos (2002), using a panel of 1,149 Spanish firms from 18 manufacturing industries, arrived at the conclusion that firm size is conducive to efficiency. Torii (1992) claimed that the efficiency can be positively related to the scale or size of a firm if it is assumed that maintaining or improving efficiency incurs costs in terms of the firm's management because larger firms tend to be less resource constrained.

Hypothesis 4 *Outsourcing activities and R&D enhance the productive efficiency of a firm.*

Grossman and Helpman (2005) emphasize that "... firms seem to be subcontracting an ever expanding set of activities, ranging from product design to assembly, from research and development to marketing, distribution, and after-sales service." A number of studies find that a high level of outsourcing has a positive effect on efficiency, but some studies state that the positive role of outsourcing is often overestimated (Heshmati, 2003). The relationship between productive efficiency and R&D investment is also ambiguous (Bartelsman and Doms, 2000). Some researchers have confirmed a positive relationship between R&D and efficiency (see Ornaghi, 2006, and the references therein), but others (see, e.g., Albach, 1980; Caves and Barton, 1990) find that R&D intensity has a negative impact on productive efficiency. In an attempt to explain this negative, effect Caves and Barton (1990, p. 76) hypothesize that the R&D expenditures of an industry are only a poor predictor of that industry's innovativeness because a large part of the innovation output will be applied in other industries. Additionally, investment in R&D is by its very nature risky and will pay off, if it even does, at a considerable time lag.

Hypothesis 5 *The average productivity level of all firms increases over time, whereas the average relative efficiency level remains constant.*

It can be expected that technical progress will yield productivity improvements over time. Moreover, it is commonly accepted in economics that competition will result in an efficient use of scarce resources. Competition is a very powerful mechanism that provides incentives for an efficient organization of production. Competition will force inefficient firms to leave the market, thereby increasing the average productivity level in the industry. If markets are predominantly competitive, the firms' average productivity level is expected to increase over time. However, in contrast to productivity, the average efficiency of firms, which is measured *relative* to the most efficient firm(s), is hypothesized to remain constant over time.

3 Production frontier and efficiency measurement

3.1 Distribution-free approach to measuring productive efficiency

A point of reference is needed in measuring the productive efficiency of a firm. The stochastic frontier model as proposed simultaneously by Aigner et al. (1977) and Meeusen and van den Broeck (1977) is the most commonly used approach for measuring productive efficiency.[2] The stochastic frontier model of Battese and Coelli (1995) can be employed if panel data are available. Though the stochastic frontier models have some advantages in distinguishing efficiency from other random influences on a firm's output, they are based on rather restrictive assumptions. First, a distributional assumption on the inefficiency term is imposed, which may not be supported by the data. For instance, Schmidt and Lin (1984) showed that if the skewness of residuals resulting from an ordinary least squares (OLS) regression is positive, the stochastic frontier approach should not be used.[3] Second, it is assumed that productive efficiency and production inputs are not correlated. In empirical applications, however, such a correlation is actually likely to exist, resulting in inconsistent parameter estimates. Third, the conditional mean model of Battese and Coelli (1995) can be estimated only for a moderate number of explanatory variables because it is based on a single-step maximum likelihood (ML) procedure. However, since the second step of our analysis includes more than 700 variables (e.g., dummies for industry and location), we cannot use available ML-based procedures. Fourth, firm-specific efficiencies in the stochastic frontier approach are computed as expected values (Jondrow, Lovell, Materov and Schmidt, 1982) and must be obtained indirectly from the residual term, whereas the fixed-effects approach provides direct estimates of the relative efficiency of a firm.

Therefore, we take advantage of the panel character of our data and measure productive inefficiency as a firm-specific effect.[4] The basic specification is a deterministic transcendental logarithmic (translog) production function, which can be written as (see Greene, 1997):

$$\ln y_{it} = \ln \alpha_i + \lambda_t + \sum \beta_k \ln x_{kit} + \sum \beta_{2\_k} (\ln x_{kit})^2 + \frac{1}{2} \sum_{q \neq w} \gamma_{qw} (\ln x_{qit}) (\ln x_{wit}) + \varepsilon_{it} \qquad (1)$$

where $k=1,\ldots,p$, $i=1,\ldots,N$, $t=1,\ldots,T_i$ and $q=1,\ldots,p$, $w=1,\ldots,p$, $q \neq w$. The term y_{it} represents output of firm i in period t; x_{kit} denotes production input k, and λ_t represents a time-specific effect. We have N firms and T_i observations for each firm. The assessment of productive efficiency is based on the firm-specific fixed effects α_i. The largest estimate of a firm-specific

[2] See Mayes, Lansbury and Harris (1995) and Kumbhakar and Lovell (2003) for an overview of different parametric approaches for assessing the efficiency of firms.

[3] An exception is Carree (2002) who proposes a stochastic frontier model with positive skewness of productive efficiency. However, we are not aware of any empirical application using this approach to date.

[4] See Schmidt and Sickles (1984) and Sickles (2005) for a more detailed discussion on such an approach.

fixed effect max $\hat{a}_j$ in each industry is used as a benchmark value that represents the highest attainable efficiency level. Productive efficiency E_i of firm i is then estimated as:

$$\widehat{E_i} = \frac{\hat{\alpha}_i}{\max \hat{\alpha}_j} \cdot 100 \quad [\%] \tag{2}$$

At least one firm in an industry will meet the benchmark value and the remaining firms will have positive efficiency estimates between 0 and 100 percent.[5]

Several caveats of the fixed effects approach should be mentioned. First, recent developments in efficiency measurement provide models that allow the distinction between a firm's inefficiency and unobserved heterogeneity (see Greene, 2005). Accordingly, the fixed effects do not only capture "pure" productive efficiency differences between firms but also other (unobserved) differences, such as diverging management or marketing strategies. However, for our sample of approximately 39,000 firms, Greene's approach is computationally too demanding.[6] Second, because prices of inputs and outputs are not available at the firm level, we do not measure a pure input-output quantity relationship with the production function, since all inputs as well as the output are measured in monetary terms. Accordingly, the estimated fixed effects indicate not only that at a given level of inputs some firms produce higher output than others, but also that some firms can obtain higher market prices for their output, or benefit from lower input prices. Our interpretation of this measurement issue is that the fixed effects also measure a type of price efficiency of firms. However, we are confident that using inputs and outputs in monetary terms is not a major drawback, which is supported by evidence from Mairesse and Jaumandreu (2005), who find that using a nominal output measure in a production function estimation yields a quite negligible difference in comparison to using a real output measure. Furthermore, monetary values allow the aggregation of multiple outputs into a single output measure as well as the aggregation of different inputs and make aggregation of inputs and outputs of different qualities feasible, since prices will adjust for those differences.

3.2 Data

Our analysis is based on micro data from the German Cost Structure Census[7] of Manufacturing for the 1992 to 2005 period (see Fritsch, Görzig, Hennchen and Stephan, 2004). The Cost Structure Census is gathered and compiled by the German Federal Statistical Office (*Statistis-*

[5]Note that in the second step analysis the fixed effects are not expressed relatively to the maximum fixed effect in the respective industry, since this would affect the scale of the estimated industry effects. All other results remain unchanged when absolute instead of relative fixed effects are used in the regression analysis.

[6]One further shortcoming of the "true" fixed effects stochastic frontier model is that it leads to biased parameter estimates and biased estimates of productive efficiencies for panels with relatively few observations, as in our case (cf. Greene, 2005).

[7]Aggregate figures are published annually in *Fachserie* 4, *Reihe* 4.3 of the German Federal Statistical Office (various years).

ches Bundesamt). The survey consists of all the large German manufacturing firms that have 500 or more employees over the entire period. To limit the reporting burden for smaller firms, firms with 20 499 employees are included only as a random sample that can be assumed as being representative for this size category as a whole. Firms with less than 20 employees are not included.[8] As a rule, the smaller firms report for four consecutive years and then are substituted by other small firms (rotating panel).[9] Because the estimation of firm-specific fixed effects requires at least two observations, firms with only one observation are excluded, thus leaving approximately 39,000 firms in the sample.

