ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Fritsch, Michael; Stephan, Andreas

Working Paper

The Distribution and Heterogeneity of Technical Efficiency within Industries: An Empirical Assessment

DIW-Diskussionspapiere, No. 453

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Fritsch, Michael; Stephan, Andreas (2004) : The Distribution and Heterogeneity of Technical Efficiency within Industries: An Empirical Assessment, DIW-Diskussionspapiere, No. 453

This Version is available at:
http://hdl.handle.net/10419/18305

Nutzungsbedingungen:
Die ZBW räumt Ihnen als Nutzerin/Nutzer das unentgeltliche, räumlich unbeschränkte und zeitlich auf die Dauer des Schutzrechts beschränkte einfache Recht ein, das ausgewählte Werk im Rahmen der unter
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
nachzulesenden vollständigen Nutzungsbedingungen zu vervielfältigen, mit denen die Nutzerin/der Nutzer sich durch die erste Nutzung einverstanden erklärt.

Terms of use:
The ZBW grants you, the user, the non-exclusive right to use the selected work free of charge, territorially unrestricted and within the time limit of the term of the property rights according to the terms specified at
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
By the first use of the selected work the user agrees and declares to comply with these terms of use.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft

Discussion Papers

453

Michael Fritsch
Andreas Stephan

The Distribution and Heterogeneity of Technical Efficiency within Industries – An Empirical Assessment

Berlin, November 2004

DIW Berlin
German Institute for Economic Research

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

DIW Berlin

German Institute
for Economic Research

Königin-Luise-Str. 5
14195 Berlin,
Germany

Phone +49-30-897 89-0
Fax +49-30-897 89-200

www.diw.de

ISSN 1619-4535

The Distribution and Heterogeneity of Technical Efficiency within Industries – An Empirical Assessment*

Michael Fritsch†

Andreas Stephan††

October 2004

† Technical University Bergakademie Freiberg, German Institute for Economic Research (DIW Berlin) and Max-Planck Institute for Research into Economic Systems, Jena, Germany.

†† European University Viadrina Franfurt/Oder and German Institute for Economic Research (DIW Berlin), Germany.

Addresses for correspondence:

Prof . Dr. Michael Fritsch
Technical University of Freiberg
Faculty of Economics and
Business Administration
Lessingstraße 45
D-09596 Freiberg, Germany
Phone: +49 / 3731 / 39 24 39
Fax: +49 / 3731 / 39 36 90
michael.fritsch@tu-freiberg.de

Prof. Dr. Andreas Stephan
German Institute for Economic
Research (DIW Berlin)

Königin-Luise Straße 5
D-14191 Berlin, Germany
Phone: +49 / 030 / 897 89 - 325
Fax: +49 / 030 / 897 89 - 103
astephan@diw.de

---

* This paper is based on a cooperation project of the German Institute for Economic Research (DIW Berlin) and the Federal Statistical Office Germany in Wiesbaden which enabled access to confidential firm-level micro data of the annual Cost Structure Census. We are grateful to Gerald Göbel, Ottmar Hennchen and Roland Sturm at the Federal Statistical Office, Germany for their support of this project. We thank Adrianna Kaminiarz for excellent research assistance. We have benefited from discussions of earlier versions at the EARIE 2003 conference in Helsinki, the 2003 meeting of the German Economic Association (Verein für Socialpolitik) in Zurich, the 2003 FiDaSt meeting in Berlin and at workshops at the Technical University of Freiberg and at the DIW Berlin.

Contents

Abstract

This paper analyzes the distribution of technical efficiency within manufacturing industries. Using a representative sample of 35,000 firms in 255 industries of the German cost structure census, technical efficiencies are estimated by applying a deterministic frontier production function with firm-specific fixed effects. A new measure is also introduced for characterizing the extent of heterogeneity within an industry that is robust with regard to extreme values of a few small firms. It was found that the level of intra-industry heterogeneity is mainly determined by an industries' average technical efficiency, average firm size, capital intensity and the rate of new firm formation. Most strikingly, we find that in about 95 percent of industries the distribution of technical efficiency is skewed to the right, not to the left as is commonly assumed.

Keywords: Technical efficiency, heterogeneity, deterministic production function frontier.

JEL classification: D24, L10, L11

Zusammenfassung

"Die Verteilung und Heterogenität von technischer Effizienz innerhalb von Branchen – Eine empirische Untersuchung"

Diese Arbeit analysiert die Verteilung von technischer Effizienz innerhalb von Branchen des Verarbeitenden Gewerbes. Die technische Effizienz wird als firmenspezifischer fixer Effekt im Rahmen einer deterministischen Frontier-Produktionsfunktion ermittelt. Als Grundlage hierfür dienen Angaben aus der Kostenstrukturstatistik für ein repräsentatives Sample von 35.000 Unternehmen in 255 Branchen. Zur Analyse der Heterogenität innerhalb der Branchen entwickeln wir ein neues Maß, das sich als relativ robust hinsichtlich einzelner extremer Werte erweist. Das Ausmaß an Heterogenität von technischer Effizienz innerhalb von Branchen wird im Wesentlichen vom durchschnittlichen Niveau an technischer Effizienz der Branche, der durchschnittlichen Unternehmensgröße, der durchschnittlichen Kapitalintensität und dem Ausmaß an Gründungen neuer Firmen in der Branche geprägt. Bemerkenswert ist insbesondere, dass die Verteilung der technischen Effizienz in ca. 95 Prozent der Branchen einen positiven Wert für die Schiefe aufweist (rechtsschief). Dies steht im Widerspruch zu der üblichen Annahme, dass die Schiefe negativ (linksschief) sei.

Schlagworte: Technische Effizienz, Heterogenität, Deterministische Frontier-Produktionsfunktion.

JEL-Klassifikation: D24, L10, L11

1. Introduction

Contrary to most textbook models, firms in real world are quite heterogeneous. One particular aspect of this heterogeneity is that even within the same industry firms are not equally technically efficient. Technical efficiency is defined as the highest attainable output a firm can produce given its inputs. While some firms operate at the technological frontier and potentially earn high profits, others lag considerably behind and are hardly able to survive.[1] The objective of this paper is to analyze the extent and the distribution of technical efficiency within different industries. Are certain industries more homogeneous than others with regard to the distribution of firm-level technical efficiency? What causes the observed heterogeneity? For example, can industries' average firm size or the average capital intensity or the rate of new firm formation explain the extent of heterogeneity?

Both the extent and distribution of technical efficiency may have important implications for competition and market evolution. If, for example, a few firms have a major efficiency advantage while the rest of the industry is operating at much higher costs, we may expect increasing market shares of the highly efficient firms. On the other hand, if the main competing firms in an industry operate at about the same level of technical efficiency, one could expect that competitive pressure in such an industry is lower compared to an industry with a few highly efficient leading firms. For these reasons an analysis of the distribution of technical efficiency within industries will provide important insights into the dynamics of competition.

Our study contributes to the literature on the distribution of efficiency within industry in several respects (see for instance Caves and Barton, 1990; Mayes, 1996). First, our approach for measuring heterogeneity is based on a new measure that is rather robust with respect to extreme values of a few small

[1] There are a few empirical studies showing high levels of heterogeneity among and within industries. See, for example, Caves and Barton (1990), Mayes, Harris and Lansbury (1994) and the contributions in Caves (1992) and Mayes (1996). For a survey, see Caves and Barton (1990, pp. 15-20) and Caves (1992).

firms. Second, in contrast to most previous studies we apply a deterministic production frontier model and estimate technical efficiency with firm-specific fixed effects. This has the advantage over stochastic frontier models that no a-priori assumption on the distribution of technical efficiency within industries has to be made. An analysis of the distribution of efficiencies appears to be more sensible without a-priori distributional assumptions. Third, in our econometric analysis on the determinants of efficiency heterogeneity we apply robust econometric methods to check the sensitivity of results with respect to extreme observations and functional form specification.

The remainder of the paper is organized as follows. Section 2 compares the deterministic and the stochastic frontier approach. Section 3 describes the data used in the analysis. Section 4 depicts estimates of the production function that serves as a basis for our measurement of technical efficiency. Furthermore, it analyzes the distribution of efficiency within industries. Section 5 introduces a two-dimensional approach for measuring the extent of within-industry heterogeneity of technical efficiency. Hypotheses regarding the determinants of heterogeneity are formulated and the econometric analyses are presented in Section 6. Finally, Section 7 summarizes the main findings and concludes the paper.

2. Frontier models for measuring (in)efficiency

Technical efficiency is defined as the generation of maximum output from a given amount of resources. A firm is technically *inefficient* if it fails to obtain the maximum possible output.[2] Note that another important type of efficiency – allocative efficiency – concerns the optimal choice of inputs.[3] Reasons for technical inefficiency can be manifold and comprise all kinds of

---

[2] The concept of technical inefficiency was introduced by Farrell (1957).

[3] A firm is allocatively efficient if its input combination is optimal, given input prices and marginal productivities. A firm can be allocatively efficient, but at the same time technically inefficient, if it chooses an optimal input combination but does not attain the highest possible isoquant of its production function.

'mismanagement' like inappropriate work organization and use of technology (cf. Fritsch and Mallok, 2002), bottlenecks with regards to material flows, etc. It can also be attributed to X-inefficiency as exposed by Leibenstein's (1966) seminal work.

An assessment of technical (in)efficiency of firms or industries requires efficiency to be measured as well as a point of reference for the relative efficiency level of the unit under inspection to be identified. This can be done in a number of ways (see Mayes, Harris and Landsbury, 1994, pp. 27-54, for an overview). What all these approaches have in common is that they define technical efficiency as the highest output level that can be attained with a given combination of inputs. Any deviation from this maximum is then regarded as inefficiency. If data about various inputs is available, the maximum technical efficiency of an industry can be directly obtained by estimating a frontier production function, i.e. a function for the input-output relationship.

To our knowledge, almost all of the analyses using this approach estimated a stochastic form of a frontier production function. A stochastic frontier production function is based on the assumption that the input-output relationship is not completely deterministic, but subject to influences that appear to be random noise.

The general form of a frontier production function for industry l can be written as

$$\text{(1)} \qquad \ln\ y_{il} = \ln\ f(x_{il},\ \beta_{il})\ +\ \varepsilon_{il}$$

where y_{il} denotes the output that firm i in industry l is producing using inputs x_{il}. The unknown parameters are represented by β_{il}, f describes the functional form and ε_{il} denotes an error term. Technical inefficiency is identified by decomposing the error term of the stochastic frontier production function into two components, i.e. $\varepsilon_{il}\ =\ u_{il}\ -\ v_{il}$. One component ($u_{il}$) reflects the random disturbances that is assumed to follow a symmetric normal distribution. The second component (v_{il}) is an asymmetrically distributed, negative error term that represents the technical inefficiency. Accordingly a firms' output lies on or below the stochastic frontier.

Assessing technical inefficiency on the basis of a stochastic frontier production function has the advantage that extreme outliers of highly efficient firms do not automatically serve as efficiency benchmarks. This is particularly important if the extreme values are due to measurement error. However, in order to separate the impact of technical inefficiency from the general stochastic effects, an a priori assumption about the distribution of technical inefficiency is required. Since the factual efficiency of a firm cannot exceed the maximum, the distribution must be truncated at this maximum. The usual hypothesis in this respect is that most firms cluster near the efficiency frontier and that their frequency decreases with rising inefficiency. Such a distribution of the v_{il} is negatively skewed and may be described as a truncated normal or a log-normal distribution.[4]

The usual rationale of this assumption is that the distribution of technical efficiency is truncated at the efficiency frontier because this frontier represents the attainable maximum. This argument may be illustrated by means of a simple graph. We assume that the original distribution of technical efficiency follows a symmetric normal distribution with asymptotic tails. Any truncation of such a distribution at an upper threshold, such as an alleged maximum of technical efficiency, results in a distribution that is negatively skewed as shown by the shaded area in figure 1a. One may, however, also expect there to be a lower threshold for observed technical efficiency resulting in a truncated

---

[4] If the values of a distribution are in increasing order from the left to the right, a negatively skewed distribution has the longer tail at the left side (skewed "to the left"), where the values are below the median. If the distribution is positively skewed the longer tail is at the right side (skewed "to the right") with the values above the median. Measures for the skewness of this distribution can then be used as indicators for the level of technical inefficiency in the respective industry (cf. Caves and Barton, 1990, pp. 47-49; Mayes, Harris and Lansbury, 1994, pp. 50-52). If the distribution of residuals is not skewed but symmetric, the level of technical inefficiency in the respective industry is assumed to be insignificant. A positively skewed distribution of residuals is not consistent with the underlying assumptions. In this case, the stochastic production function approach of measuring technical efficiency is inappropriate and may, therefore, be misleading.

normal distribution that is positively skewed (figure 1b). An important lower threshold for firms is the necessity to be sufficiently efficient for surviving market competition. Those firms which are not sufficient technically efficient will exit the market and fall out of the efficiency distribution. We presume that the lower threshold is determined by costs, in particular by labor cost.

