DEUTSCHE
BUNDESBANK

Measurement matters – Input price proxies and bank efficiency in Germany

Michael Koetter
(Utrecht School of Economics and The Boston Consulting Group)

Discussion Paper
Series 2: Banking and Financial Studies
No 01/2005

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1
Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax No +49 69 9566-3077

ISBN 3-86558-046-7

Abstract

Most bank efficiency studies that use stochastic frontier analysis (SFA) employ each bank's own implicit input price when estimating efficient frontiers. But the theoretical foundation of most studies is a cost minimisation and/ or profit maximisation problem assuming perfect input markets. At the very least, traditional input price proxies therefore contain substantial measurement error. In this paper, we examine the magnitude and direction of this error in cost and profit efficiency (CE and PE) measurement. We suggest two input market definitions to approximate exogenous input prices alternatively and estimate CE and PE of German banks between 1993 and 2003. Our main findings are threefold. First, after accounting for systematic differences across banks, mean CE is sensitive to alternative input prices. Second, distortions of mean PE due to traditional input prices are small. Third, across CE models small cooperative banks located in large western states are identified as top performers. Large banks and those located in eastern states rank lowest.

Keywords: Germany, Banks, cost efficiency, profit efficiency, stochastic frontier analysis, measurement error.

JEL Codes: G20, L11, C51

Non-technical Summary

The last decade witnessed continuous changes in regulation, technology and competition in the global financial services industry. German banks are no exception. Rising cost-income ratios and declining profitability reflect increased competitive pressure. To assess the stability of the banking system, it is therefore crucial to benchmark the performance of banks operating in Germany. This allows identification of role models and problem banks.

One major obstacle to any benchmark methodology is to what extent investigated subjects are comparable. To this end, traditionally employed key performance indicators (KPI) enjoy the advantage of straightforward calculation. But they fail to account for systematic differences between, say, commercial, savings and cooperative banks.

We therefore suggest a complementary methodology to benchmark banks on the basis of their cost and profit efficiency. These measures quantify how successful banks are in minimising cost and maximising profits when converting inputs into outputs. We assume that all banks in Germany have access to the same production technology. At the same time, we acknowledge that systematic differences prevail in terms of size, asset risk, location, regional market concentration and bank pillar membership. Therefore, we explicitly account for these differences.

To estimate optimal profits and costs we furthermore suggest two alternatives to approximate input prices. In contrast to most studies, we utilise regional markets to calculate market prices that banks face when arriving at their production decisions.

Our results identify cost inefficiencies on the order of 8 to 14 percentage points of actually incurred costs. Foregone profits amount to 35 percentage points of actually accrued profits. These inefficiencies emanate primarily among large commercial banks and apex institutions of cooperative and savings banks.

From a regional perspective, banks located in new states suffer from above-average cost inefficiency, but enjoy higher profit efficiency. From a banking group perspective, cooperative banks are among the top performers in terms of cost efficiency, while savings banks enjoy higher mean profit efficiency. From a size perspective, small banks outperform large banks. Utilisation of alternative input prices leads to approximately 5 percentage points lower cost efficiency compared to traditional input prices. Therefore, alternative input price measurement matters.

Nichttechnische Zusammenfassung

Finanzdienstleister sind in den letzten Jahren massiven Veränderungen in den Bereichen Regulierung, Technologie und Wettbewerb ausgesetzt gewesen. Das deutsche Bankwesen stellt dabei keine Ausnahme dar. Steigende Aufwand-Ertrags-Relationen und sinkende Profitabilität reflektieren gestiegenen Wettbewerbsdruck. Zur Beurteilung der Stabilität des Bankensystems ist daher ein Vergleich von Banken zur Identifikation der effizientesten und ineffizientesten Banken von entscheidender Bedeutung.

Bei jeder Benchmarking-Methode ist ausschlaggebend, dass die betrachteten Banken auch tatsächlich vergleichbar sind. Traditionelle Vergleichsgrößen haben zwar den Vorteil, dass sie leicht zu berechnen sind. Ein wichtiger Nachteil liegt jedoch darin, dass systematische Unterschiede, z. B. zwischen den Banksektoren, nicht explizit berücksichtigt werden.

Wir schlagen deshalb eine ergänzende Benchmarking-Methode vor. Kosten- und Profiteffizienz quantifizieren die Fähigkeit einer Bank, bei der Produktion Kosten zu minimieren bzw. Profite zu maximieren. Wir nehmen an, dass alle deutsche Banken Zugang zu der gleichen Produktionstechnologie haben. Gleichzeitig berücksichtigen wir dabei existierende Unterschiede hinsichtlich Größe, Risiko, Einsatzort, Marktkonzentration und Sektorzugehörigkeit.