The Cost Structure Census contains information for a number of input categories, including payroll; employer contributions to the social security system; fringe benefits; and expenditures for material inputs, self-provided equipment, goods for resale, and for energy. Also included is information on expenditures for external wage-work, external maintenance and repair, tax depreciation of fixed assets, subsidies, rents and leases, insurance costs, sales tax, other taxes, public fees, and interest on outside capital, as well as "other" costs such as license fees, bank charges, and postage or expenses for marketing and transport. Further information available in the Cost Structure Census includes industry affiliation; type of business (craft or manufacturing); location of headquarters; value of the stocks of raw materials, goods for resale, and final output; and the amount of R&D expenditure as well as the number of R&D employees.[10] The information on employment comprises the number of owners actively working in the firm and the number of full-time, part-time, home-based, and temporary workers.

3.3 Estimation results of the production frontier

Table 1 displays the parameter estimates of a translog production function according to Equation (1) based on the micro data of the individual firms.[11] We include dummy variables for the different years of the observation period, with 2005 being the year of reference to account for yearly shifts in the frontier. The fit of the regression (R^2) is remarkably high (0.995) and the fixed firm effects as well as the year effects are highly significant.[12]

Several specification tests were performed to see whether our estimated technology is consistent with predictions from neoclassical production theory. First, we investigated whether the translog specification is superior to a simple Cobb-Douglas specification by testing the null

[8]Since 2001 the statistics also contain firms with 1-19 employees. However, these firms are not included in our analysis due to a rotating sampling scheme; only one observation is available for most of these small firms.

[9]Due to mergers or insolvencies, some firms have less than four observations. Note, however, that firms are legally obligated to respond to the Cost Structure Census; thus, there are actually almost no missing observations due to nonresponse.

[10]Information on resources devoted to R&D has been gathered in the Cost Structure Census since 1999.

[11]Least squares dummy variables (LSDV) method for panel data; see Baltagi (2001) and Coelli, Rao and Battese (2002) for details on this approach.

[12]The results of a Hausman-Wu test indicate correlation between fixed effects and the other explanatory variables. Thus, a random effects model or a stochastic frontier framework is not appropriate in this case.

Table 1: Estimation results of the logarithmic Translog production function with fixed effects, years 1992-2005

Variable	Coefficient	p-value
β_{mat}	0.209	<.0001
β_{lab}	0.229	<.0001
β_{ene}	0.028	0.0016
β_{cap}	0.245	<.0001
β_{oth}	0.167	<.0001
β_{ext}	0.116	<.0001
β_{2_mat}	0.055	<.0001
β_{2_lab}	0.073	<.0001
β_{2_ene}	0.007	<.0001
β_{2_cap}	0.027	<.0001
β_{2_oth}	0.026	<.0001
β_{2_ext}	0.016	<.0001
γ_{mat_lab}	−0.057	<.0001
γ_{mat_ene}	−0.001	0.2093
γ_{mat_cap}	−0.02	<.0001
γ_{mat_oth}	−0.008	<.0001
γ_{mat_ext}	−0.017	<.0001
γ_{lab_ene}	−0.009	<.0001
γ_{lab_cap}	−0.033	<.0001
γ_{lab_oth}	−0.036	<.0001
γ_{lab_ext}	−0.012	<.0001
γ_{ene_cap}	0.003	0.0016
γ_{ene_oth}	−0.003	<.0001
γ_{ene_ext}	−0.001	0.0002
γ_{cap_oth}	−0.009	<.0001
γ_{cap_ext}	−0.001	0.0191
γ_{oth_ext}	0.001	0.1135
D_{1992}	0.022	<.0001
D_{1993}	0.009	<.0001
D_{1994}	0.012	<.0001
D_{1995}	0.017	<.0001
D_{1996}	0.011	<.0001
D_{1997}	0.015	<.0001
D_{1998}	0.013	<.0001
D_{1999}	0.017	<.0001
D_{2000}	0.014	<.0001
D_{2001}	0.006	<.0001
D_{2002}	−0.004	0.0146
D_{2003}	−0.002	0.0711
D_{2004}	0.003	0.0236
R^2		0.995
Number of observations		217,415

Notes: mat: material inputs, lab: labor compensation, ene: energy consumption, cap: capital, oth: other inputs, ext: external services.

hypothesis $\beta_{2\_i} = 0$ and $\gamma_{ij} = 0$ for all i and j. This null hypothesis is strongly rejected (p-value < 0.0001) indicating that the translog specification is more appropriate. Second, the H_0 that $(\sum \beta_{2\_i} + \sum \gamma_{ij})$ ($j \neq i$) is equal to zero[13] is not rejected (p-value = 0.41). This indicates a

[13] This sum of estimates is 0.000474, with a standard error of 0.000572.

homothetic production technology; that is, the marginal rate of technical substitution is homogeneous of degree zero with regard to inputs. Third, given homotheticity and because the test of H_0 that $\sum\beta=1$ yields a p-value of 0.89, we conclude that the estimated technology is linearly homogeneous.[14]

Output elasticities can be calculated from the translog estimates using the formula $\sigma_{yi} = {\partial \ln y}/{\partial \ln x_i} = \beta_i + \beta_{2\_i} \ln x_i + \sum_{i \neq j} \beta_{ij} \ln x_j$. The output elasticities at different values of production inputs (1, 5, 25, 50, 75, 95, and 99 quantiles) are shown in Table 2. Note that they all add up to about unity and are not very different from median production shares of production inputs as reported in Table A.1 in the data appendix, exactly what one would expect according to neoclassical theory (Chambers, 1988). This is further support for the plausibility of our production function estimates.[15]

Table 2: Output elasticities of input factors at different input levels

Input factor	Output elasticity at input level						
	p1	p5	Q1	Median	Q3	p95	p99
Material inputs	0.194	0.332	0.392	0.418	0.441	0.460	0.470
Labor compensation	0.612	0.489	0.394	0.351	0.320	0.293	0.277
Energy consumption	0.015	0.020	0.026	0.030	0.035	0.043	0.051
Capital	0.096	0.081	0.075	0.067	0.056	0.045	0.038
External services	0.046	0.052	0.070	0.081	0.088	0.095	0.098
Other inputs	0.032	0.032	0.052	0.065	0.073	0.082	0.086
Sum	0.995	1.004	1.009	1.012	1.015	1.018	1.020

Notes: p1, p5, p95 and p99 are the 1^{st}, 5^{th}, 95^{th}, and 99^{th} percentiles, respectively; Q1 and Q3 are lower and upper quantiles.

Comparing the output elasticities at different hypothetical scales of production tells us a few more things about production technology. First of all, the sum of elasticities is never statistically different from one. This is because the elasticities are obtained from parameter estimates that are in accordance with a homothetic production function. Second, as the input scale increases, the marginal products of labor and capital are decrease, whereas the marginal productivity of the material (intermediates) is increases, thus making the substitution of labor and capital by material more profitable. This implies that the larger the scale of a firm in terms of its inputs, the more profitable it is for the firm to rely on intermediate inputs. Note that the elasticity gradually increases from 0.194 for the first percentile of the input value to 0.470 for the 99th percentile. This finding is in line with evidence from previous studies that large manufacturing firms, in particular, have increased their outsourcing intensity in recent years (Görzig and Stephan, 2002).

[14] The sum of single input estimates is 0.9945 with a standard error of 0.01691.

[15] As an alternative to a single production function for all industries we also estimated industry-specific translog function at the 3- and 4-digit level respectively, but obtained less satisfactory results, e.g. negative output elasticities or returns to scale significantly outside the range [0.5, 1.5]. Given that the common production function estimation over all industries yields plausible results, we are convinced that this approach is appropriate.