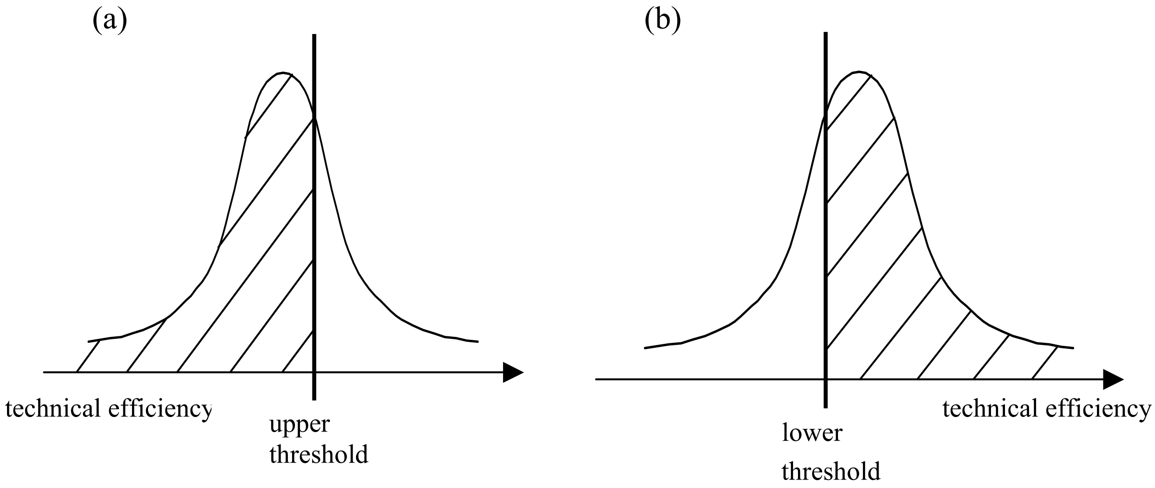

Figure 1: Distribution of technical efficiency truncated at an (a) upper and (b) lower threshold

Considering this, an important advantage of our approach over estimating a stochastic frontier production function is that we do not need to specify a priori a particular functional form (e.g., a certain direction of skewness) for the distribution of inefficiency (cf. Schmidt and Sickles, 1984). We can, therefore, analyze the distributional properties of inefficiency within industries without any restrictions.

For measuring technical efficiency we apply a deterministic production function using a panel of firms. The production function is of the Cobb-Douglas type and, in its logarithmic form can be written as (cf. Greene, 1997)

(2) $$\ln y_{ilt} = \ln \alpha_{il} + \lambda_{ilt} + \sum \beta_{kl} \ln x_{kilt} + \varepsilon_{ilt},$$
$$i = 1,\ldots,N;\, l = 1,\ldots,l;\, k = 1,\ldots,p;\, t = 1,\ldots,T.$$

The term y_{ilt} represents the output of firm i in industry l in period t, x_{kilt} denotes the production input k, β_{kl} indicates the industry-specific elasticities of production for the different inputs, λ_{lt} represents a time-specific effect, and α_{il} stands for the technical efficiency of a specific firm in industry l. There are N firms and T_i observations for each firm. We estimate technical efficiency as firm-specific fixed effect. According to our approach, the largest estimate of technical efficiency $\hat{\alpha}_{jl}$ within a certain industry l is used as a benchmark value. An estimate of the technical efficiency $\widehat{TE}_{il}$ of the firm i in industry l is then calculated as

$$(3) \qquad \widehat{TE}_{il} = (\hat{\alpha}_{il} / \max\ \hat{\alpha}_{jl}) \cdot 100 \quad [\%].$$

According to this approach, at least one firm in industry l will meet the benchmark value and the remaining firms will have positive efficiency estimates between 0 and 100 percent.

This approach of estimating technical inefficiency as a firm-specific fixed effect implies that the relative efficiency level of a firm is invariant over time. This may be regarded as a critical assumption because a firm's efficiency level is likely to change over time. Principally, it would be possible to estimate time varying firm efficiencies with a deterministic approach (see for instance Heshmati, Kumbhakar and Hjalmarsson (1995)). However, this requires a sufficient number of observations for each firm. In our sample, more than eighty percent of the firms have only five or less observations (see table 1), which renders estimation of a time-varying firm-specific effects inapplicable.

Generally, one might suspect that the estimates of technical efficiencies would be correlated with the number of observations for each firm. To check this matter, we computed the correlation between firm-specific fixed effects (as deviations from the industry's median level of efficiency) and the number of observations which were used to estimate the firm-specific effect. Indeed, we find a statistically significant – albeit quite low - correlation of -0.04, which indicates that the estimated efficiency levels of firms with a larger number of observations tend to be closer to the average than of firms with fewer observations. However, taking into account the rather small value, and the small fraction of cases in our sample with more than 5 observations, this effect can be largely neglected.[5]

Finally, a major advantage of the deterministic approach is that it does not require the rather strong assumption that the α_i values are uncorrelated with the explanatory variables. In fact, we can show that in our sample there is a considerable correlation between estimated technical efficiency and the factor inputs. This correlation would yield inconsistent parameter estimates in the stochastic production frontier framework.

3. Data and variable definitions

We utilize data of the German Cost Structure Census[6] of manufacturing for the period 1992-2001. This survey is conducted by the German Federal Statistical Office (Statistisches Bundesamt). It comprises almost all large German manufacturing firms with 500 or more employees.[7] Firms with 20-499 employees are included as a random sample which is representative for the respective size category and industry.[8] Firms with less than 20 employees are not sampled.[9] Usually, smaller firms report for four subsequent years and are then substituted by other small firms (rotating panel).[10] Since the estimation of

[5] Moreover, since only firms with 500 or more employees in our data set are likely to have more than five observations, the effect may well be caused by respective size differences. If such an impact of size really applies to our data, this means that the extreme values of technical efficiency tend to be due to the smaller firms with only a few observations included in our data.

[6] Aggregate figures are published annually by the German Federal Statistical Office in Statistisches Bundesamt, Fachserie 4, Reihe 4.3, Kostenstrukturerhebung im Verarbeitenden Gewerbe.

[7] Firms are legally obliged to participate in the survey and have no legal means of holding back any of the information required.

[8] This is done to keep the reporting effort of smaller firms at a reasonable level.

[9] Since the year 2001, the statistic also contains firms with 1-19 employees. These firms are, however, not included in our analysis because there was only one observation for these firms at the time this analysis was conducted.

[10] Due to mergers or insolvencies, some firms have less than four observations. Note, however, that firms are legally obliged to respond to the Cost Structure survey, so there are actually almost no missing observations due to non-response. On the other hand, some of the

firm-specific fixed effects requires at least two observations, firms with only one observation are not included in our sample that comprises a total of about 35,000 firms. Table 1 shows the frequency of firms with different numbers of observations.

Table 1: Frequency of firms with regard to the numbers of observations in the sample

Number of observations (years)	*Number of firms*	*Share of all firms (percent)*	*Cumulated share of all firms (percent)*
2	11,248	32.14	32.14
3	7,756	22.16	54.30
4	2,682	7.66	61.97
5	6,929	19.80	81.77
6	1,485	4.24	86.01
7	1,554	4.44	90.45
8	1,341	3.83	94.28
9	422	1.21	95.49
10	1,579	4.51	100
Total	34,996	100	–

We use gross production as measure of output. Gross production comprises the turnover plus the net change of the stock of final products. Turnover from activities that are classified as miscellaneous such as license fees, commissions, rents and leasing, etc. is excluded because we assume that such revenue would be inadequately described by means of a production function.

The Cost Structure Census contains information for a large number of input categories. These categories are payroll, employers' contribution to the social security system, fringe benefits, expenditure on material inputs, self-

smaller firms have more than four observations. This is more likely in industries with a low number of firms.

provided equipment and goods for resale, energy, external wagework, external maintenance and repair, tax depreciation of fixed assets, subsidies, rents and leases, insurance costs, sales tax, other taxes and public fees, interest payments as well as “other” costs such as license fees, bank charges, postage or expenses for marketing and transport. Further information available in the Cost Structure Census includes industry affiliation, location of headquarters, stock of raw materials, goods for resale and final output, R&D expenditure and number of R&D employees[11]. Information on employment comprises the number of active owners, the number of employees, trainees, part-time employees and home workers and the number of temporary workers.

Some of the cost categories such as expenditure on external wagework and external maintenance and repair include a relatively high proportion of reported zero values because many firms do not utilize these types of inputs. Since all inputs of the Cobb-Douglas production function are included in logarithms, these zero values lead to missing observations and accordingly to the exclusion of the respective firm from the analysis. It is worth pointing out that zero values for inputs are not compatible with a Cobb-Douglas technology because they imply zero output.

To reduce the number of reported zero input quantities, we decided to aggregate single inputs into the following categories: material inputs (intermediate material consumption plus commodity inputs), labor compensation (salaries and wages plus employer's social insurance contributions), energy consumption, user cost of capital (depreciation plus rents and leases), external services (e.g., repair costs and external wagework) 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.

It turned out that using yearly depreciations as a proxy variable for the capital input leads to implausibly low estimated production elasticities of

---

[11] Information on R&D expenditures has been included in the Cost Structure Census since 1999.

capital. One can presume that this is due to the relatively high year-by-year variations of depreciations at the firm level. In order to reduce this volatility, we calculated average annual depreciations by adding up for each year the depreciations in the current year and in all the preceding years that we observe. This sum is then divided by the respective number of observation years.[12] Taking this average value of annual depreciations we obtain considerably higher estimates of capital elasticity.

Table 2: Input cost shares in total production

Variable	*Mean*	*Median*	*Standard deviation*	*Minimum*	*Maximum*	*Coefficient of variation*
Material inputs	0.4090	0.4060	0.1651	0.0177	0.8547	40.37
Labor compensation	0.3305	0.3199	0.1367	0.0530	0.8470	41.36
Energy consumption	0.0210	0.0134	0.0232	0.0008	0.1729	110.49
User cost of capital	0.0671	0.0564	0.0427	0.0081	0.2800	63.58
External services	0.0471	0.0277	0.0530	0.0012	0.3318	112.59
Other inputs	0.0921	0.0791	0.0586	0.0096	0.3616	63.60

Average cost shares of these input categories and other summary statistics for the cost shares are reported in table 2. The dominant cost categories are material inputs and the payroll that add up to about 75 percent of all expenses. Summing all cost shares gives a total of 0.9668. The difference of about 3.3 to 100 percent can be interpreted as the average share of net profits. Note that there is substantial variation of the cost shares between the industries, indicating significant differences with regard to production technologies. To account for such industry-specific heterogeneity of production technologies, we estimated production functions for each industry separately.

[12] Example: Assume that the data set provides information on depreciations of a certain firm for the years ’93, ’94, ’95 and ’96. Average annual depreciation for ’95 is the average of the years ’93 – ’95. For the year ’96, it is the average of the years ’93 - ’96, etc. For ’93, the average is the value for that year.