Um optimale Profite und Kosten zu schätzen, benutzen wir alternative Faktorpreise. Im Gegensatz zu den meisten Studien berechnen wir die Faktorpreise, denen eine Bank ausgesetzt war, in Abhängigkeit des jeweiligen regionalen Faktormarktes.

Die Ergebnisse zeigen, dass die durchschnittliche Kostenineffizienz zwischen 8 und 14 Prozent der tatsächlich realisierten Kosten beträgt. Nicht realisierte Profite betragen etwa 35 Prozent der aktuellen Profite. Ineffizienzen entstehen vornehmlich bei großen Geschäftsbanken sowie den Zentralbanken des Genossenschafts- und Sparkassensektors.

Banken in den neuen Bundesländern leiden unter höherer Kostenineffizienz, zeigen jedoch höhere Profiteffizienz. Regional tätige Genossenschaftsbanken sind überdurchschnittlich kosteneffizient, während lokale Sparkassen überdurchschnittlich profiteffizient sind. Kleine Banken arbeiten effizienter als große Institute. Die Verwendung von alternativen Faktorpreisen beeinflusst insbesondere die Kosteneffizienzmaße. Die so ermittelte Kosteneffizienz liegt im Mittel um etwa 5 Prozent niedriger als die traditionell geschätzte Kosteneffizienz. Der Gebrauch von alternativen Faktorpreisen ist somit von Relevanz.

Contents

Measurement Matters – Input Price Proxies and Bank Efficiency in Germany [#]

1 Introduction

Since its introduction by Aigner et al. (1977), Battese and Corra (1977) and Meeusen and Broeck (1977) stochastic frontier analysis (SFA) has received increasing attention in the literature as a methodology to measure a single firm's efficiency.[1] Decisive virtues of SFA are to account for both (i) random noise, e.g. due to well-known measurement problems, and (ii) systematic differences between banks in the sample due to heterogeneity across banks (Kumbhakar and Lovell, 2000).[2] These features allow a relative comparison of markedly different banks, for example large commercial versus small cooperative or savings banks.[3] In contrast to accounting-based key performance indicators (KPI), SFA rankings therefore account explicitly for both environmental factors and random error.[4]

Independent of whether heterogeneity is accounted for or not, the specification of a cost or profit function under the assumption of perfect input markets is at the core of most of these applications. The error term is composed of white noise and inefficiency. By means of maximum-likelihood techniques, we estimate best-practice frontiers, relative to which all firms in the sample are compared. For each firm, we decompose the error term. Firms deviate to varying degrees from this frontier according to the estimated inefficiency.

But as noted by Mountain and Thomas (1999) the vast majority of SFA studies share a potentially severe measurement error when calculating input prices. The assumption of perfect input markets requires a bank to be a price taker. It purchases inputs given prices that are determined exogenously in the respective market. But most studies approximate

[#] This paper is a successor study to the Marjolin paper presented at the 25th SUERF Colloquium on October 22, 2004. I am grateful for helpful discussions with participants in general and David Humphrey in particular. The paper is the result of a research cooperation between the Utrecht School of Economics and the Research Group at the Banking and Financial Supervision department of Deutsche Bundesbank. I want to thank Clemens Kool, Jaap Bos, Daniel Porath and Christoph Memmel for inspiration and comments. Financial support from The Boston Consulting Group is gratefully acknowledged. I thank the Bundesbank for permission to employ data. The views expressed in this paper are mine and neither reflect those of the Bundesbank nor those of The Boston Consulting Group. All remaining errors are mine.

[1] As early as 1997, Berger and Humphrey surveyed 130 bank efficiency studies.

[2] If systematic differences are not accounted for they are mistakenly identified as inefficiency (Mester, 1997).

[3] Coelli et al. (1998) note that relative rankings imply that only efficiency measures relative to an identical frontier can be compared. This in turn means that efficiency measures from different frontiers, e.g. across alternative studies or sub-samples within studies, cannot readily be compared.

[4] Accounting-based KPI are cost-income ratios, return on assets and equity and interest margins.

input prices by relating individual banks' factor payments to employed production factors.[5] Consequently, these prices are bank-specific rather than market-determined.[6] Such proxies might therefore be poor measures of theoretical counterparts. Poor measurement of true explanatory variables, namely input prices, could distort efficiency estimations substantially (Greene 1993, chapter 9).