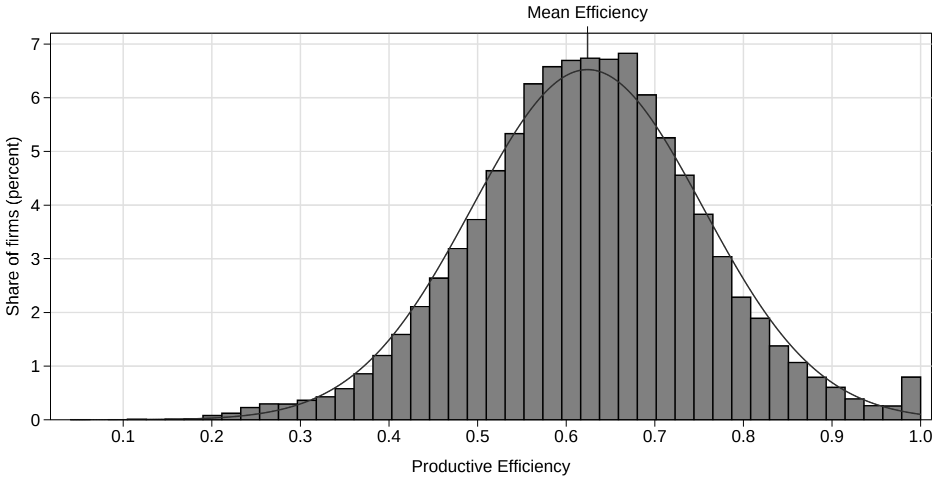

Figure 1: Histogram of efficiency at the micro level and normal density (38,641 observations)

3.4 Distribution of productive efficiency

Table 3 shows the parameters of the distribution of productive efficiency scores calculated according to Equation (2). In general, the distribution of productive efficiency is centered and most firms are clustered close to the mean (Figure 1). The peak seen in distribution at the maximum level is because, by definition, at least one firm in each industry is fully efficient; that is, each industry has a different $\max\hat{\alpha}_j$s used as the benchmark in Equation (2) for the other firms in that industry. Symmetry as well as skewness of the distribution of productive efficiency largely coincides with the normal distribution. This is reassuring as it confirms the appropriateness of using OLS in the second step of the analysis.

4 Determinants of productive efficiency

4.1 Partial R^2s and variables used in the second step of analysis

To analyze the determinants of productive efficiency, we relate the estimated productive efficiencies to a number of explanatory variables. We employ analysis of covariance (ANCOVA), where independent variables can be both metric and categorical, as the regression method. Since

Table 3: Distribution of productive efficiency

Variable	N	Mean	CV	p90	Q3	Median	Q1	p10	min
Efficiency	38641	0.625	0.209	0.785	0.707	0.624	0.542	0.461	0.041

Notes: p10 and p90 are the 10^{th} and 90^{th} percentiles; CV is the coefficient of variation; Q1 and Q3 are lower and upper quantiles.

categorical variables (e.g., industry affiliation) may have a large number of levels (categories), we do not report the single estimates for each category (i.e. for each dummy variable) but instead provide partial R^2 for each variable or effect. Partial R^2s are preferred over t-statistics in analyses with a large number of observations since the significance of simple t-tests does not express the explanatory power of a variable or an effect (McCloskey and Ziliak, 1996). Partial R^2 are defined as (see Greene, 2003, p. 36):

$$R^2_{x|z} = \frac{R^2_{x,z} - R^2_z}{1 - R^2_z} \tag{3}$$

where $R^2_{x|z}$ is the partial R^2 of variable(s) x, $R^2_{x,z}$ is the R^2 for the model including all variables x and z, and R^2_z is the model R^2 where only the z-variables are included.

The partial R^2 of a variable expresses how much of the variation of the dependent variable can be explained by this particular variable, or by a subset of dummy variables (representing a categorical variable) *given that the other variables are included in the model*. Therefore, the partial R^2 measures the difference of the model's R^2 with and without a certain variable or effect. Theil (1971) emphasizes the importance of measuring the incremental contribution of a variable for explaining the dependent variable. Furthermore, Flury (1989) and Shea (1997) argue that partial statistics should be especially taken into consideration when analyzing the relevance of variables in multivariate models. Moreover, Hamilton (1987) highlights the merit of partial correlations in determining which explanatory variables to keep in the case of correlated variables.

Since the productive efficiency estimate for each firm is time invariant, the second step of the analysis is based on the cross-section of firms. All explanatory variables are included as firm-specific averages over the observation period. Even in this cross-sectional setup it is possible to include year dummies for the years a firm is included in the sample. The respective year dummy is set to 1 if the firm is observed in that year; 0 otherwise. The estimation of year dummies with cross-sectional data is possible since not all firms are observed over the entire period; some firms are only included only in subperiods. The year dummies capture the overall trend of the firms' average efficiency. For instance, if average efficiency improves over time we should find significantly higher estimates of the year dummy variables for the later years compared to the first years of the sample period.

Table 4 provides an overview of the firm-level information available in the Cost Structure Census that is included in the second step of our analyses. The dataset provides a unique opportunity to investigate the relative importance of a broad range of determinants of efficiency that have not been investigated in previous studies due to data constraints. In our single study, we are able to combine the effects of both internal and environmental factors and also control for a number of other variables.[16] Further details about the data can be found in the Appendix.

With the standard errors of efficiencies estimated in the first step, it is possible to apply the more efficient weighted least squares method, instead of OLS, in the second step, using the reciprocals of the standard errors of efficiency as weights. However, the results changed so little that we decided to report only the OLS results in the following sections.[17]

4.2 Empirical results

Table 5 displays the partial R^2 values, which indicate the relative importance of a variable for the entire observation period, 1992-2005 (Model I), or for the last six years, 1999-2005 (Model II). Conducting the analyses for the subperiod of 1999 to 2005 allows the inclusion of information on R&D intensity and temporarily employed (subcontracted) labor, which is only available for from 1999 onward. Table 6 provides the signs, magnitudes, and t-values for all continuous and some selected categorical variables. We include the number of observation periods as a control variable for sample selection. Of potential concern in these estimates is that some inefficient firms exit the market and are consequently not included in the sample in later years, a situation known as panel attrition. This could lead to an attrition bias since efficiency is the dependent variable of the analysis. If this is the case, we should find a significantly positive relationship between a firm's observation periods and its efficiency. However, we find that the number of observation periods is negatively correlated with efficiency, although with low explanatory power measured in terms of partial R^2. Hence, we cannot preclude that there is a sample selection bias, but in the opposite direction of attrition – firms that stay in the sample longer, presumably the larger ones, tend to be less efficient. An indication of an attrition bias is found only for the subgroup of least efficient firms (Table 8), which is probably due to a moderate survivor bias for this group of firms.

Several conclusions can be drawn from the results in Tables 5 and 6. First, in both models, for the 1992-2005 and 1999-2005 period, all included independent variables – except the year effects – have significant explanatory power at the 1 percent level. This might in part be driven by the huge size of the dataset. However, with regard to the magnitudes of partial R^2s, we can state that industry affiliation, firm size, and location have by far the most important effects on productive efficiency. Jointly, the effects adds up to 84 percent (Model I) and 82 percent (Model II) of the models' explanatory power.

[16] Note that the industry classification changed in 1995 from WZ1979 to WZ1993, the latter corresponding to the international NACE classification. We kept only those firms in the sample for which an industry affiliation according to WZ1995 is available, i.e. which have at least one observation after the year 1994. Furthermore, in the second step of our analysis of the determinants of efficiency, we excluded all firms that changed industry affiliation, location, or legal form during the observation period.

[17] The WLS results can be obtained from the authors upon request.