For 12 out of 254 industries, no industry-level production function could be estimated due to an insufficient number of cases. Furthermore, in estimating the production functions we use a weight for each observation, which reflect the relationship between the population number of firms in a certain industry and given size category and the number of firms in the corresponding sample.[13] Weights are greater or equal to one for firms with less than 500 employees. Since these weights do not change much over time, we decided to apply the weights of 1997 for the other years as well.

Though the quality of our data is excellent and measurement error can be largely neglected, we noticed that a few observations with extreme values of inputs or output exert a significant impact on the magnitudes of estimated production function parameters. We therefore decided to exclude those extreme values from the analysis for which the cost share for a certain input category in relation to gross production is less than the lowest (1%) or greater than the highest (99%) percentile. In total, these excluded cases (plus firms with zero values for certain input categories) constitute about 10 percent of all observations. We find that the exclusion of these cases leads to a considerable improvement of stability and plausibility of obtained estimates.

4. Estimation results of the industry-specific frontier models

A Cobb-Douglas production function according to (2) has been estimated for each industry using the Least Squares Dummy Variables (LSDV) estimator for panel data (Baltagi, 2001; Coelli et al., 1998).[14] Table 3 displays the average of parameter estimates of the frontier production functions for the different

---

[13] Example: If only 25 percent of the firms are included in the statistics, each observation is multiplied by a factor of 4.

[14] Attempts to estimate other types of production function did not give satisfactory results. Estimates of a translog type of production function frequently produced rather implausible results (for example, negative production elasticities for certain inputs or estimated production elasticities larger than one). We suspect that the problems we experienced in estimating forms of production function other than the Cobb-Douglas type were caused by the relatively high number of different inputs we used and the dependence (multicollinearity) between these inputs. Non-linear forms of a production function, e.g., the CES, could not be estimated due to the computational limitations of a non-linear regression involving more than 35,000 parameters.

industries. The estimates of α_{il} represent our measures of firms' technical efficiency. The highest estimated α_{il} in an industry serves as a benchmark for determining relative technical efficiency according to Equation (3).

Overall, the fit of regression ($R^2 = 0.997$) is remarkably high. The dummies for the different years are highly significant as well as the firm-specific fixed effects. The sum of averages of the estimated output elasticities is 0.9758. According to neoclassical production theory, profit maximizing firms will choose such a combination of inputs that the input's cost shares equal the respective production elasticities. The fact that there are no large deviations between average cost shares (table 2) and average production elasticities (table 3) indicates that the parameters of our production functions are in a plausible range and that the model is properly specified.

Table 3: Average parameter estimates of industry-specific Cobb-Douglas production functions (Least Square Dummy Variable estimation)

Variable	*Average parameter estimate*		*Standard deviation*	*Minimum*	*Maximum*
	Mean	*Median*			
Material inputs	0.4177	0.4202	0.1394	0.0353	0.8245
Labor compensation	0.3505	0.3475	0.1513	-0.0446	1.0452
Energy consumption	0.0363	0.0266	0.0573	-0.2913	0.3660
User cost of capital	0.0606	0.0608	0.0950	-0.3017	0.5596
External services	0.0373	0.0379	0.0287	-0.1466	0.1807
Other inputs	0.0734	0.0707	0.0457	-0.0663	0.4078
1992 dummy	0.0069	0.0051	0.0628	-0.2329	0.2726
1993 dummy	-0.0051	-0.0039	0.0640	-0.2573	0.2421
1994 dummy	0.0029	0.00135	0.0570	-0.1600	0.2115
1995 dummy	0.0018	0.0035	0.0591	-0.5553	0.2520
1996 dummy	0.0002	0.0003	0.0552	-0.4582	0.1712
1997 dummy	0.0024	0.0023	0.0480	-0.1896	0.1418
1998 dummy	0.0068	0.0073	0.0405	-0.1581	0.2104
1999 dummy	0.0065	0.0046	0.0351	-0.1293	0.1246
2000 dummy	0.0062	0.0040	0.0248	-0.0855	0.0766

There is considerable variation in the estimated production elasticities among industries. This variation shows that, as we presumed above, industries are indeed quite heterogeneous regarding their production technologies. The positive values of most time dummies indicate a higher productivity in those years compared to the reference year 2001. Note that the year dummies do not purely measure technical progress but also reflect macroeconomic conditions. These were relatively unfavorable with a considerable underutilization of capacities in 2001 and also in 1993, for which the obtained estimates of the respective dummy variable is negative.

Our measure of technical efficiency describes a firm's performance in relation to the most efficient firm in the respective industry. Therefore, low efficient firms in industries with no comparably highly efficient firm will appear to perform relatively well. By estimating a common production function for all industries in our sample, it is possible to generate a measure for a firms' efficiency in relation to the most efficient firm for the total manufacturing sector. It turns out that there is some correspondence between both measures of technical efficiency, though the correlation (Pearson correlation coefficient is 0.403 and a (Spearman) rank correlation is 0.392) is not very high. The correlation coefficient between the technical efficiency level of firms and their operating surplus over sales is about 0.6. This indicates that a high level of technical efficiency may lead to relatively high profits but that there are also other factors that determine profitability. According to such other determinants technically low efficient firms may gain relatively high profits (e.g. because of market power) and highly efficient firms may not succeed to operate with a positive surplus (e.g. due to high competitive pressure and decreasing demand).

We find variance of technical efficiency between firms in all industries of our sample. This means that there is no industry without at least some degree of technical inefficiency. It is remarkable that the distribution of technical efficiency is positively skewed in about 95 percent of the industries in our sample (230 out of 242). In 77.7 percent of the industries (188 out of 242

Table 4: Share of industries with positive or negative skewness of the technical efficiency distribution

Skewness	Share (percent) / number of industries		
	All industries	Statistically significant at the 5 percent level	Statistically significant at the 1 percent level
Negative	5.0 /12	0.4 / 1	0 / 0
Positive	95.0 / 230	77.7 / 188	69.8 / 169

industries), the positive skewness of the distribution is even statistically significant at the 5 percent level. In 69.8 percent, it is even statistically significant at the 1 percent level. Only for one of the 242 industries is the negative skewness statistically significant at a 5 percent level and for none is it significant at a 1 percent level (table 4). The average value of the skewness measure for the industries of our sample is 1.639 (the median is 1.225) with a maximum of 9.139 and a minimum value of -1.735. This result clearly shows that for almost all of the industries the usual assumption of negatively skewed distribution of efficiency is rejected. As we have emphasized above, using a stochastic frontier model for the assessment of technical efficiency would, therefore, be inappropriate.

5. Assessing the heterogeneity of technical efficiency within industries

In this section we describe a new approach for describing the heterogeneity of technical efficiency within industries. We start with a graphical representation of the efficiency distribution curve. In the following we show examples of efficiency distribution for selected industries.

5.1 Graphical representation of the efficiency distribution curve

Figure 2 shows a graphical exposition of a (fictive) sample of firms in a particular industry with diverging efficiency levels.[15] In this graph the firms are arranged according to their efficiency in descending order, starting with the most efficient firm. This most efficient firm constitutes the 100 percent benchmark for measuring relative technical efficiency of the other firms in the respective industry; that is, efficiency of a firm is measured in relation to the value of the most efficient firm that represents the 100 percent value in this distribution. The length of the line for each firm corresponds to the relative size measured as share of gross production in the respective industry (see figure 2).[16] Small firms are accordingly represented by short lines and large firms by longer lines. The resulting curve provides an informative portrayal of the distribution of efficiency within the respective industry. The value of about 30 percent for the least efficient firm in figure 1 means that technical efficiency is about 70 percent lower than for the most efficient firm defining the 100 percent benchmark. The total range of the efficiency distribution is calculated by subtracting the percent value of the least efficient firm from 100 percent.

---

[15] This exposition is inspired by diagrams in Salter (1969). Salter displayed productivity levels of firms in ascending order, starting with the least efficient firm.

[16] Other possible measures of size to be used here are the number of employees and the volume of turnover that gauges the importance of the relevant firm on the market. The number of employees is highly correlated with gross production and measures virtually the same thing, i.e. the level of economic activity in the firm. Using the volume of gross production or the amount of turnover as a measure of size may lead to considerably diverging results according to the firms' share of value added. If firms differ with regard to their vertical range of manufacture, turnover does not provide comparable information about the amount of economic activity. A further advantage of gross production as a measure of size is that gross production is not affected by stock-keeping behavior.

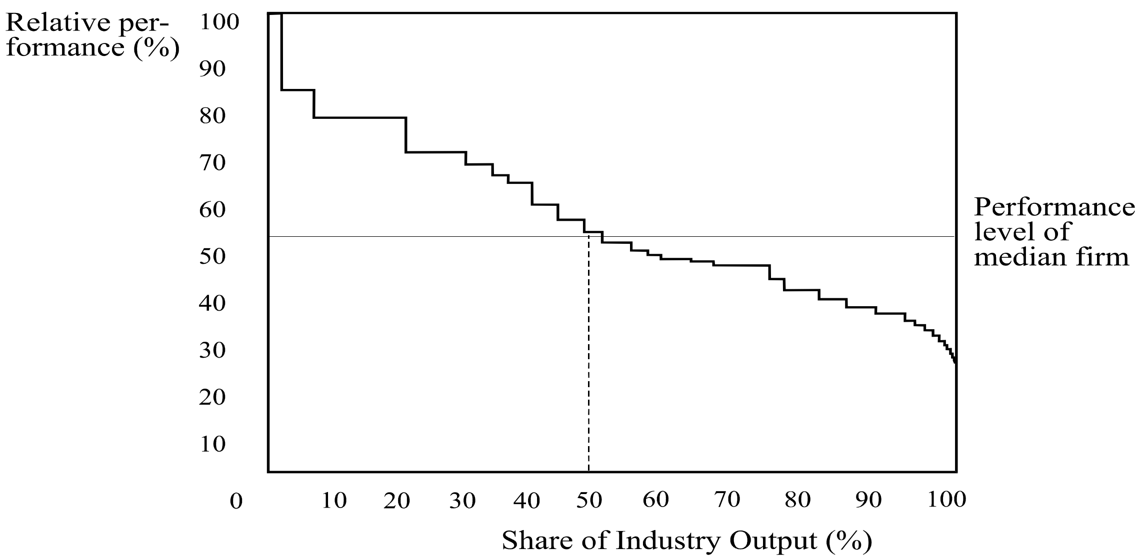

Figure 2: The efficiency distribution curve

Since the range of the efficiency distribution might be affected by single outliers, it is not a robust description of the efficiency heterogeneity of an industry. However, the relative technical efficiency level of the firm at the median output provides a more reliable and robust description of the average efficiency level of an industry. Additionally, efficiency levels related to other output shares could also be taken as measures for relative efficiency of an industry.

Table 5: Summary of relative technical efficiency within industries at different output shares

At share of industry output	*Relative efficiency level (percent)*			
	Mean	*Median*	*Maximum*	*Minimum*
10 %	73.22	73.87	100	11.52
25 %	66.75	66.46	100	10.75
50 %	59.42	57.93	100	10.06
75 %	53.86	52.83	100	9.57
90 %	49.39	48.09	100	8.20
100 %	38.50	37.11	84.11	1.33

On average, the median output firm attains about 59 percent of the maximum efficiency level in the respective industry (table 5). The average minimum efficiency level is 38.5 percent. There is enormous variation of this minimum efficiency level among industries with a highest value of about 84 percent and a lowest value of only 1.3 percent**.** That the highest value of the minimum of technical efficiency within an industry is below 100 percent means that there is at least some technical inefficiency in all industries.