This paper addresses the described potential measurement error. We suggest two alternatives to approximate input prices and compare efficiency estimates when using traditional versus alternative input price proxies. The paper is organised as follows. In the next section, we review German bank efficiency studies and the respective approaches to accommodate the discussed measurement error. Section three introduces the theoretical model underlying efficiency measurement, our approach to proxy input prices alternatively and the empirical specification to estimate cost and alternative profit efficiency (CE and PE) for a heterogeneous sample of banks. Section four elaborates on the data used in this study. In section five, we present and discuss our results. We conclude in section six.

2 Input prices and German efficiency studies

To our knowledge, Mountain and Thomas (1999) is the only study that explicitly discusses the potential measurement error due to the use of bank-specific input prices. They argue that under perfect (input market) competition banks should face at a given point in time identical prices.[7] Then, traditional input price proxies lead to biased efficiency results.

To assess the importance of this measurement error, they compare CE from two cost frontier specifications with and without input prices. Efficiency estimates between the full and the parsimonious model differ by 7 percentage points. However, they reject the hypothesis that input price coefficients are simultaneously equal to zero. Despite this result they conclude that input prices should be dropped. They evaluate the benefit from obtaining more precise estimates to outweigh the cost due to a higher degree of inconsistency because of omitting mismeasured proxy variables.[8]

They acknowledge that the use of "true" input prices is highly desirable, but argue that the loss in precision requires that poor proxies be dropped. A first suggestion of Mountain and

[5] Dividing interest expenses by borrowed funds is an example found in virtually every study.
[6] Additionally, they are derived from the dependent variable, which should be explained - namely total cost.
[7] In sum, concentration and competition studies (Bikker and Haaf, 2002 and Hempell, 2004) conclude that a theoretical model from the realm of perfect competition is a defendable assumption for German banking.
[8] According to Aigner (1974), proxy variables should be included even if they contain measurement error.

Thomas (1999) to alleviate this measurement error and to enhance estimation is to either adjust the frontier for regional differences or to restrict estimation of the benchmark to a sample of banks sufficiently akin to each other.[9]

German bank studies follow this suggestion only implicitly. Among the few available studies, three follow the approach to use regional information primarily due to data availability constraints. Lang and Welzel (1996, 1998a, 1998b) employ a sample that is restricted to cooperative banks from Bavaria only.[10] The authors use data collected from a non-commercial database. In sum, these three studies find improving CE during the early 1990s. Average cost inefficiency is around 7 percent. Also, their evidence suggests that smaller cooperatives perform better than larger cooperatives and that mergers between low efficiency targets and better performing ones increase overall efficiency. The authors carefully note, however, that their results only apply to the sector of cooperative banks in general and to those located in the state of Bavaria in particular. Apparently, regional and ownership differences prohibit a generalisation of their results.

Altunbas et al. (2001) further investigate efficiency differences across banking sectors and size classes between 1989 and 1998.[11] To our knowledge, this is the only parametric study that simultaneously examines cost and profit efficiency of German banks. The study supplies estimates of CE and PE against one identical frontier and against sector-specific benchmarks. They find that inefficiencies are highest for commercial institutes and large banks. For the full sample, CE amounts to 83 percent for the banking sector as a whole. The ability of banks to realise potential profits is worse, as average PE amounts to 80 percent. Efficiency improves if measured against a benchmark for a more homogenous sample, i.e. banks of one type only. This indicates that systematic differences that prevail between commercial, savings and cooperative banks are wrongly dubbed inefficiency if compared against a uniform frontier that is not adjusted for these factors.

These results are consistent with studies explicitly examining the importance to account for heterogeneity when estimating bank-specific efficiency (Bos et al., 2004).[12] But the

[9] Note, however, that especially the approach to estimate, say, sector-specific frontiers suffers from the drawback that efficiency measures can then no longer be compared across samples (Coelli et al., 1998).
[10] Hence, systematic differences due to bank size, business mix and regional economic conditions are low.
[11] The sectors are the three "pillars" in German banking: commercial, savings and co-operative banks.
[12] Factors identified to cause heterogeneity are e.g. different regional macroeconomic conditions (Mester, 1997), varying risk profiles (Hughes and Mester, 1993) and bank market characteristics such as concentration (Lozano-Vivas et al., 2002). Apparently, banks can be “grouped” according to many different criteria. But true groups, i.e. markets, remain unknown and therefore these choices are ultimately arbitrary.

potential measurement error pointed out by Mountain and Thomas (1999) is in our view not sufficiently accounted for by adjusting efficient frontiers and/or deviations from it.