Table 4: Names and definitions of variables

Name	Description
Environmental factors	
– Industry affiliation	Industry dummies at the 4-digit level (255 industries)
– Location	District (Kreis) of the headquarter of the enterprise (440 districts)
– Year effects	Dummy variable for each year, 1992-2005
Firm-specific factors	
a) Firm characteristics	
– Size	Six categories: less then 49 employees (= 1), 50-99 employees (= 2), 100-249 employees (= 3), 250-499 employees (= 4), 500-999 employees (= 5), more than 1000 employees (= 6)
– Share in industry	Relative production share of German suppliers in the respective industry
– R&D intensity	Share of R&D personnel over total employment (available from 1999 on)
b) Outsourcing activities	
– Quota of external contract work	Expenditure for external contract work / internal labor cost
– Quota of external services	Expenditure for external services / internal labor cost
– Quota of material inputs	Expenditure for material inputs / internal labor cost
– Quota of temporarily employed labor	Expenditure for temporary employed labor / internal labor cost; available from 1999 on
– Quota operating leases	Operating leasing expenses / capital depreciations; available from 1999 on
c) Ownership	
– Type of business	Manufacturing (= 1) / craft (= 0) dummy variable
– Number of owners working in the firm	Number of owners working in the firm

Second, the results suggest that efficiency is largely explained by the industry in which the firm is operating. The great importance of industry effects is echoed in the literature, which emphasizes the role of industry in explaining firm profitability (Cubbin and Geroski, 1987; Schmalensee, 1985). These results are broadly consistent with hypothesis 1. Industry effects might capture different degrees of competition in the respective markets (Fritsch and Stephan, 2004a) or might accrue from different stages of the industry lifecycle or different technological regimes (Fritsch and Stephan, 2004b). The "black box" of industry effects may also have something to do with the necessity of firms in certain industries to innovate, for example, the

Table 5: Partial R^2s (in percent)

Variable	Model I: 1992-2005		Model II: 1999-2005	
	df	Partial R^2	df	Partial R^2
Environmental factors				
Industry affiliation	256	9.34*	256	10.29*
Location (district)	439	3.12*	443	2.77*
Year-effects	14	0.72	7	0.41
Firm-specific factors				
a) Firm characteristics				
Size category	5	4.51*	5	3.38*
Production share in industry	1	0.01*	1	0.04*
Number of owners working in the firm	1	0.43*	1	0.44*
R&D intensity			1	0.20*
b) Outsourcing activities				
Quota of material inputs	1	1.27*	1	1.41*
Quota of external contract work	1	0.74*	1	0.77*
Quota of external services	1	0.03*	1	0.17*
Quota of temporarily employed labor			1	0.01
Quota rents and leases			1	0.00005
Sample selection control				
Number of years observed	1	0.02*	1	0.10*
Overall R^2	21.78		21.11	
Sum of all partial R^2s	20.19		20.00	
Number of observations	38,641		24,339	

Notes: Dependent variable: productive efficiency; df is degrees of freedom; statistical significance at the 1 percent level is indicated (*).

chemical industry. Industrial differentiation might also stem from differences in average quality of inputs, the degree of implied product differentiation, or be due to characteristics of production technology (e.g., Carlsson, 1972).

Third, firm size contributes about 20 percent to the model's explanatory power. This finding supports Hypothesis 3, and also confirms the results of other studies finding different efficiency performance among different firm size classes (e.g., Alvarez and Crespi, 2003; Caves, 1992; Torii, 1992). However, our results are in the opposite direction of the effects found in these other studies: according to our analysis, firms become less efficient as size increases. Thus, smaller firms are, on average, significantly more efficient than larger ones (Table 6). For example, the group of firms with less than 49 employees is on average 15 percent more efficient than the group of firms with more than 1,000 employees. Similarly, we find that relative size, measured in terms of production share in total industry production, is negatively related to efficiency. Therefore, Hypothesis 3 holds with respect to the significance, but not with regard to the direction, of the size effect.

Table 6: Parameter estimates for selected variables

Variable	Model I: 1992-2005	Model II: 1999-2005
Firm-specific factors		
a) Size category		
Less than 49 employees	0.15^{*} (25.25)	0.12^{*}(18.67)
50−99 employees	0.11^{*}(19.49)	0.09^{*}(14.77)
100−249 employees	0.08^{*}(14.34)	0.07^{*}(11.30)
250−499 employees	0.06^{*}(9.50)	0.04^{*}(7.16)
500−999 employees	0.04^{*}(5.73)	0.03^{*}(4.60)
More than 1000 employees	–	–
Production share in industry	-0.07^{*} (−2.2)	-0.11^{*}(−3.26)
Number of owners working in the firm	0.01^{*}(12.79)	0.01^{*}(10.23)
R&D intensity		-0.14^{*}(−6.9)
b) Outsourcing activities		
Quota of material inputs	0.01^{*}(22.06)	0.01^{*}(18.39)
Quota of external contract work	0.04^{*}(16.82)	0.04^{*}(13.57)
Quota of external services	-0.02^{*}(−3.3)	-0.05^{*}(−6.36)
Quota of temporarily employed labor		0.03(1.35)
Quota rents and leases		1E−07 (0.10)
Sample selection control		
Number of years observed	-0.005^{*}(−2.56)	-0.002^{*}(−4.87)
Number of observations	38,641	24,339

Notes: It is not possible to present all estimates, since ANCOVA gives an estimate for every category of a nominal variable, resulting in 256 estimates for each industry *etc*. Estimates for all categories are available upon request; statistical significance at the 1 percent level is indicated (*).
t-values in parentheses.

Fourth, the location effect is captured by including 440 dummy variables for the German districts (*Kreise*). It is worth noting that with this approach we not only capture differences in the performance of the firms located in the eastern or western part of Germany (e.g., Funke and Rahn, 2002), but also assess the efficiency of firms at a much smaller geographical scale. The results for firm location suggest that regional factors play a fairly important role. The explanatory power of location in terms of partial R^2 is 3.12 percent for the 1992-2005 period and 2.77 percent for the 1999-2005 period (Table 5). Thus, these finding are grounds for accepting Hypothesis 2. The location variable refers to the firm's headquarters, not to the location of branch plants, which may be located in other regions. However, since more than 90 percent of the firms in the Cost Structure Census are single-establishment firms, the effect of branch plants located in other regions is not expected to be large or important.

Furthermore, firm size is the only firm-specific determinant that explains a large part of productive efficiency (Table 5). Other factors, such as the share of R&D expenditure, the firm's legal form, and indicators for the degree of outsourcing are not important. The parameter estimates (Table 6) show a negative effect of R&D on productive efficiency. This confirms the empirical findings of Albach (1980), Caves and Barton (1990) and Hoskisson et al. (1994), but is counterintuitive since it seems as though R&D should lead to improved products or cost reduction (Aghion and Howitt, 1992; Grossman and Helpman, 1991). One explanation for this odd finding may be that there can be a considerable time lag between R&D spending and R&D results (Helpman, 1992). If this is the case, R&D expenditure is simply an additional cost at the time it is incurred, thereby reducing productive efficiency at that time, whereas the benefits can be appropriated only later. Unfortunately, we cannot test for longer time lags since information on R&D activity is available in our data for only the last six years. In addition, R&D is risky and a considerable share of projects are likely to fail, thus possibly making it an inefficient use of resources, no matter what time period is examined. We also find that most outsourcing activities enhance efficiency, which goes toward proving 4, however, the effect of R&D is negative, which contradicts this hypothesis. Moreover, the partial R^2s for both variables are of fairly small magnitude. In sum, then, Hypothesis 4 must be rejected.

Finally, the year dummy variables are not significant.[18] Since we are looking at the average efficiency of firms, this is not surprising: some firms improve their efficiency, others become less efficient. The resulting net effect is zero. This explains why we do not find an improvement of average efficiency over time, a finding in support of Hypothesis 5.

4.3 Subgroups of different efficiency performance

To obtain a more detailed understanding of the factors that contribute to the observed efficiency differences between firms, we conduct the analyses for three subgroups: (i) the 10 percent least efficient firms ("worst performers"), (ii) the 10 percent most efficient firms ("best performers"), and (iii) firms with an efficiency level between these groups ("medium performers"). The partial R^2s and parameter estimates appear in Tables 7 and 8, respectively. Each of these tables contains six models. We first present the analyses of three subgroups for the period 1992-2005; the remaining results refer to the same subgroups for the later period, 1999-2005.

The results for the subgroups show that the significance as well as the relative importance of certain influences differ tremendously across the three different groups of firms. In particular, many of the previously statistically significant effects are no longer import. Several of these findings deserve special mention.

[18] Parameters are not reported here to conserve space, but they are available upon request from the authors.