5.2 Description of heterogeneity using an area measure

From the described efficiency distribution curve we derive a measure of efficiency heterogeneity within an industry that accounts for the relative size of the individual firms and that is also rather robust with regard to extreme values. It is defined as the area between the efficiency distribution curve and the efficiency level of the median output share firm in the industry. We label this measure *h-area*, where *h* stands for heterogeneity. This heterogeneity area (*ha*) is defined as follows:

$$ha = \frac{\sum_{1}^{I} \left| e_i - e_m \right| os_i}{0.5}$$

where e_i, $(0 \le e_i \le 1)$, denotes the relative level of technical efficiency of a unit i ($i = 1, \ldots, I$) as a percentage and e_m is the technical efficiency level of the median unit. This median is defined according to the share of industry output as measure of relative size that is used for constructing the curve. The percentage output share of a firm is denoted by os_i $(0 \le os_i \le 1)$. The term in the numerator can assume values between 0 and 0.5. It is zero if all units have the same performance value and, conversely, it is 0.5 if half of the group performs at 100% and the other half has a performance of 0%. Dividing this term by 0.5 gives our measure *ha* with values between 0 and 1. In contrast to other measures of heterogeneity such as the standard deviation or the coefficient of variation, our area measure is sensitive to the size of the firms. For example, it takes into account whether the highly efficient firms have a relatively large share or only a small share of total output in industry. This also implies that the measure is reasonably robust with regard to small firms with extreme values that may not be considered as being representative of the industry. A further advantage of the *h-area* measure is that since both efficiency and firm size are expressed relatively, it can be directly compared between industries. On the average, the value of the area measure amounts to 0.152 (mean) and 0.129 (median) respectively. The maximum value is 0.756 for the striking of coins industry (NACE 36.21). The minimum value of 0.014 (in the manufacture of motor vehicles industry, NACE 36.10) for our measure indicates that there is some heterogeneity of technical efficiency in all industries of our sample (see table A2 in the Appendix).

It is possible to modify the h-area measure so that the most efficient and inefficient five (or ten) percent shares of gross production are excluded. If the lower/upper five percent is omitted we label it "h-area 5-95" measure, if the lower/upper ten percent is cut off we call it "h-area 10-90". Due to the omission of extreme values, these indicators should be even more robust with regard to outliers. On the other hand, one could also suspect that the upper/lower parts of the efficiency distribution curve are important for characterizing the extent of efficiency heterogeneity of an industry.[17]

Correlation coefficients between different indicators of heterogeneity (table 6) clearly prove the advantages of our measure. The comparison includes the different versions of the area measure (h-area 1-100, 5-95 and 10-90), the coefficient of variation and the percentage range between the minimum and maximum value (Range 0-100), as well as the range between the 5 and 95 percentile (Range 5-95) and between the 10 and the 90 percentile (Range 10-90). All three versions of our area measure are closely correlated, indicating great robustness with regard to extreme values. We also find a relatively high degree of correspondence between the area measures and the range indicators with omitted extreme values (Range 5-95, 10-90). Correlation between the area measure and the full range (Range 0-100) is, however,

---

[17] For some illustrative numerical examples of the properties of our measure as compared to the range and the median, see Fritsch and Stephan (2004).

relatively weak. This becomes particularly clear if Spearman rank correlation coefficients are employed as a measure for statistical relationship. A similarly low level of correspondence is found between the full range and the range measures with omitted extreme values. This reveals the impact of extreme cases at the upper and lower end of the spectrum on the range 0-100. The comparison of the different indicators suggests that extreme values also have a relatively strong impact on the coefficient of variation. While there is a considerable correlation between the variation coefficient and the full range, correspondence with the other indicators is considerable weaker, particularly when measured using the Spearman correlation coefficient. This demonstrates the superiority of our measure over the alternative indicators, particularly the range and the coefficient of variation.

Table 6: Correlation coefficients for different measures of heterogeneity within industries[†]

	H-area 0-100	*H-area 5-95*	*H-area 10-90*	*H-area 25-75*	*Range 0-100*	*Range 5-95*	*Range 10-90*	*Range 25-75*
Coefficient of variation	0.517** 0.344**	0.488** 0.280**	0.467** 0.253**	0.425** 0.200**	0.637** 0.686**	0.549** 0.352**	0.449** 0.277**	0.438** 0.229**
	H-area 0-100	0.994** 0.983**	0.982** 0.962**	0.911** 0.878**	0.275** 0.092	0.863** 0.928**	0.935** 0.920**	0.936** 0.906**
		H-area 5-95	0.996** 0.992**	0.937** 0.916**	0.243** 0.037	0.813** 0.872**	0.925** 0.930**	0.952** 0.935**
			H-area 10-90	0.959** 0.945**	0.230** 0.034	0.761** 0.820**	0.896** 0.905**	0.960** 0.945**
				H-area 25-75	0.197** 0.053	0.636** 0.715**	0.760** 0.774**	0.905** 0.892**
					Range 0-100	0.301** 0.091	0.228** 0.008	0.198** 0.012
						Range 5-95	0.845** 0.855**	0.716** 0.765**
							Range 10-90	0.836** 0.836**
								Range 25-75

† First row: Pearson correlation coefficients. Second row: Spearman rank correlation coefficients. N=242 four-digit industries. **: statistically significant at the 1 percent level; *: statistically significant at the 5 percent level.

Correlation coefficients for the relationship between the area measures of overall heterogeneity and the size of the upper/lower part of the efficiency distribution are shown in table 7. Whilst we find significant positive statistical relationship between most of these measures, it is remarkable that correlation coefficients for the size of the upper and the lower part of the heterogeneity area (h-area 0-5, 0-10, 0-25 and 75-100, 90-100, 95-100) are negative. This indicates that heterogeneity tends to be concentrated at one end of the efficiency distribution, the upper *or* the lower part.

Table 7: Correlation coefficients for different area measures of heterogeneity within industries†

	H-area 10-90	*H-area 25-75*	*H-area 0-25*	*H-area 0-10*	*H-area 0-5*	*H-area 75-100*	*H-area 90-100*	*H-area 95-100*
H-area 0-100	.9820** .9617**	.9115** .8777**	.7747** .6601**	.6284** .5543**	.5064** .4480**	.5657** .7017**	.5255** .6786**	.4902** .6490**
H-area 10-90		.9591** .9454**	.7406** .5846**	.5577** .4498**	.4270** .3381**	.5289** .7114**	.4693** .6733**	.4278** .6386**
H-area 25-75			.6666** .5263**	.4837** .4066**	.3690** .3129**	.4278** .6231**	.3753** .5830**	.3379** .5531**
H-area 0-25				.9925** .9424**	.8563** .8635**	-.0390 .0844	-.0495 .0732	-.0609 .0556
H-area 0-10					.9684** .9675**	-.1527* -.0105	-.1719** -.0105	-.1831** -.0476
H-area 0-5						-.2293** -.0967	-.2525** -.1186	-.2657** -.1377*
H-area 75-100							.9638** .9775**	.9279** .9459**
H-area 90-100								.9912** .9885**
H-area 95-100								

† First row: Pearson correlation coefficients. Second row: Spearman rank correlation coefficients. N=242 four-digit industries. **: statistically significant at the 1 percent level; *: statistically significant at the 5 percent level.

5.3 Examples for selected industries

Figure 3 shows the efficiency distribution for all private sector manufacturing industries. The efficiency measures here were derived from a frontier production function estimation for the total manufacturing sector. The displayed efficiency distribution curve indicates that there is a relatively small share of highly efficient firms that represent a very small share of less than 1 percent of total gross production. At the lower end, the share of very low efficient firms is also rather small and constitutes only about 1-2 percent of gross production. This suggests that if the lower and the upper five percent are omitted, most of the extreme cases should be removed from the analysis. The fact that the median efficiency is only about 35 percent of the maximum shows the large dispersion of efficiency levels. The curve breaks off quite abruptly at the lower efficiency end, appearing much more truncated here than at the upper end with the highly efficient firms. The reason for this truncation is presumably the exit of those firms that are not efficient enough to earn their cost. This truncation at the lower end explains why the distribution is positively skewed in most industries.

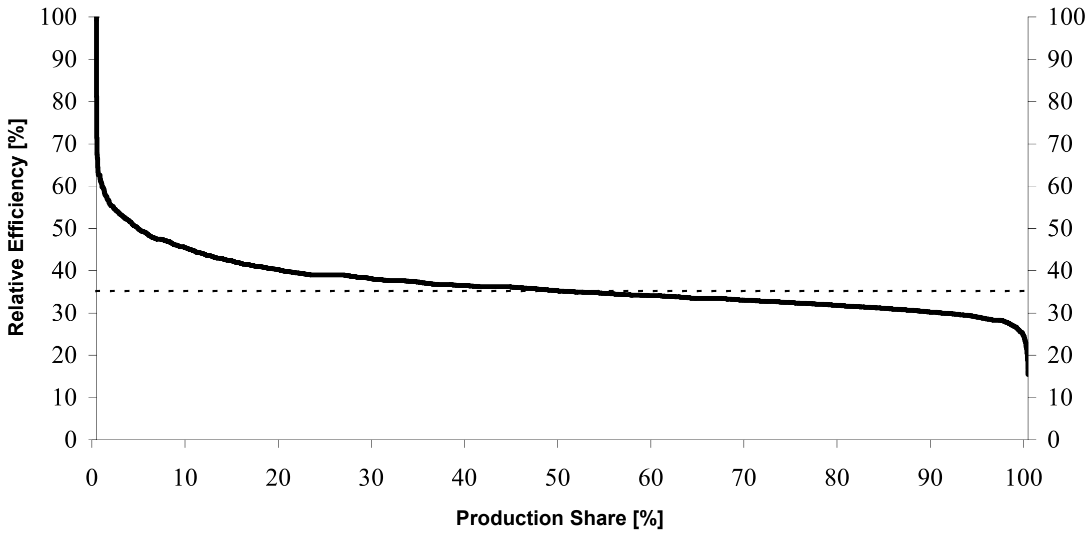

Figure 3: Efficiency distribution curve for all private sector industries

There are considerable differences between industries with regard to the degree of heterogeneity indicated by different measures. Table 8 depicts some measures of the heterogeneity of technical efficiency in these industries. As empirical illustrations, figure 4 to 7 show efficiency distribution curves in four selected types of industries. The first type is industries where one or a few large firms are the most efficient and where the smaller firms are on the inefficient side (figure 4). This pattern could indicate some size advantages in these industries. The second category is industries where large and small firms may be found in all efficiency ranges (figure 5). Size economies do not seem to play a role as far as the technical efficiency in these industries is concerned. A third type appears to be characterized by some size disadvantages because here the small firms are the relatively efficient and the larger firms attain only low levels of technical efficiency (figure 6). Finally, figure 7 displays efficiency distribution curves for industries that have no large firms.

Table 8: Measures for heterogeneity of efficiency in selected industries†

Industry [NACE code]	H-area measure 0-100%	H-area measure 5-95%	Range 0-100%	Range 5-95%	Coef. of variation	Number of firms
15.85: Manufacture of macaroni, noodles, couscous and similar farinaceous products	0.1686	0.1298	43.93	34.70	13.35	34
17.24: Silk-type weaving	0.1502	0.1101	55.31	29.52	17.28	46
20.30: Manufacture of builders' carpentry and joinery	0.1369	0.1009	52.24	28.57	11.98	456
22.12: Publishing of newspapers	0.3976	0.3028	98.67	93.50	82.01	252
24.14: Publishing of sound recordings	0.1393	0.0914	67.59	23.34	22.64	61
24.30: Manufacture of paints, varnishes and similar coatings, printing ink and mastics	0.0760	0.0518	68.51	17.44	13.91	301
24.51: Manufacture of soap and detergents, cleaning and polishing preparations	0.1358	0.1009	57.68	28.49	13.06	118
26.13: Manufacture of hollow glass	0.1326	0.0953	52.85	26.79	14.72	80
26.21: Manufacture of ceramic household and ornamental articles	0.2915	0.2384	72.38	45.73	28.40	92
28.22: Manufacture of central heating radiators and boilers	0.0948	0.0628	69.39	21.01	25.17	49
29.21: Manufacture of furnaces and furnace burners	0.2407	0.1901	60.14	44.21	20.81	87
29.51: Manufacture of machinery for metallurgy	0.2036	0.1567	82.92	39.90	26.08	70
29.56: Manufacture of other special purpose machinery n.e.c.	0.1259	0.0871	68.49	30.47	15.70	816
33.50: Manufacture of watches and clocks	0.2149	0.1642	60.20	41.36	19.93	75
34.10: Manufacture of motor vehicles	0.0144	0.0076	85.53	4.66	48.37	77

† Range of relative efficiency calculated as 100 percent minus lowest value.