To address the potential mismeasurement of input prices more directly, Mountain and Thomas (1999) suggest as a second approach to measure input prices per factor market. We argue that this "second" approach is an additional rather than an equivalent alternative to measure efficiency appropriately. The only study we are aware of that pursues this approach is Berger and Mester (2003). Using data on US banks of three cross-sections in 1984, 1991 and 1997, they derive prices per geographical market. For each bank they calculate input prices faced as an average by all other banks in that particular market, weighted by each bank's respective market share. Thus, they use for each bank an exogenous price determined by competitors in the respective market. However, the study fails to simultaneously account for additional sources of systematic differences across banks. In addition, the authors do not provide a comparison with traditionally specified input prices. Therefore, we do not know how severe this potential measurement error is.

This paper fills the gap. We estimate efficiency adjusted for heterogeneity and employ alternative proxies of input prices as suggested by Mountain and Thomas (1999).

3 Methodology

To quantify the magnitude of efficiency distortions we provide a comparison. First, we introduce a theoretical model and define our variables. Second, we turn to input price proxies according to a benchmark model and two alternative approaches. Third, we introduce the empirical specification that also accounts for heterogeneity.

3.1 Model specification

We follow the intermediation approach to model bank production and define three according outputs.[13] The first output captures interbank loans, y_1, provided. Next, banks produce commercial loans, y_2. Securities are the third output, y_3. A bank uses three production factors to produce outputs. These are fixed assets, x_1, labour, x_2, and total borrowed funds, x_3. These in- and output categories represent the most frequently used ones in the literature. But note as a caveat that across different banks the quality of inputs

[13] For an extensive discussion of possible models of bank production see Freixas and Rochet (1997).

may differ systematically.[14] Also, especially large banks produce outputs such as custody and advisory services that are not explicitly specified here. As far as the data permits, we therefore account for a number of systematic differences across banks of different pillars, region and further characteristics as discussed below in our empirical specification.

A bank faces an input price vector of w_i, where i indexes the input used. As the calculation of these input prices is at the core of this paper, we devote the next sub-section to a discussion of our baseline and alternative models. In transforming inputs into outputs, we account for the role of equity, z, as an alternative to fund outputs (Hughes and Mester, 1993). The transformation function of the banking firm is depicted by $T(y,x,z)$. Note that this implies the assumption that all banks employ the same production technology. We assume therefore that a single federal frontier can be estimated for German banks.[15] We employ total operating cost, *TOC,* as dependent variable in the cost minimisation problem and profits before tax, *PBT*, in the alternative profit maximisation problem.[16] We assume that banks are price takers in input markets. To produce a given vector of outputs y, banks minimise cost by choosing input quantities x_i at given input prices, w_i. Using these definitions, the cost minimisation problem is written as[17]

$$\begin{aligned} &C(y, w_i) = \min_x \sum_i w_i \, x_i \\ &\text{s.t. } T(y, x, z) \leq 0. \end{aligned} \tag{1}$$

The Lagrangian of this constrained optimisation is written as

$$L = \sum_i w_i \, x_i - \lambda T(\bullet). \tag{2}$$

We take partial derivatives with respect to each input, x_i, and the multiplier, λ. Setting these equal to zero and simultaneously solving for x_i results in optimal input demand functions, $x_i^*(y,w_i,z)$, which in our model are also conditional on the available level of

[14] For example, investment bankers at large commercial banks are a different input compared to relatively low-skilled labour employed at branches of regionally active small banks. Unfortunately, the available data does not allow for a more detailed granularity of inputs according to such quality differentials.
[15] Alternatively, one could estimate separate frontiers per group. But that approach suffers from the incomparableness of efficiency measures from different frontiers and the ultimate arbitrariness of "correct" groups. Instead, we choose here to adjust the frontier for systematic differences across banks.
[16] To conserve on space we illustrate here only the cost minimisation problem. The alternative profit model by Humphrey and Pulley (1997) is similar and we refer to differences via footnotes.
[17] In the alternative profit model, banks possess pricing power on the output side. The pricing opportunity set is $H(p,y,w,z)$, where p denotes output prices, and constitutes an additional constraint next to technology.