Table 7: Partial R^2 (in percent): groups of the 10 percent least efficient, the 10 percent most efficient, and firms between 10 and 90 percent efficiency

Variable	1992-2005						1999-2005					
	(I)		(II)		(III)		(IV)		(V)		(VI)	
	df	10% least efficient	df	Between 10% least and 10% most efficient	df	10% most efficient	df	10% least efficient	df	Between 10% least and 10% most efficient	df	10% most efficient
Environmental factors												
Industry affiliation	231	12.67*	255	6.49*	223	16.42*	218	27.33*	254	7.55*	214	20.77*
Location (district)	428	15.68*	446	2.29*	429	11.56	385	24.72	443	2.99*	419	16.43
Year-effects	14	1.08	14	0.3	14	0.62	7	0.28	7	0.36	7	0.05
Firm-specific factors												
a) Firm characteristics												
Size category	5	0.35	5	5.26*	5	0.66*	5	0.08	5	4.88*	5	0.94*
Share in industry	1	0.06	1	0.0002	1	0.24*	1	0.09	1	0.002	1	0.35
Number of owners working in the firm	1	0.01	1	0.56*	1	0.02	1	0.001	1	0.57*	1	0.06
R&D intensity	–	–	–	–	–	–	1	0.04	1	0.09*	1	0.09
b) Outsourcing activities												
Quota of material inputs	1	2.35*	1	1.64*	1	0.05	1	0.62*	1	2.05*	1	0.02
Quota of external contract work	1	0.09	1	0.09*	1	2.42*	1	0.0001	1	0.10*	1	1.54*
Quota of external services	1	1.98*	1	0.003	1	0.07	1	0.54	1	0.01	1	0.14
Quota of temporarily employed labor	–	–	–	–	–	–	1	0.11	1	0.0001	1	0.01
Quota rents and leases	–	–	–	–	–	–	1	0.001	1	0.001	1	0.03
Sample selection control												
Number of years	1	0.30*	1	0.03*	1	0.19*	1	1.91*	1	0.19*	1	0.41
Sum of partial R^2s	34.56		16.66		32.26		55.72		18.8		40.85	
Overall R^2	31.09		16.98		29.95		44.55		19.37		37.02	
Number of obs.	3,865		30,911		3,865		1,720		20,115		2,504	

Notes of Table 5 apply.

The fourth and six models utilize the tenth and ninetieth percentiles cut-off values which are used for the first and third models, respectively.

Table 8: Parameter estimates of selected variables: groups of the 10 percent least efficient (“worst performers”), the 10 percent most efficient (“best performers”), and firms between 10 and 90 percent efficiency (“medium performers”)

Variable	1992-2005			1999-2005		
	(I)	(II)	(III)	(IV)	(V)	(VI)
	10% least efficient	b/n 10% and 90%	10% most efficient	10% least efficient	b/n 10% and 90%	10% most efficient
Firm-specific factors						
a) Size category						
less than 49 employees	0.02 (1.51)	0.08* (24.27)	0.07* (3.89)	−0.01 (−0.51)	0.08* (20.50)	0.07* (3.07)
50−99 employees	0.01 (0.98)	0.07* (19.64)	0.07* (3.64)	−0.01 (−0.27)	0.07* (17.02)	0.06* (2.99)
100−249 employees	0.001 (0.05)	0.05* (14.53)	0.06* (3.06)	−0.004 (−0.20)	0.05* (12.92)	0.05 (2.27)
250−499 employees	−0.0003 (−0.02)	0.03* (9.52)	0.05* (2.77)	0.001 (0.05)	0.03* (8.42)	0.04 (1.74)
500−999 employees	0.01 (0.86)	0.02* (5.48)	0.05* (2.44)	−0.01 (−0.40)	0.02* (5.39)	0.03 (1.20)
More than 1000 employees						
Share in industry	−0.12 (−1.39)	−0.01 (−0.27)	0.34* (2.79)	−0.09 (−1.00)	−0.01 (−0.57)	0.37* (2.55)
Number of owners working in the firm	0.002 (0.57)	0.01* (13.05)	0.002 (0.75)	0.001 (0.10)	0.01* (10.56)	0.003 (1.08)
R&D intensity	–	–	–	−0.04 (−0.65)	−0.06* (−4.26)	−0.08 (−1.28)
b) Outsourcing activities						
Quota of material inputs	0.03* (8.75)	0.01* (22.41)	−0.001 (−1.24)	0.01* (2.62)	0.01* (20.14)	−0.0004 (−0.62)
Quota of external contract work	0.03 (1.67)	0.01* (5.19)	0.03* (8.89)	0.001 (0.03)	0.01* (4.50)	0.02* (5.38)
Quota of external services	1.2015	0.01 (0.89)	0.02 (1.53)	−0.05* (−2.43)	−0.01* (−1.64)	0.04 (1.62)
Quota of temporarily employed labor	–	–	–	−0.14 (−1.11)	0.002 (0.12)	−0.02 (−0.32)
Quota rents and leases	–	–	–	4.88E−05 (0.12)	2.44E−07 (0.34)	−4.59E−05 (−0.74)
Sample selection control						
Number of years observed	0.02* (3.07)	−0.003* (−3.20)	−0.01* (−2.47)	0.01* (4.62)	−0.002* (−6.09)	−0.004* (−2.77)
Number of observations	3,865	30,911	3,865	1,720	20,115	2,504

Notes of Table 6 apply.

First, the magnitudes of partial R^2s for the effect of industry affiliation (Table 7) clearly reinforce the results of the previous section and therefore support Hypothesis 1. Despite the fact that in absolute terms, the partial R^2s of industry affiliation for the best and worst performers are larger than for the medium performers, in relative terms, industry affiliation provides approximately 40 percent (more than 50 percent for the 1999-2005 period) of the explanatory power of the models. Thus, Hypothesis 1 holds true irrespective of the firms' level of productive efficiency.

Second, within the subgroup of medium performers, the size effects are similar to those (Table8) observed for the entire sample (Table 5) for both periods, 1992-2005 and 1999-2005. Moreover, in this subgroup, larger firms are, again, less efficient than their smaller counterparts. For the worst performers, however, size has no explanatory power. In the group of best performers, the size effects have only 0.02 percent explanatory power and lead us to reject Hypothesis 3 for the worst and best performing firms.

Third, location effects are notably different across the three subgroups. Location effects are not statistically significant for the group of best performers. However, they are pronounced for the worst performers in period 1992-2005, but, oddly, not significant for the period 1999-2005. The parameter estimates of the district dummies reflect the average efficiency of the firms located in the respective district. Though in the beginning of the 1990s, firms in East Germany have been rather inefficient as a result of the transition of the former socialist regime, this clear East versus West separation in the efficiency of districts can not be found for the later period of 1999-2005. Rather, there is a mixture of East and West German districts among the least and most efficient locations, indicating that locational effects are not solely due to East or West German regional differences but might be caused by other (nonobserved) reasons. Thus, Hypothesi 2 is supported with regard to medium performers, but not for worst and best performers.

Fourth, the results for the medium performer subgroup also confirm Hypotheses 4 and 5. A heterogenous picture emerges for the best and worst performing firms (Table 8). For example, the quota of material inputs has a positive impact for the worst and medium performers but is not significant for the best performers. The quota of external services has a negative impact on efficiency for worst performers but is not significant for the two other groups. However, external contract work is conducive to efficiency for the best performers. Thus, in addition to the relatively low explanatory power of outsourcing activities the evidence on the direction of effects for efficiency are ambiguous. Likewise, for the worst and best performers, R&D intensity is statistically insignificant. Both partial R^2s as well as the coefficient are statistically significant only for medium performers. Thus, surprisingly, R&D does not explain any statistically significant variation of productive efficiency at the two ends of the efficiency distribution.

Overall, three effects are responsible for most of the explanatory power: (i) industry, (ii) size, and (iii) location. All other factors, both firm-specific and environmental, yield statistically

Table 9: Distribution of estimated linear efficiency trends θ_i

Variable	N	mean	cv	p90	q3	median	q1	p10
Trend	3,876	-0.004	0.017	0.013	0.004	-0.004	-0.011	-0.021

Notes: p10 and p90 are the 10^{th} and 90^{th} percentiles; cv is the coefficient of variation; q1 and q3 are lower and upper quantiles.

significant parameters estimates in some cases, but have only rather little explanatory power. This evidence again corroborates our preference regarding interpreting partial R^2s instead of simple t-ratios in assessing the relative importance of various factors.