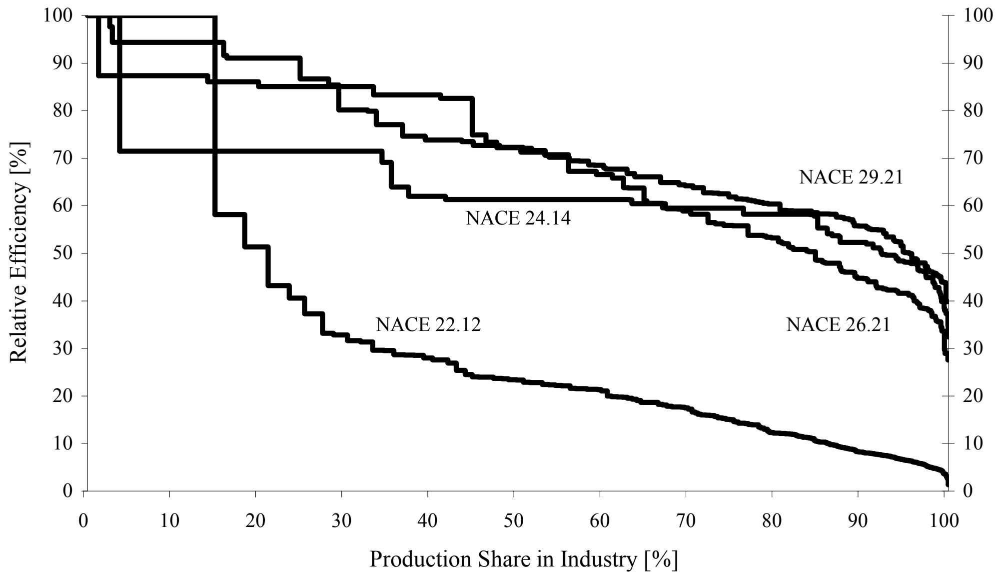

Figure 4: Efficiency distribution curves for industries with large highly efficient firms and smaller inefficient firms

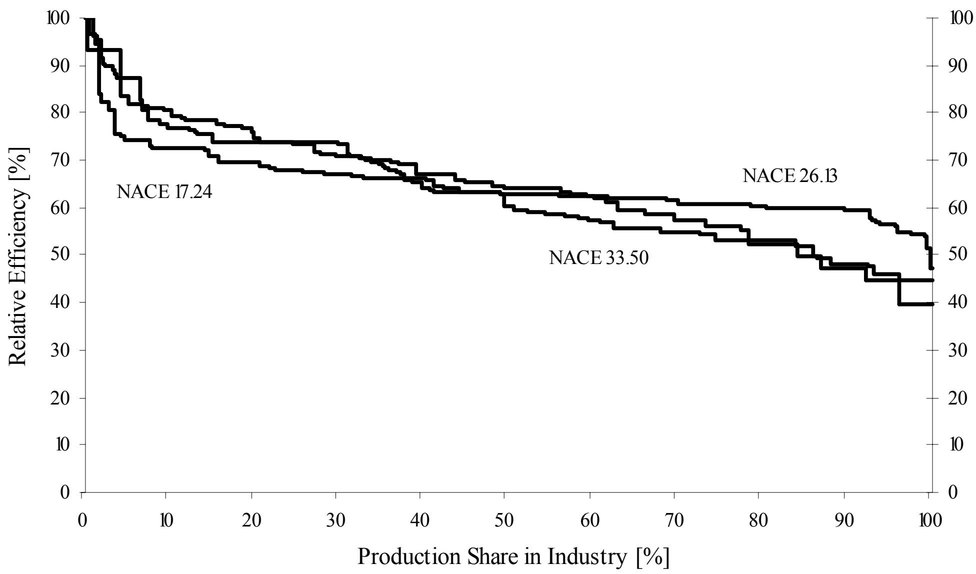

Figure 5: Efficiency distribution curves for industries with large and small firms in all efficiency ranges

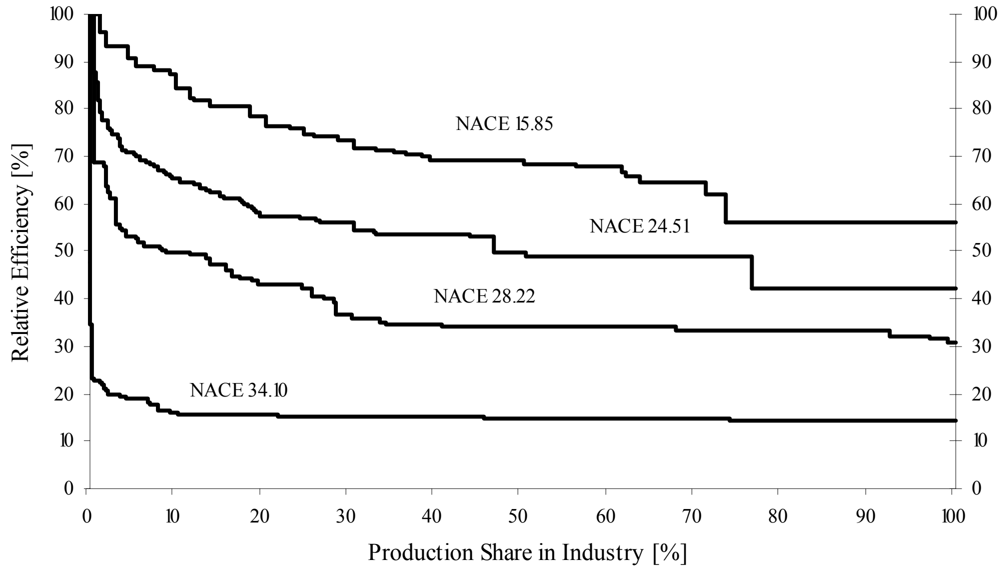

Figure 5: Efficiency distribution curves for industries with efficient small firms and inefficient large firms

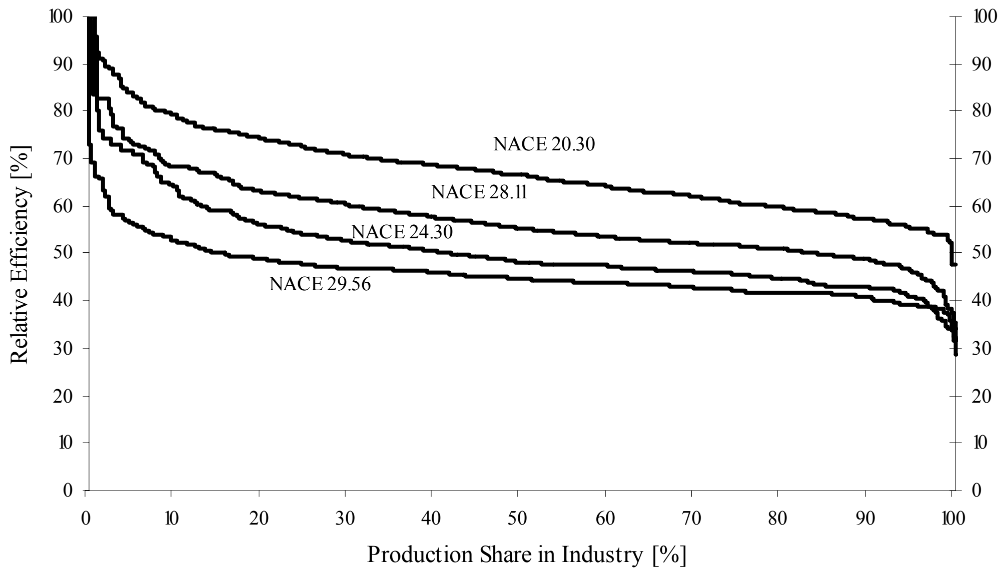

Figure 7: Efficiency distribution curves for industries without large firms

.

In all these kinds of industry, the general picture seems to be that a few firms play a leading role with regard to technical efficiency and the other firms are considerably behind. Consequently, most of firms tend to be clustered at the lower end, not at the upper end of the efficiency distribution.

6. Determinants of intra-industry heterogeneity of efficiency

We first expose our theoretical considerations regarding the determinants of efficiency heterogeneity and the variables we employ for testing our hypotheses (section 6.1). In the second part we present the econometric estimation results (section 6.2).

6.1 Theoretical considerations and hypotheses

It is plausible to assume that a relatively high degree of competition on output markets leads to a low heterogeneity of efficiency within an industry (Scarpetta, 2003). The reason is that in markets with a low level of competition there are fewer opportunities for comparing firms' performances. For this reason, survival of firms is less threatened by inefficient practices as in markets with a high level of competition. Hence, slack and sub-optimal use of factor inputs can persist longer when competition is not so pronounced. Because competitive pressure should provide a powerful motivation for adjusting technology and work organisation to best practice we may expect relatively low heterogeneity of performance in highly competitive markets.

If intensity of competition in an industry increases with the number of firms, then there should be a negative relationship with the degree of heterogeneity. Because a relatively large number of firms in an industry may also constitute a source of diversity, the relationship between the number of firms and performance heterogeneity could as well be positive. Accordingly, a high level of start-ups in an industry may, on the one hand, indicate high pressure of competition that works as a limit to inefficiency, but, on the other hand, new firms could also be a source of diversity. Also the effect of market concentration on intra-industry heterogeneity is unclear. In case that concentration leads to lower intensity of competition the impact on heterogeneity should be positive. If, however, those authors are correct which argue that increasing market concentration leads – up to a certain degree – to intensified competition, there could be a negative relationship with heterogeneity (Salter, 1969, 90-93; Scherer and Ross, 1990). Another reason for expecting a negative relationship between market concentration and heterogeneity is that highly concentrated markets are populated by only some few firms, so that the potential for diversity is limited by the number of firms.

Many approaches which try to explain relative efficiency of firms refer to the vintage of the physical capital stock (Salter, 1969; Aghion and Howitt, 1992; Caballero and Hammour, 1994, 1996; Stein, 1997; Chari and Hopenhayn, 1991). Salter (1969) as one of the first economists who called attention to heterogeneity within industries assumed that the most efficient firm is the best-practice user of the most technologically advanced equipment available. Accordingly, less efficient firms are thought not to apply the best available technology or not having implemented a best practice use of that technology. This view implies that not all firms in an industry adopt a new technology and the method for best practice usage simultaneously. Accordingly, heterogeneity is caused by the time that is necessary for the diffusion of new technology and its best-practice usage. If technology and its usage do play an important role for technical efficiency, then a relatively slow speed of diffusion will lead to a large spread of efficiency levels and a high degree of heterogeneity. Therefore, all the determinants of diffusion speed discussed in the literature may have an influence on the level of heterogeneity (cf. Stoneman, 2002). The main such influences on the speed at which new technology is disseminating within an industry are the intensity of competition, the remaining time for an economically reasonable use of the existing capital stock, capital intensity as an indicator of the amount of investment that is required for a switch to new equipment, other cost of switching to a superior technology (e.g. amount and price of complementary resources like human capital), availability of financial funds (e.g. profits) as well as the magnitude of technical progress incorporated in the new machinery. Moreover, the time required for the diffusion of knowledge about advantages of the new technology and its efficient usage may play a role in this respect.

We expect a negative relationship between an industries' average firm size and the level of performance heterogeneity for at least two reasons. First, markets with high average firm size provide room for only few firms and, therefore, low potential for diversity. And second, larger firms are not as likely to suffer from internal bottlenecks with regard to the adoption of new technology as smaller firms. Hence, adoption of new technology should be faster in industries with relatively high average firm size than in industries where average firm size is smaller.