5 Determinants of the dynamics of productive efficiency

Finally, we examined the development of productive efficiency at the firm level. To do so, the approach outlined in Equation (1) was easily extended by adding the term $\theta_i t$, where θ_i denotes a firm-specific parameter and t is a time trend, $t = 1, \ldots, T_i$. This model allows for firm-specific (linear) changes in productive efficiency over time (Kumbhakar, Heshmati and

Table 10: Partial R^2 (in Percent): determinants of the dynamics of firm efficiency

Variable	Model I: 1992-2005		Model II: 1999-2005	
	Df	Partial R^2	Df	Partial R^2
Environmental factors				
Industry affiliation	247	22.41*	247	22.60*
Location (district)	413	17.36*	413	17.51*
Year-effects	14	0.4	7	0.001
Firm-specific factors				
a) Firm characteristics				
Size category	5	0.41	5	0.41
Production share in industry	1	0.001	1	0.005
Number of owners working in the firm	1	0.02	1	0.01
R&D intensity			1	0.17
b) Outsourcing activities				
Quota of material inputs	1	0.02	1	0.01
Quota of external contract work	1	0.59*	1	0.56*
Quota of external services	1	0.25	1	0.23
Quota of temporarily employed labor			1	0.005
Quota rents and leases			1	0.02
Sample selection control				
Number of years observed	1	0.02*	1	0.10*
Overall R^2	36.31		36.51	
Sum of all partial R^2s	41.45		42.00	
Number of observations	3,147		3,116	

Notes: Dependent variable: θ_i, notes of Table 5 apply.

Table 11: Parameter estimates of selected variables: determinants of the dynamics of firm efficiency

	Model I: 1992-2005	Model II: 1999-2005
Firm-specific factors		
a) Size category		
Less than 49 employees	−0.002 (−1.34)	−0.002 (−1.18)
50−99 employees	−0.000236 (−0.16)	0.0001 (0.10)
100−249 employees	0.001 (0.67)	0.001 (0.83)
250−499 employees	0.001 (0.89)	0.001 (1.08)
500−999 employees	0.001 (1.03)	0.001 (1.14)
More than 1000 employees	–	–
Production share in industry	−0.001 (−0.14)	−0.001 (−0.25)
Number of owners working in the firm	−0.0002 (−0.62)	−0.0002 (−0.49)
R&D intensity	–	0.02^{**} (2.02)
b) Outsourcing activities		
Quota of material inputs	0.00006 (0.61)	0.00005 (0.46)
Quota of external contract work	0.004^{*} (3.8)	0.004^{*} (3.71)
Quota of external services	-0.003^{*} (-2.5)	-0.003^{*} (-2.35)
Quota of temporarily employed labor	–	0.003 (0.34)
Quota rents and leases	–	0.000009 (0.67)
Year Dummies		
D_{1992}	0.001 (0.50)	–
D_{1993}	−0.003 (−0.91)	–
D_{1994}	0.001 (0.49)	–
D_{1995}	0.00004 (0.01)	–
D_{1996}	0.00009 (0.03)	–
D_{1997}	−0.0002 (−0.05)	–
D_{1998}	−0.0001 (−0.04)	–
D_{1999}	−0.001 (−0.41)	−0.0008 (−0.35)
D_{2000}	−0.002 (−0.59)	−0.002 (−0.63)
D_{2001}	0.002 (0.58)	0.002 (0.69)
D_{2002}	0.002 (0.80)	0.002 (0.97)
D_{2003}	0.001 (0.48)	0.001 (0.54)
D_{2004}	−0.003 (−1.23)	−0.003 (−1.26)
D_{2005}	0.004^{*} (2.29)	0.004^{*} (2.84)
Number of observations	3,147	3,116

Dependent variable: θ_i, notes of Table 6 apply.

Hjalmarsson, 1999). The parameter θ_i indicates whether a firm's efficiency increases $(\theta_i > 0)$ or decreases $(\theta_i < 0)$ with time t. Therefore, in this part we extended the translog production function framework by including firm-specific time trends. We performed this analysis only for firms with at least 10 observations in order to obtain more reliable estimates of θ_i. We also refrained from including a quadratic time trend in the translog production function, as the high collinearity between the linear and quadratic time trends leads to imprecise estimates of both trends. The sample in this step is comprised of about 3,900 firms, which nonetheless cover almost all industries and locations.

The distribution of estimated time trends is presented in Table 9. While about 10 percent of the best performing firms improved their efficiency about 1.3 percent per year, the average (or median) firm experienced a slight efficiency decline. For the 10 percent of the worst performing firms, efficiency decreased by an annual rate of about 2 percent. This finding serves as an additional argument in support of Hypothesis 5.

In the last step of the empirical analysis, we explore the determinants for the positive or negative firm-specific time trends in efficiency. We regress the parameter estimates θ_i as in the previous analyses on the same set of explanatory variables. The partial R^2s are reported in Table 10 and the parameter estimates (selected variables) are displayed in Table 11. The picture that emerges from this analysis of firm-specific efficiency trends is in line with the former results: the overwhelming part of the variation in efficiency trends is explained by industry and location. Other environmental or firm-specific factors have only minor impact.

The estimates presented in Table 11 suggest that, first, a change in efficiency is independent of the size of the firm. Second, two factors determine the development of efficiency: the industry in which the firm is operating and its location. Third, only two of the outsourcing activities have a significant impact: quota of external contract work (positive sign) and quota of external services (negative sign). However, these effects appear to offset one another. One further remarkable contrast to the analysis for the level is that R&D has now a positive impact on the development of efficiency, albeit with extremely low explanatory power. We infer from these findings that there is an inverse relationship between R&D and the level of efficiency, but that firms with a higher R&D intensity tend to improve their efficiency over time.

6 Conclusions

This paper analyzed the importance of a variety of factors to the productive efficiency of firms, with particular emphasis on industry, location, R&D, and size. In a first step, we obtained estimates from a translog production frontier and then, in a second step, performed analysis of covariance to investigate the determinants for firm-specific productive efficiency and its dynamics. We employed the concept of partial R^2 to gauge the relative importance of the various factors.

The translog production function estimates for firms covering the entire manufacturing sector are in accordance with predictions from neoclassical theory for competitive product and factor markets, that is, the average firm operates with constant returns to scale technology. Second, industry affiliation is the most important factor, having the largest share in the model's explanatory power. This holds both for the level and the development of efficiency. Third, size effects have the second largest explanatory power. However, contrary to previous studies, we find that on average smaller firms are more efficient than larger ones. Moreover, our results support the view that size is not important in explaining the development of efficiency. Fourth, location is an important factor which influences productive efficiency. Fifth, the explanatory power of other firm characteristics, such as R&D intensity, outsourcing activity, and legal form, is relatively small. Most remarkably, we find a negative effect of R&D intensity on efficiency, albeit with very low explanatory power. However, R&D appears to positively affect the development of efficiency over time. Furthermore, some types of outsourcing activities have a positive impact on productive efficiency but, again, with rather low explanatory power. Finally, although the results show that the efficiency of many firms increases or decreases over time, the average efficiency of all firms taken together does not change over time, since positive and negative efficiency changes across firms appear to cancel each other out.

Overall, our findings provide a number of novel insights into the factors that determine the productive efficiency of a firm. In particular, they indicate the relative importance of different influences. Given the heterogeneity of firms in a certain industry, it is quite surprising that industry affiliation explains such a large share of the efficiency differences while many of the firm-specific factors turn out to be relatively unimportant. This could mean that the internal factors are, indeed, comparatively unimportant, but it could also be regarded as an indication that the variables of our relatively rich dataset do not adequately reflect the management decisions that are relevant to a firm's productive efficiency. The effects of factors such as industry affiliation, size, and location deserve further investigation in order to discover the mechanisms behind these effects, which will require additional in-depth micro-level analyses. The influence of R&D effort on efficiency is in particular need of further analysis.

References

Agell, Jonas (2004) 'Why are small firms different? Managers' views.' *Scandinavian Journal of Economics* 106(3), 437–452

Aghion, Philippe, and Peter Howitt (1992) 'A model of growth through creative destruction.' *Econometrica* 60(2), 323–351

Aigner, Dennis, C.A. Knox Lovell, and Peter Schmidt (1977) 'Formulation and estimation of stochastic frontier production function models.' *Journal of Econometrics* 6(1), 21–37

Albach, Horst (1980) 'Average and best-practice production functions in German industry.' *The Journal of Industrial Economics* 29(1), 55–70

Alvarez, Roberto, and Gustavo Crespi (2003) 'Determinants of technical efficiency in small firms.' *Small Business Economics* 20, 233–244

Antonelli, Christiano (2003) *The Economics of Innovation New Technologies and Structural Change* (London): Routledge)

Audretsch, David B. (1999) ‘Small firms and efficiency.’ In *Are Small Firms Important? Their Role and Impact,* ed. Zoltan J. Acs (Springer) pp. 241–272

Baltagi, Badi H. (2001) *Econometric Analysis of Panel Data, 2d ed.* (John Wiley and Sons)

Baptista, Rui, and Peter Swann (1998) ‘Do firms in clusters innovate more?’ *Research Policy* 27, 525–540

Bartelsman, Eric J., and Mark Doms (2000) ‘Understanding productivity: Lessons from longitudinal microdata.’ *Journal of Economic Literature* 38(3), 569–594

Battese, George E., and Tim Coelli (1995) ‘A model for technical inefficiency effects in a stochastic frontier models.’ *Empirical Economics* 20, 325–332

Beeson, Patricia E., and Steven Husted (1989) ‘Determinants of productive efficiency in state manufacturing.’ *Journal of Regional Science* 29(1), 15–28

Carlsson, Bo (1972) ‘The measurement of efficiency in production: An application to Swedish manufacturing industries.’ *Swedish Journal of Economics* 74(4), 468–485

Carree, Martin A (2002) ‘Technological inefficiency and the skewness of the error component in stochastic frontier analysis.’ *Economic Letters* 77, 101–107

Caves, Richard E. (1989) ‘International differences in industrial organization.’ In *Handbook of Industrial Organization,* ed. Richard Schmalensee and Robert Willig, vol. 2 of *Handbook of Industrial Organization* (Elsevier) chapter 21, pp. 1225–1250

_ (1992) ‘Determinants of technical efficiency in australia.’ In *Industrial Efficiency in Six Nations,* ed. Richard E. Caves (MIT Press) pp. 241–272

Caves, Richard E., and David R. Barton (1990) *Efficiency in U.S. manufacturing industries* (Cambridge (Mass.): MIT Press)

Chambers, Robert G. (1988) *Applied Production Analysis: A Dual Approach* (New York: Cambridge University Press)

Coelli, Tim, D.S. Prasada Rao, and George E. Battese (2002) *An Introduction to Efficiency and Productivity Analysis* (London: Kluwer Academic Publishers)

Cooke, Philip N., Martin Heidenreich, and Hans-Joachim Braczyk (2004) *Regional Innovation Systems: The Role of Governance in a Globalized World, 2nd edition* (London: Routledge)

Cubbin, John S., and Paul A. Geroski (1987) ‘The convergence of profits in the long run: Interfirm and inter-industry comparisons.’ *Journal of Industrial Economics* 35, 427–442

Debreu, Gérard (1951) ‘The coefficient of resource utilization.’ *Econometrica* 19(3), 273–292

Farrell, Michael J. (1957) ‘The measurement of productive efficiency.’ *Journal of the Royal Statistical Society. Series A (General)* 120(3), 253–290

Flury, Bernhard W. (1989) 'Understanding partial statistics and redundancy of variables in regression and discriminant analysis.' *The American Statistician* 43(1), 27–31

Fritsch, Michael, and Andreas Stephan (2004a) 'The distribution and heterogeneity of technical efficiency within industries: An empirical assessment.' Discussion Paper 453, DIW Berlin

_ (2004b) 'What causes cross-industry differences of technical efficiency?—An empirical investigation.' Discussion Paper 457, DIW Berlin

Fritsch, Michael, Bernd Görzig, Ottmar Hennchen, and Andreas Stephan (2004) 'Cost structure surveys in Germany.' *Journal of Applied Social Science Studies* 124, 557–566

Funke, Michael, and J. Rahn (2002) 'How efficient is the East German economy? an exploration with micro data.' *The Economics of Transition* 10(1), 201–223

German Federal Statistical Office (various years) *Cost Structure Census (Kostenstrukturerhebung im Verarbeitenden Gewerbe),* vol. Fachserie 4, Reihe 4.3 (Stuttgart: Metzler-Poeschel)

Greene, William H. (1997) 'Frontier production functions.' In *Handbook of Applied Econometrics, vol. II,* ed. M. Hashem Pesaran and Peter Schmidt (Blackwell Publishers) pp. 81–166

_ (2003) *Econometric Analysis* (New Jersey: Prentice Hall)

_ (2005) 'Reconsidering heterogeneity in panel data estimators of the stochastic frontier model.' *Journal of Econometrics* 126, 269–303

Grossman, Gene, and Elhanan Helpman (1991) 'Quality ladders in the theory of growth.' *Review of Economic Studies* 58, 43–61

_ (2005) 'Outsourcing in a global economy.' *Review of Economic Studies* 72, 135–159

Görzig, Bernd, and Andreas Stephan (2002) 'Outsourcing and firm-level performance.' Discussion Paper 309, DIW Berlin

Gumbau-Albert, Mercedes, and Joaquiin Maudos (2002) 'The determinants of efficiency: the case of the Spanish industry.' *Applied Economics* 34, 1941–1948

Hamilton, David (1987) 'Sometimes $r^2 > r^2_{yx1} + r^2_{yx2}$: Correlated variables are not always redundant.' *The American Statistician* 41(2), 129–132

Helpman, Elhanan (1992) 'Endogenous macroeconomic growth theory.' *European Economic Review* 36, 237–276

Heshmati, Almas (2003) 'Productivity growth, efficiency and outsourcing in manufacturing and service industries.' *Journal of Economic Surveys* 17(1), 79–112

Hoskisson, Robert E., Richard A. Johnson, and Douglas D. Moesel (1994) 'Corporate divestiture intensity in restructuring firms: Effects of governance, strategy, and performance.' *The Academy of Management Journal* 37(5), 1207–1251

Jondrow, James, C. A. Knox Lovell, Ivan S. Materov, and Peter Schmidt (1982) 'On the estimation of technical inefficiency in the stochastic frontier production function model.' *Journal of Econometrics* 19, 233–238

Klepper, Steven (1997) 'Industry life cycle.' *Industrial and Corporate Change* 6, 145–181

Klepper, Steven, and Kenneth Simons (2005) 'Industry shakeouts and technological change.' *International Journal of Industrial Organization* 23, 23–43

Koopmans, T. C. (1951) 'An analysis of production as an efficient combination of activities.' In *Activity Analysis of Production and Allocation,* ed. T. C. Koopmans (New York: Wiley)

Krugman, Paul (1991) *Geography and Trade* (Cambridge: MIT-Press)

Kumar, Jayesh (2003) 'Ownership structure and corporate firm performance.' Technical Report 0304004, EconWPA

Kumbhakar, Subal C., Almas Heshmati, and Lennart Hjalmarsson (1999) 'Parametric approaches to productivity measurement: A comparison among alternative models.' *Scandinavian Journal of Economics* 101(3), 405–424

Kumbhakar, Subal C., and C.A. Knox Lovell (2003) *Stochastic Frontier Analysis* (Cambridge University Press)

Leibenstein, Harvey (1966) 'Allocative efficiency vs. "x-efficiency".' *American Economic Review* 56(3), 392–415

Li, Yang, and Jin-Li Hu (2004) 'Technical efficiency and location choice of small and medium-sized enterprises.' *Small Business Economics* 19(1), 1–12