Another group of factors for explaining divergent efficiency levels is the diversity of supplies in an industry. The less homogeneous and integrated an industry, the higher the level of heterogeneity that can be expected. One measure of the diversity of supply is the dispersion of firm size indicating different production processes. The share of activity in research and development (R&D) and human capital intensity indicate non-standardized products and may, therefore, also be taken as measures for the heterogeneity of supply. Another source of heterogeneity of efficiency levels within an industry could be diversity of the firms' locational conditions.[18]

Many of our hypotheses regarding the determinants of efficiency heterogeneity are related to the development stage of an industry as characterized by its technological regime (cf. Audretsch, 1995, 39-64; Winter, 1984). We can, therefore, expect that the development stage of an industry constitutes an important determinant of the heterogeneity of efficiency levels. In a new industry with an entrepreneurial regime[19], products and processes are

---

[18] We use the share of firms with headquarters in Western Germany as an indicator for homogeneity of locational conditions due to the large differences that can be found between the two parts of the country. Except for two industries, this share is well above 50 percent with an average value of 86.88 percent (mean) and 89.47 percent (median) respectively (see table A2 in the Appendix). The higher the proportion of Western firms the lower the heterogeneity of locational conditions.

[19] We use the share of R&D employees in small firms with less than 50 employees over the share of R&D employees in firms of all size classes. This indicator corresponds to the "small-

rather diverse inducing a relatively high level of heterogeneity and high importance of competition by quality as compared to price competition. Relevant knowledge is relatively new and dispersed. In such an industry it is unlikely that scale economies play a role. Therefore, the larger firms have no significant advantage over the smaller firms. This is different in mature industries with a routinized technological regime. In this regime-type, the technological path is more progressed so that the stock of path-specific knowledge is considerably older giving the incumbent large firms an advantage over their smaller competitors. Under a routinized regime, the efficiency distribution curve should run relatively flat due to high intensity of competition between suppliers of homogeneous products that are manufactured by highly standardized processes.

The effect of the industries' growth rate on heterogeneity of technical inefficiency is unclear. On the one hand, growth may induce high investment and speedy adoption of new technology. On the other hand, economic prosperity could be associated with only low pressure to modernize machinery and, thus, may allow for relatively low efficiency and a correspondingly high degree of heterogeneity. If some part of the observed heterogeneity results from incomplete adjustment to changing economic conditions, then a turbulent environment, e.g. rapid technological developments or fluctuation of demand, should lead to a relatively pronounced level of diversity.

We have already argued (in section 2) that the minimum efficiency level could be determined by costs, in particular by labor costs. The company must earn the money to pay for these costs in order to survive in the market. If wages do not diverge much among the industries then we should also expect roughly the same minimum efficiency levels required for economic survival in all industries. In this case, we can expect a relationship between an industries' average level of relative efficiency and the intra-industry heterogeneity of

firm innovation rate / total innovation rate" used by Audretsch (1995) as a measure of the entrepreneurial character of an industry. In contrast to Audretsch's indicator, which is based on the number of innovations introduced, our measure refers to R&D input.

efficiency. Such a relationship may occur because the spread between the bottom-line of efficiency and the efficiency leader will be the larger the higher the average efficiency of the industry. Accordingly, there should be some firms with rather low levels of relative efficiency even in those industries that on average perform relatively well.

Table 9: Overview of hypotheses about the effect of different factors on heterogeneity of efficiency within industries

Determinant	*Expected sign for relationship with heterogeneity*
Number of firms in industry	– / +
New firm formation rate	– / +
Market concentration	– / +
Capital intensity	+
Average firm size	–
Diversity of firm size	+
Human capital intensity	+
R&D intensity	+
Homogeneity of locational conditions	+
Entrepreneurial character of industry	+
Output growth rate	– / +
Average value of relative efficiency	+

Table 9 summarizes the hypotheses outlined above giving the expected sign for the relationship with heterogeneity of intra-industry technical efficiency. Table A1 in the Appendix gives the definition of the independent variables as used in the empirical analysis; table A2 provides descriptive statistics of dependent and independent variables. Because the number of firms in an industry as well as the new firm formation rate stand for intensity of competition *and* diversity of supply, the impact on heterogeneity is a priori unclear. The same holds for market concentration because its effect on the intensity of competition is also undecided. Capital intensity is a measure for the amount of investment that is required to switch to new equipment or for the sunk cost in form of the old equipment that can no longer be used. The higher the capital intensity in an industry, the higher the expected level of heterogeneity. Average firm size should have a negative effect on heterogeneity of efficiency. There are a number of industry characteristics that may reflect the variety of products and processes such as diversity of firm size, human capital intensity and R&D intensity. We expect that high variety of products and processes results in more pronounced heterogeneity of efficiency levels. The same holds for the differences of locational conditions within an industry. High growth of an industry's output may have heterogeneity increasing as well as decreasing effects. A relatively high average efficiency level allows for pronounced diversity of efficiency levels.

6.2 Econometric results

The econometric analysis was performed with different versions of our area measure as the dependent variable, as well as with different estimation methods. For describing the overall heterogeneity of efficiency performance, we use the area measure for the full spectrum (h-area 0-100). We also analyzed heterogeneity for different parts of this curve. Cutting off the upper and lower 10 percent or 25 percent gives a measure for the middle part (h-area 10-90, 25-75, see table 10). The results for this middle part are, however, quite similar to those for the complete heterogeneity area (0-100). We also analyzed heterogeneity at the upper/lower tail of the efficiency curve, taking the area for the upper/lower 5 percent, 10 percent and 25 percent as dependent variables (table 11 and 12).

Since ordinary least squares (OLS) estimation is rather sensitive with regard to extreme observations, we also applied Reweighted Least Squares which is based on outlier robust Least Trimmed Squares (LTS) regression in a first step (see Rousseeuw and Leroy, 1987, for details). In estimating Reweighted Least Squares the identified extreme observations from LTS receive a weight of zero. As a third method, we applied OLS estimation based on the rank values of the variables (Conover and Iman, 1982; Iman and Conover, 1979). Compared to the other regression methods, this approach has three advantages. First, like LTS regression (or Spearman correlation coefficient), it is quite robust with regard to outliers. Second, because values are based on ranks, non-linear monotonous relationships can be identified that may not be found using the linear regression methods. Third, heteroscedasticity and multicollinearity of the original variables will be reduced with rank regression. However, as far as the 'true' relationships are linear, rank regression has the disadvantage of being a relatively inefficient estimation method.

Table 10 shows the results of regressions for heterogeneity measures h-area 0-100, 10-90 and 25-75. Regressions for explaining heterogeneity at the upper end (h-area 0-5, 0-10 and 0-25) and at the lower end (h-area 75-100, 90-100 and 95-100) of the distribution are reported in tables 11 and 12. The indicators for the number of firms, market concentration and the entrepreneurial character of an industries' technological regime were not included into the final versions of the models because they turned out to be hardly statistically significant. The indicator for market concentration showed some positive correlation with average firm size. However, when including the market concentration indicator instead average firm size we found that average firm size is of considerably higher significance and explains the heterogeneity within industries much better than market concentration. We also tested the impact of an industries' exposure to foreign international competition, i.e. imports and exports, but did not achieve any statistically significant results. A summary of the findings is provided in table 13 where the direction of the impact is indicated by positive and negative signs.

Table 10: Regression analyses for different heterogeneity areas (h-areas 0-100, 10-90 and 25-75) †

	h-area 0-100 %			h-area 10-90 %			h-area 25-75 %		
Variable	*OLS*	*RLS*	*Rank*	*OLS*	*RLS*	*Rank*	*OLS*	*RLS*	*Rank*
Intercept	-2.636** (-5.01)	-3.125** (-7.60)	91.161** (3.94)	-3.143** (-4.08)	-5.979** (-12.13)	89.679** (3.86)	-4.777** (-5.56)	-8.471** (-12.86)	86.708** (3.53)
Average efficiency value	0.393** (4.51)	0.446** (5.96)	0.262** (4.42)	0.484** (3.85)	0.936** (11.11)	0.297** (4.95)	0.531** (3.80)	1.101** (10.78)	0.246** (3.91)
Average firm size	-0.130** (-2.72)	-0.076* (-2.12)	-0.101 (-1.10)	-0.164* (-2.36)	-0.154** (-3.67)	-0.061 (-0.65)	-0.177* (-2.27)	-0.055 (-0.88)	-0.069 (-0.69)
Diversity of firm size	-0.012 (-0.36)	-0.062* (-2.51)	-0.077 (-1.00)	-0.015 (-0.31)	-0.058* (-1.98)	-0.081 (-1.04)	-0.148** (-2.76)	-0.072 (-1.27)	-0.063 (-0.77)
Capital intensity	5.092** (2.75)	1.454 (1.00)	0.129* (2.08)	4.982 (1.87)	4.165** (2.62)	0.082 (1.31)	4.760 (1.60)	3.551 (1.71)	0.090 (1.38)
Human capital intensity	2.321** (3.68)	1.465** (3.15)	0.256** (3.11)	2.518** (2.77)	2.663** (4.90)	0.256** (3.07)	3.915** (3.86)	2.959** (4.35)	0.236** (2.70)
R&D intensity	-5.705 (-1.16)	0.212 (0.06)	-0.089 (-1.16)	-4.395 (-0.62)	-7.523 (-1.69)	-0.120 (-1.54)	-12.730 (-1.61)	-10.272 (-1.89)	-0.114 (-1.40)
New firm formation rate	0.188 (1.69)	0.256** (3.10)	0.198** (2.73)	0.095 (0.59)	0.237* (2.42)	0.163* (2.23)	0.088 (0.49)	0.329** (2.49)	0.125 (1.59)
Output growth rate	-3.028** (-3.57)	-2.222** (-3.70)	-0.206** (-3.21)	-4.839** (-3.97)	-2.054** (-2.75)	-0.181** (-2.77)	-2.437 (-1.74)	-3.517** (-3.62)	-0.151* (-2.21)
Homogeneity of locational conditions	-0.007* (-2.43)	-0.005* (-2.30)	-0.123* (-2.06)	-0.011* (-2.49)	0.001 (0.07)	-0.098 (-1.62)	-0.004 (-0.78)	0.004 (1.19)	-0.025 (-0.40)
R-squared	0.246	0.331	0.220	0.194	0.528	0.204	0.203	0.468	0.142
Root mean squared error	0.470	0.318	63.013	0.676	0.376	63.398	0.745	0.455	65.015
No. of observations	242	219	242	241	216	241	238	208	238

† T-values in parentheses. **: statistically significant at a 1 percent level; *: statistically significant at a 5 percent level.

The estimations for the overall level of heterogeneity (table 10) show a strong and robust positive relationship with the median efficiency level of the industry. This phenomenon can be explained by more or less equal minimum wage levels across industries. The wage level is an important determinant of the minimum efficiency that a firm has to attain in order to survive. If a firm is not efficient enough to cover its costs, then it will sooner or later have to leave the market. This effect cuts off the firms with insufficient performance at the lower end of the efficiency distribution (see section 2). Given the

Table 11: Regression analyses for different heterogeneity areas at the upper end of the efficiency distribution (h-areas 0-5, 0-10 and 0-25) †

	h-area 0-5 %			h-area 0-10 %			h-area 0-25 %		
Variable	*OLS*	*RLS*	*Rank*	*OLS*	*RLS*	*Rank*	*OLS*	*RLS*	*Rank*
Intercept	-0.637 (-1.04)	-1.352** (-3.10)	137.649** (6.31)	-0.311 (-0.46)	-0.925 (-1.88)	127.595** (5.67)	0.239 (0.31)	-1.078* (-2.13)	119.361** (5.17)
Average efficiency value	-0.579** (-5.72)	-0.501** (-6.65)	-0.317** (-5.53)	-0.475** (-4.29)	-0.361** (-4.58)	-0.223** (-3.77)	-0.392** (-3.09)	-0.111 (-1.35)	-0.132* (-2.18)
Average firm size	-0.183** (-3.32)	-0.132** (-3.27)	-0.130 (-1.45)	-0.201** (-3.33)	-0.191** (-4.52)	-0.108 (-1.17)	-0.215** (-3.12)	-0.225** (-5.07)	-0.062 (-0.65)
Diversity of firm size	-0.023 (-0.61)	0.033 (0.97)	-0.005 (-0.07)	-0.012 (-0.28)	0.061 (1.79)	-0.005 (-0.06)	-0.006 (-0.13)	0.115** (3.27)	-0.025 (-0.32)
Capital intensity	5.892** (2.79)	6.994** (5.10)	0.235** (3.91)	6.069** (2.62)	5.584** (3.94)	0.207** (3.35)	6.410* (2.42)	6.043** (3.90)	0.188** (2.96)
Human Capital Intensity	0.134 (0.19)	0.132 (0.27)	0.038 (0.48)	0.324 (0.41)	1.586** (3.32)	0.064 (0.78)	0.477 (0.53)	1.765** (3.46)	0.099 (1.17)
R&D intensity	-7.402 (-1.31)	-4.436 (-1.22)	-0.114 (-1.52)	-7.746 (-1.25)	-20.096** (-5.09)	-0.132 (-1.71)	-8.087 (-1.14)	-16.279** (-3.97)	-0.160* (-2.03)
New firm formation rate	0.244 (1.92)	0.213** (2.80)	0.183** (2.61)	0.273* (1.97)	0.229** (2.62)	0.231** (3.19)	0.292 (1.84)	0.339** (3.71)	0.262** (3.53)
Output growth rate	-0.552 (-0.57)	-0.725 (-1.21)	-0.078 (-1.26)	-0.914 (-0.86)	-1.142 (-1.78)	-0.101 (-1.57)	-1.100 (-0.91)	-1.471* (-2.21)	-0.102 (-1.54)
Homogeneity of locational conditions	-0.002 (-0.51)	0.001 (-0.24)	0.031 (0.53)	-0.004 (-1.09)	-0.003 (-1.21)	-0.007 (-0.12)	-0.008 (-1.83)	-0.007** (-2.83)	-0.071 (-1.15)
R-squared	0.279	0.405	0.281	0.235	0.430	0.234	0.193	0.410	0.196
Root mean squared error	0.532	0.306	59.293	0.582	0.333	61.195	0.666	0.344	62.676
No. of observations	237	211	237	237	209	237	237	209	237

† T-values in parentheses. **: statistically significant at a 1 percent level; *: statistically significant at a 5 percent level.

approximately equal wage levels, dispersion of technical efficiency can be greater in industries with high average efficiency values.[20] Estimations for the lower and upper end of the efficiency distribution (table 11 and 12) clearly show that industries with a high level of median efficiency are less heterogeneous in the upper part of the distribution, but much more

[20] This implies that there is considerable positive correlation between the measures of relative efficiency within a certain industry and within the manufacturing sector as a whole as we have reported above (section 4).

Table 12: Regression analyses for different heterogeneity areas at the lower end of the efficiency distribution (h-areas 75-100, 90-100 and 95-100) †

	h-area 75-100 %			h-area 90-100 %			h-area 95-100 %		
Variable	*OLS*	*RLS*	*Rank*	*OLS*	*RLS*	*Rank*	*OLS*	*RLS*	*Rank*
Intercept	-6.802** (-2.57)	-8.017** (-17.87)	58.920** (3.02)	-7.761** (-3.00)	-8.067** (-21.83)	53.144** (2.78)	-8.511** (-3.35)	-8.784** (-24.33)	51.148** (2.69)
Average efficiency value	1.516** (3.41)	1.350** (16.84)	0.543** (10.6)	1.547** (3.56)	1.168** (18.08)	0.572** (11.38)	1.567** (3.68)	1.183** (18.85)	0.587** (11.75)
Average firm size	0.081 (0.35)	0.003 (0.07)	0.092 (1.16)	0.069 (0.31)	-0.053 (-1.68)	0.047 (0.60)	0.054 (0.24)	-0.070* (-2.22)	-0.003 (-0.03)
Diversity of firm size	-0.043 (-0.26)	-0.058* (-2.20)	-0.187** (-2.79)	-0.021 (-0.13)	-0.054** (-2.65)	-0.170** (-2.59)	-0.009 (-0.06)	-0.034 (-1.73)	-0.148* (-2.27)
Capital intensity	-20.151* (-2.26)	2.329 (1.51)	-0.030 (-0.56)	-19.414* (-2.22)	3.015* (2.53)	-0.029 (-0.55)	-18.978* (-2.22)	2.851* (2.46)	-0.030 (-0.58)
Human capital Intensity	3.142 (1.03)	2.416** (4.07)	0.290** (4.09)	3.070 (1.04)	3.674** (8.20)	0.300** (4.31)	2.948 (1.01)	3.655** (8.18)	0.307** (4.43)
R&D intensity	1.460 (0.06)	4.980 (1.19)	-0.024 (-0.37)	0.231 (0.01)	4.022 (1.31)	-0.013 (-0.19)	-0.354 (-0.02)	2.372 (0.79)	-0.004 (-0.07)
New firm formation rate	0.563 (1.03)	0.322** (3.50)	0.162* (2.60)	0.551 (1.03)	0.304** (4.31)	0.153* (2.50)	0.541 (1.03)	0.258** (3.72)	0.133* (2.19)
Output growth rate	-5.107 (-1.24)	-2.427** (-3.50)	-0.263** (-4.73)	-4.469 (-1.11)	-2.259** (-4.17)	-0.246** (-4.50)	-3.828 (-0.97)	-2.003** (-3.81)	-0.224** (-4.11)
Homogeneity of locational conditions	-0.028* (-1.97)	-0.010** (-3.99)	-0.076 (-1.48)	-0.026 (-1.83)	-0.006** (-3.04)	-0.058 (-1.16)	-0.024 (-1.73)	-0.004* (-2.00)	-0.047 (-0.93)
R-squared	0.107	0.664	0.432	0.108	0.750	0.452	0.108	0.754	0.457
Root mean squared error	2.262	0.353	52.918	2.209	0.266	51.962	2.168	0.260	51.703
No. of observations	238	212	238	238	198	238	238	198	238

† T-values in parentheses. **: statistically significant at a 1 percent level; *: statistically significant at a 5 percent level.

heterogeneous in the lower part. Evidently, relatively high efficiency of the median output unit leads to a longer 'tail' of the efficiency distribution.

Average firm size has a marked negative impact on the degree of heterogeneity within an industry, in particular at the upper part of the efficiency curve. This negative relationship between average firm size and heterogeneity of technical efficiency has two main explanations. First, if firm size is a result of relatively large-scale production technology, the respective

Table 13: Summary of findings[†]

Independent variable	Overall	*Part of h-area* Upper	Middle	Lower
Average value of relative efficiency	+	-	+	+
Average firm size	-	-	(-)	(-)
Diversity of firm size	(-)	(+)	(-)	n.s.
Capital intensity	+	+	(+)	+/-
Human capital intensity	+	(+)	+	+
R&D intensity	n.s.	(-)	n.s.	n.s.
New firm formation rate	+	+	+	+
Output growth rate	-	(+)	-	-
Homogeneity of locational conditions	-	(-)	(-)	n.s.
Number of firms in industry	n.s.	n.s.	n.s.	n.s.
Market concentration	n.s.	n.s.	n.s.	n.s.
Entrepreneurial character of industry	n.s.	n.s.	n.s.	n.s.

[†] Signs in parentheses: variable was statistically significant at the 5 percent level in less than half of the models reported; n.s.: variable was not statistically significant at the 5 percent level in any of the models reported.

industry may be in the later stages of its life cycle when products are rather standardized and firms tend to apply about the same 'dominant' type of technology. Second, high average efficient size implies a relatively small number of firms and, thereby, a low potential for heterogeneity. High capital intensity of an industry is related to a high level of heterogeneity. This is consistent with our hypothesis that high capital intensity slows down the diffusion of new technology (section 6.1). If capital intensity is high, production conditions in the industry may be quite diverse resulting in different efficiency levels. The impact of human capital intensity on heterogeneity is, as predicted, positive. This is in accordance with the assumption that industries with a high level of human capital intensity have rather diverse products (section 6.1). However, for R&D intensity, which can also be regarded an indicator for output diversity, we found nearly no statistically significant impact.[21]

A high new firm formation rate goes hand in hand with a high level of heterogeneity in the respective industry that is well pronounced at the lower and upper ends of the efficiency distribution. If entries were only marginal firms with low technical efficiency operating at the fringe of the market, then we would expect a positive impact only on the lower part of the distribution as is confirmed by our estimates. That we also find a positive effect of the startup rate on the upper part suggests that a considerable proportion of the new firms are characterized by a relatively high efficiency level and that these new firms have succeeded in establishing a market position at the upper end of the scale.

The impact of the average sales growth rate of an industry on heterogeneity is found to be negative, in particular at the lower end of the distribution curve. This may be explained by arguing that economic prosperity is conducive for the adoption of new technology so that production conditions within an industry are rather similar and heterogeneity of technical efficiency is low (see section 6.1). The more homogeneous the locational conditions, as measured by the share of West German firms, the lower the level of heterogeneity with regard to technical efficiency. This effect is, however, not very pronounced.

Summarizing the results, one can say that a high average level of relative technical efficiency, of capital intensity, of human capital intensity and of new firm formation lead to a relatively high heterogeneity of technical efficiency in an industry. Large average firm size, a high rate of output growth and homogeneity of locational conditions result in a relatively low degree of heterogeneity.

[21] This relatively low impact of R&D intensity can hardly be explained by the considerable positive statistical correlation with the indicator for human capital intensity that we find in the data. Omitting the indicator for human capital intensity in the model does not lead to any considerably higher coefficient of R&D intensity and vice versa.

7. Summary and Conclusions

In this paper we have analyzed the heterogeneity of technical efficiency within industries. We could show that differences in the level of technical efficiency between firms can be found in all the industries of our sample. Industries differ a lot with regard to the dispersion of technical efficiency indicating a respective variety of competitive conditions. The results suggest that the lower boundary for technical efficiency is given by the costs which must be covered in order to be able to survive in the market. This necessary minimum of technical efficiency seems to truncate the distribution at its lower end. In industries that show a low level of heterogeneity with regard to technical efficiency, firms are clustered near this minimum. It appears that in such markets firms are rather similar with regard to technology and innovation, so that other factors are more important for gaining market shares. Such industries are characterized by large average firm size, homogeneity of locational conditions, low capital intensity and relatively high growth rates of output (cf. table A3 in the Appendix). There is no complete match of these characteristics with commonly used categorization schemes like the industry life-cycle concept (Klepper, 1997) or the concept of technological regimes (Audretsch, 1995; Winter, 1984). However, one may say that many of these industries are positioned in the latter stages of their life-cycle and that the characteristics of innovation activity show some correspondence with a routinized technological regime.

A main cause for pronounced heterogeneity of technical efficiency in industries is that some firms exceed the minimum efficiency level that is required to survive in the market. Industries with a relatively high level of heterogeneity are characterized by some highly efficient firms whose performance level is much higher than the necessary minimum. In these industries, innovation and technology seem to play an important role for economic success. We found that the main drivers of the variation of technical efficiency within an industry are the diversity of firms and conditions as well as the ease of adopting technical change. Industries with a relative pronounced heterogeneity of technical efficiency tend to be characterized by low average firm size, high capital intensity, high human capital intensity as well as moderate output growth. Additionally, high entry rates of new firms correspond to high levels of heterogeneity (see table A3 in the Appendix). Our results indicate that not all entries are marginal firms but that some of the entries are characterized by a relatively high degree of technical efficiency. The industries with high values of the heterogeneity indicator appear rather diverse. It can hardly be said that they show significant correspondence to the concept of an early stage of the product life-cycle and an entrepreneurial technological regime. Maybe, these industries are per se rather fragmented and a large part of the observed heterogeneity is caused by differences between the sub-markets in an industry.