Lovell, C.A. Knox (1993) 'Production frontier and productive efficiency.' In *The Measurement of Productive Efficiency. Techniques and Applications,* ed. Harold O. Fried, C.A. Knox Lovell, and Shelton S. Schmidt (Oxford: Oxford University Press) pp. 3–67

Mairesse, Jacques, and Jordi 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

Mayes, David, Melanie Lansbury, and Christopher Harris (1995) *Inefficiency in Industry* (Pearson Education)

McCloskey, Deirdre N., and Stephen T. Ziliak (1996) 'The standard error of regressions.' *Journal of Economic Literature* 34(1), 97–114

Meeusen, Wim, and Julien van den Broeck (1977) 'Efficiency estimation from Cobb-Douglas production functions with composed error.' *International Economic Review* 18(2), 435–444

Oczkowski, Edward, and Kishor Sharma (2005) 'Determinants of efficiency in least developed countries: Further evidence from nepalese manufacturing firms.' *Journal of Development Studies* 41(4), 617–630

Ornaghi, Carmine (2006) 'Spillovers in product and process innovation: Evidence from manufacturing firms.' *International Journal of Industrial Organization* 24, 349–380

Porter, Michael E. (1998) 'Clusters and the new economics of competition.' *Harvard Business Review* (November–December), 77–90

— (2003) ‘The economic performance of regions.’ *Regional Studies* 37, 549–578

Roudaut, Nolwenn (2006) ‘Influences of the business environment on manufacturing firms technical efficiencies: The côte d’ivoire case.’ *Journal of Productivity Analysis* 25(1-2), 93–109

Schmalensee, Richard (1985) ‘Do markets differ much?’ *American Economic Review* 75(3), 341–351

Schmidt, Peter, and Robin C. Sickles (1984) ‘Production frontier and panel data.’ *Journal of Business and Economic Statistics* 2(4), 367–374

Schmidt, Peter, and Tsai-fen Lin (1984) ‘Simple tests of alternative specifications in stochastic frontier models.’ *Journal of Econometrics* 24(3), 349–361

Shea, John (1997) ‘Instrument relevance in multivariate linear models: A simple measure.’ *The Review of Economics and Statistics* 79(2), 348–352

Sickles, Robin C. (2005) ‘Panel estimators and the identification of firm-specific efficiencylevels in parametric, semiparametric and nonparametric settings.’ *Journal of Econometrics* 126, 305–334

Söderbom, Måns, and Francis Teal (2004) ‘Size and efficiency in African manufacturing firms: Evidence from firm-level panel data.’ *Journal of Development Economics* 73(1), 369–394

Theil, Henri (1971) *Principles of Econometrics* (Santa Barbara: Wiley and Sons)

Torii, Akio (1992) ‘Dual structure and differences of efficiency between Japanese large and small enterprises.’ In *Industrial Efficiency in Six Nations*, ed. Richard E. Caves (Cambridge (Mass.): MIT Press) pp. 385–424

Wernerfelt, Birger, and Cynthia A. Montgomery (1988) ‘Tobin’s q and the importance of focus in firm performance.’ *American Economic Review* 78(1), 246–250

Appendix

We use the value of gross production net of subsidies and excise taxes as a measure of output. This mainly comprises the turnover plus the net change in the stock of final products. We do not include turnover from goods for resale or from activities that are classified as miscellaneous, such as license fees, commissions, rents, leasing, and etc., because we assume that such revenue cannot adequately be explained on the basis of a production function.

Median production shares of these input categories and other descriptive statistics are reported in Table A.1. The dominant categories are material inputs and payroll, the median values of which add up to about 73 percent of the overall expenses. The median values of the shares sum up to 92.4 percent. The difference to unity of approximately 7.6 percent can be interpreted as the share of gross profits in production. Since firms with less than 500 employees are included in the Cost Structure Census only as a representative random sample, we use weights

Table A.1: Production shares of inputs – descriptive statistics

Variable	Min	p1	Median	p99	Max
Material inputs	6.00E-07	0.013	0.382	0.855	661
Labor compensation	3.00E-03	0.059	0.349	0.957	2177
Energy consumption	0	0.001	0.014	0.180	325
Capital	9.00E-09	0.009	0.061	0.312	377
External services	2.00E-06	0.001	0.031	0.361	188
Other inputs	3.00E-05	0.010	0.087	0.472	329

Notes: p1 and p99 are the 1^{st}, and 99^{th} percentiles.
Number of observations 219,293

greater than or equal to one for estimating production for the firms in these size categories. Each of these firms is multiplied by a factor that represents the relationship between the number of firms in the respective industry and size that is included in our sample and the number of firms in an industry and size category in the full population.[19] Since these weights are rather stable over time, we use the weights for 1997 in all the estimations.

Some of the cost categories, including expenditure for external wage-work and for external maintenance and repair, contain a relatively high share of reported zero values since many firms do not utilize these types of input. Since all inputs in a translog production function are included in logarithms, such zero values for certain input categories would lead to missing values and result in the exclusion of the respective firm from the analysis. Moreover, zero input values are not consistent with a translog production technology and would imply zero output. To reduce the number of reported zero input values, we aggregated the inputs into the following broader categories: material inputs (intermediate material consumption), labor compensation (salaries and wages plus employer's social insurance contribution), energy consumption, capital input (depreciation of fixed assets plus rents and leases), external services (e.g., repair costs and external wage-work), and other inputs related to production (e.g., transportation services, consulting, or marketing). All input and output series were deflated using the producer price index for the respective industry. Table A.2 presents the basic descriptive statistics for logarithmic values of all output and input categories.

The yearly values of the depreciations as a proxy for capital input led to a rather low estimate for the elasticity of the capital input. The obvious reason for this low value is the relatively high year-to-year variation of the depreciations. To reduce this volatility, we calculated the average yearly depreciations by adding up the depreciations in the current year and for all the preceding years that we have in our data. This sum was then divided by the number of respective years.[20]

[19] As an example, if only 25 percent of the firms of a particular size class are included in the sample, each observation is multiplied by a factor of 4.

[20] Example: Assume that the dataset provides information on depreciations of a certain firm for 1993, 1994, 1995, and 1996. Average yearly depreciation for 1995 is the average for 1993-1995. For 1996, it is the average for 1993-1996, etc. For 1993, the average equals the value for this year.

Table A.2: Descriptive statistics of inputs and outputs (in logs)

Variable	Mean	St. Dev	CV	Min	Q1	Median	Q3	Max
Output	16.89	1.49	8.85	12.36	15.78	16.71	17.85	25.29
Material inputs	15.77	1.84	11.64	0	14.58	15.71	16.96	24.91
Labor compensation	15.76	1.36	8.64	10.29	14.73	15.55	16.58	23.97
Energy consumption	12.65	1.72	13.61	0.61	11.39	12.47	13.75	22.25
Capital	14.07	1.51	10.70	8.94	13.02	13.94	15.00	22.50
External services	13.38	2.01	15.00	2.87	11.99	13.36	14.73	22.34
Other inputs	14.40	1.76	12.22	7.93	13.14	14.28	15.54	23.38

Notes: CV is the coefficient of variation; Q1 and Q3 are lower and upper quantiles.

Such average values of yearly depreciation result in a considerably higher estimate of the output elasticity of capital. We are aware that using a proxy variable instead of a direct measure of the capital stock input could be of concern. However, even with such a crude proxy based on the tax depreciations for the capital input, we obtain estimates of the elasticity of capital that appear to be quite reasonable.

The sample contains a number of observations with extreme values (see maximum and minimum columns in Table A.2) that proved to have a considerable impact on the estimated parameters of the production function and led to implausible results. Therefore, we exclude such "outliers" from the analysis when the cost for a certain input category in relation to gross value added is less than the lowest 0.1 percent and the highest 99.9 percent. In total, these excluded cases plus firms with zero values for at least one input category (the major part of excluded cases) account for about 10 percent of all observations. We find that the exclusion of these extreme cases leads to a considerable improvement of robustness and plausibility of the estimation results for the production function.