An important finding of our analysis was that in the overwhelming majority of industries, the distribution of technical efficiency is positively skewed, i.e. with a longer tail at the end of the relatively high efficient firms. This contradicts the widespread assumption in the literature that the distribution is negatively skewed with a relatively wide range of values among the low efficiency level firms. The pronounced positive skewness that we found implies that it would not be appropriate to use a stochastic frontier production function for assessing technical efficiency because this type of function is based on the assumption of negative skewness. The positive skewness of the distribution of technical efficiency can be explained by a truncation of this distribution at the lower end. This truncation occurs because firms whose efficiency falls below a certain level that is given by production cost make losses and will sooner or later have to exit the market.

We also introduced a new measure for intra-industry heterogeneity of technical efficiency that gives a more reliable description of the intra-industry distribution than conventional statistical measures such as the range or the coefficient of variation. In particular, this new measure is quite robust with regard to extreme values. Our measure is derived from efficiency distribution curves that provide an interesting portrayal of the competitive situation within an industry. The population of firms in an industry is heterogeneous and this heterogeneity reflects important characteristics of competition, innovation processes and development that deserve further attention.

References

Aghion, Philippe and Peter Howitt (1992): A Model of Growth through Creative Destruction, *Econometrica*, 60, 323-351.

Audretsch, David. B. (1995): *Innovation and Industry Evolution*, Cambridge (Mass.): MIT Press.

Bailey, Sheryl D. (1992): The Intraindustry Dispersion of Plant Productivity in the British Manufacturing Sector, 1963-79, in: Richard E. Caves (ed.), pp. 329-384.

Baltagi, Badi H. (2001): *Econometric Analysis of Panel Data*, 2nd ed.: John Wiley & Sons.

Caballero, Ricardo and Mohamad Hammour (1994): The Cleansing Effects of Recessions, *American Economic Review*, 84, 1352-1368.

Caballero, Ricardo and Mohamad Hammour (1996): On the Timing and Efficiency of Creative Destruction, Quarterly Journal of Economics, 111, 805-852.

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

Caves, Richard E. (ed.) (1992): *Industrial efficiency in six nations*, Cambridge (Mass.): MIT Press.

Caves, Richard E. (1992): Introduction and Summary, in: Richard E. Caves (ed.), pp. 1-27.

Chari, V.V. and Hugo Hopenhayn (1991): Vintage Human Capital, Growth, and the Diffusion of New Technology, *Journal of Political Economy*, 99, 1142-1165.

Coelli, Tim, D.S. Prasada Rao and George E. Battese (1998): *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers.

Conover, W.J. and R.L. Iman (1982): Analysis of Covariance Using Rank Transformation, *Biometrics*, 38, pp. 715-724.

Cubbin John S. and Paul Geroski (1997), The convergence of profits in the long-run: Inter-firm and inter-industry comparisons, *Journal of Industrial Economics*, 35, pp. 427-442.

Farrell, M.J. (1975): The Measurement of Productive Efficiency, *Journal of the Royal Statistic Society*, 120, pp. 253-282.

Fritsch, Michael and Joern Mallok (2002): Machinery and Productivity - A Comparison of East and West German Manufacturing Plants, in:

Ludwig Schätzl and Javier Revilla Diez (eds.), *Technological Change and Regional Development in Europe*, Heidelberg/New York: Physica, pp. 61-73.

Fritsch, Michael and Andreas Stephan (2004): Measuring Performance Heterogeneity within Groups – A Two-Dimensional Approach, Working Paper 03/2004, Faculty of Economics and Business Administration, Technical University Bergakademie Freiberg.

Görzig, Bernd and Claudius Schmidt-Faber (2001): *Wie entwickeln sich die Gewinne in Deutschland? Gewinnaussagen von Bundesbank und Volkswirtschaftlicher Gesamtrechung im Vergleich* (How do profits develop in Germany? A comparison of results about profits from the balance-sheet statistics of the Bundesbank and from the National Accounting System), Berlin: Duncker & Humblot (Deutsches Institut für Wirtschaftsforschung, Sondeft 17).

Greene, William (1997): Frontier Production Functions, in: M. Hashem Pesaran and Peter Schmidt (eds.)*: Handbook of Applied Econometrics*, Vol. II, Blackwell Publishers.

Greene, William (2000): Econometric Analysis, 4th ed.: Prentice-Hall.

Heshmati, Almas; Subal C. Kumbhakar and Lennart Hjalmarsson (1995): Efficiency of the Swedish Pork Industry: a Firm-level Study Using Rotating Panel Data 1976-1988, *European Journal of Operational Research*, 80, pp. 519-533.

Iman, R.L. and W.J. Conover (1979): The Use of Rank Transformation in Regression, *Technometrics*, 4, pp. 499-509.

Klepper, Steven (1997): Industry life cycles, Industrial and Corporate change, 6, pp. 145-181.

Mayes, David, G., Christopher Harris and Melanie Lansbury (1994): *Inefficiency in Industry*, Hemel Hampstead, Harvester Wheatsheaf.

Mayes, David G. (ed.) (1996): *Sources of productivity growth*, Cambridge: Cambridge University Press.

Rousseeuw, P.J. and A.M. Leroy (1987): *Robust Regression and Outlier Detection*, New York: John Wiley & Sons, Inc.

Salter, Wilfred. E. (1969): *Productivity and Technical Change*, 2nd. ed., Cambridge 1969.

Scarpetta, Stefano (2003): *The Sources of Economic Growth in OECD Countries*, Paris: OECD.

Scherer, Frederic M. and David Ross (1990): *Industrial Market Structure and Economic Performance*, 3rd edition, Boston: Houghton Mifflin.

Schmidt, Peter and Robin C. Sickles (1984): Production Frontiers and Panel Data, *Journal of Business & Economic Statistics*, 2, pp. 367-374

Statistisches Bundesamt (German Federal Statistical Office) (annual series): *Kostenstrukturerhebung im Verarbeitenden Gewerbe*, Fachserie 4, Reihe 4.3, Stuttgart: Metzler-Poeschel.

Stoneman, Paul (2002): *The economics of technological diffusion*, Oxford: Blackwell.

Winter, Sidney G. (1984): Schumpeterian Competition in Alternative Technological Regimes, *Journal of Economic Behavior and Organization*, 5, pp. 287-320.

Appendix

Table A1: Definition of independent variables

Number of firms in industry	Log of the number of firms in the industry (for the smaller firms estimated on the basis of sampling rates)
New firm formation rate	Mean annual number of new firms[a] per employee[b] at the 4-digit industry level 1992-2001
Market concentration	Mean value of the Herfindahl index in the 1992 to 2001.
Capital intensity	Mean of annual depreciations plus expenditures for rents and leases over sales at firm-level from 1992 to 2000
Average firm size	Log of mean number of employees in respective industry from 1992 to 2001
Diversity of firm size	Coefficient of variation of production shares in industry
Human capital intensity	Number of employees with a university degree divided by number of untrained employees
R&D intensity	Mean of R&D over gross production in the 1999 to 2001 period
Homogeneity of locational conditions	Proportion of firms with headquarter in West Germany
Entrepreneurial character of industry	Share of R&D expenditure on gross production in firms with less than 50 employees over share of R&D expenditure in firms of all size categories. Mean value of the 1999-2001 period
Output growth rate	Average of annual firms' growth rate of sales in the industry
Average value of relative efficiency	Log of relative efficiency level of median output unit

Source: German Cost Census Statistics if not indicated otherwise.
a Source: Firm foundations panels of the Center for European Economic Research (ZEW, Mannheim).
b Social Insurance Statistics of the German Employment Office.

Table A2: Descriptive statistics of dependent and independent variables

Variable	Mean	Median	Standard deviation	Minimum	Maximum
H-area 0-100	0.152	0.129	0.090	0.014	0.756
H-area 0-100 (log)	-2.024	-2.052	0.531	-4.244	-0.279
H-area 10-90 (log)	-2.684	-2.691	0.738	-7.409	-0.551
H-area 0-10 (log)	-3.350	-3.300	0.653	-7.973	-1.820
H-area 0-5 (log)	-3.838	-3.760	0.614	-7.973	-2.513
H-area 90-100 (log)	-4.000	-3.867	2.294	-37.43	-2.106
H-area 95-100 (log)	-4.543	-4.377	2.252	-37.43	-2.540
Number of firms in industry	144.5	71.5	180.1	5	897
New firm formation rate	0.242	0.124	0.304	0	2.216
Market concentration	0.109	0.057	0.134	0.002	0.846
Capital intensity	0.044	0.017	0.040	0.012	0.108
Average firm size	5.235	5.134	0.805	3.684	9.269
Diversity of firm size	1.726	1.065	1.504	0.399	11.70
Human capital intensity	0.078	0.054	0.067	0.003	0.477
R&D intensity	0.007	0.009	0.003	0	0.041
Homogeneity of locational conditions	86.88	89.47	10.57	33.33	100.00
Entrepreneurial character of industry	775.0	142.4	1526.5	0.0	9061.1
Output growth rate	0.020	0.038	0.019	-0.095	0.165
Average value of relative efficiency	4.031	4.059	0.352	2.308	4.605

Table A3: Characteristics of industries with highest and lowest value of heterogeneity measure

NACE code: Industry	H-area measure 0-100%	Number of firms	Relative efficien-cy	Average firm size	Capital intensity	Human capital intensity	New firm forma-tion rate	Average output growth rate
36.21: Striking of coins	0.756	5	0.190	91	0.0347	0.0218	0.072	-0.071
37.10: Recycling of metal waste and scrap	0.523	49	0.473	86	0.0493	0.0544	1.338	0.001
14.40: Production of salt	0.453	9	0.447	319	0.0714	0.0655	0.052	0.009
22.13: Publishing of journals and periodicals	0.449	197	0.537	167	0.0256	0.2780	0.703	-0.012
14.50: Other mining and quarrying	0.427	7	0.434	80	0.0793	0.0541	0.126	-0.010
22.12: Publishing of newspapers	0.398	252	0.233	342	0.0421	0.1762	0.165	0.041
26.24: Manufacture of other technical ceramic products	0.386	19	0.555	359	0.0730	0.0842	0.077	0.067
25.12: Retreading and rebuilding of rubber tyres	0.385	17	1.000	167	0.0378	0.0239	0.599	-0.004
15.11: Production and preserving of meat	0.381	151	0.654	131	0.0293	0.0235	0.255	-0.024
29.72: Manufacture of non-electric domestic appliances	0.375	53	0.731	230	0.0357	0.1126	0.845	-0.021
31.61: Manufacture of electrical equipment for engines and vehicles n.e.c.	0.057	107	0.444	537	0.0411	0.1105	0.050	0.058
32.30: Manufacture of television and radio receivers, sound or video recording or reproducing apparatus etc.	0.053	167	0.405	365	0.0358	0.1402	0.146	0.070
15.94: Manufacture of cider and other fruit wines	0.045	7	0.876	80	0.0555	0.0424	1.155	0.002
24.65: Manufacture of prepared unrecorded media	0.045	5	0.569	666	0.0574	0.1820	0.666	0.165
31.20: Manufacture of electricity distribution and control apparatus	0.042	723	0.359	541	0.0369	0.1403	0.143	0.046
24.16: Manufacture of plastics in primary forms	0.040	141	0.288	776	0.0527	0.1091	0.096	0.041
15.12: Production and preserving of poultrymeat	0.034	43	0.218	188	0.0292	0.0138	0.109	0.052
15.61: Manufacture of grain mill products	0.020	87	0.101	106	0.0354	0.0344	0.358	0.002
24.17: Manufacture of grain mill products	0.019	5	0.956	89	0.0365	0.0854	0.098	0.047
34.10: Manufacture of motor vehicles	0.014	77	0.149	6783	0.0312	0.1092	0.004	0.045