equity, z.[18] The minimum cost level is then obtained by substituting the optimal input demand functions into the total cost function given by equation (1), resulting in[19]

$$C^* = \sum_i w_i \, x_i(y, w, z) = C^*(y, w, z). \tag{3}$$

If truly exogenous prices under perfect competition are available, any empirical specification of equation (3) suffers from identification problems.[20] However, we assume that input prices can differ between banks due to some degree of market imperfections. In fact, the vast majority of studies make this assumption implicitly. In the most extreme case, banks can set prices in input markets. Consequently, they would minimise cost by choosing prices w_i that they are willing to pay for given amounts of input.[21] Here, we abstain from these extreme assumptions. Instead, we focus on the magnitude of potentially distorted efficiency measures due to the use of traditional proxy variables. We therefore turn next to our suggestion as how to obtain alternative input price measures.

3.2 Input price measurement

In the cost model, the factor prices for inputs, w_i, are exogenous variables.[22] They are assumed to be determined in their respective factor markets and enter the production model in section 3.1 as given. Then, utilisation of observed equilibrium factor market prices would be strictly speaking the correct choice to empirically specify equation (3).

But due to the lack of such data, researchers employ according to Greene (1993) proxy variables that represent appropriate measures of theoretical counterparts. Ideally, we would therefore use observed contractual wage rates for the price of labour, rental prices per square metre of office space and interest rates per unit of borrowed funds, respectively. But

[18] The profit maximisation problem yields, in addition, optimal output prices $p^*(y,w,z)$.

[19] Accordingly, maximum profits, $\pi^*(y,w,z)$, depend only on exogenously determined input prices and available equity to produce a given output. This curbs the frequently encountered lack of available output prices. It allows thereby to evaluate banks' success to realise opportunities in output markets.

[20] In a given year, input prices are in such a case constant and could not be disentangled from the intercept.

[21] We also test empirically if banks minimise cost with respect to prices. Then, an optimal cost function $C^*(y,x_i,z)$ depends on input quantities rather than prices. Also, we followed Mountain and Thomas (1999) and simply dropped the w_i's entirely, leaving us with an optimum cost function depending on output quantities and equity $C^*(y,z)$, only. These specifications were not estimable due to collinearity or residuals exhibiting a skew that contradicted the requirements of cost and/or profit inefficiency to prevail (Waldmann (1982)). This suggests that the respective specifications are inappropriate.

[22] The same holds throughout for the alternative profit model.

such data are usually also unavailable.[23] To obtain input price proxies, most studies (e.g. Altunbas et al. (2001), Lang and Welzel (1996)) calculate input prices per bank.

Our baseline model, model 1, mimics this traditional approach. For each bank, we obtain the price of fixed assets, w_1, by dividing depreciation and other expenditure on fixed assets over the volume of fixed assets. The price of labour is calculated as an average wage rate, w_2, by relating the Euro amount of personnel expenses to the number of full-time equivalent employees (FTE). Finally, we approximate the price of borrowed funds, w_3, by dividing interest expenses over total borrowed funds.

We suggest to derive alternative input prices by an approach akin to Berger and Mester (2003). We rely on the indications provided by studies on German banking and define markets, *m*, in terms of region and banking type. In the first alternative input price model, model 2, we assume that all banks operate on regional markets. In the second alternative price model, model 3, we assume that large banks compete on a federal market.[24]

Consider the market for fixed assets first. In model 2, a bank that is restricted to branch in a certain region will naturally rent physical assets in the vicinity. It competes for real estate with other banks also branching in this particular region.[25] On the one hand, it is a safe assumption for small banks that office equipment or main administrative real estate is only demanded on the regional market. On the other hand, one might object to this definition of an input market on grounds of nationwide operations of large banks. In fact, large banks might set-up regional administrative centres only where prices are most favourable. Therefore, we assume in model 3 that large banks constitute one federal market for real estate on which they compete with each other. On this federal market, then, each bank faces exogenous rather than traditionally employed endogenous input prices.

With respect to labour, we assume that the majority of banks draw from regional labour pools. We assume that the labour force in Germany is not flexible enough to move from, say, Bavaria to Bremen to conduct standardised banking tasks. In model 2, we therefore assume that the majority of employees are recruited in the respective region. However, we assert that highly specialised tasks like risk management, investment banking or financial engineering require experts. This highly skilled labour force is probably readily willing to

[23] In addition, there are at times substantial conceptual problems as to what the respective measurement units are. For example, interest paid does not cover all costs associated with borrowed money as fees are neglected.
[24] This group comprises large commercial banks, apex institutions of savings and cooperatives and Postbank.

move. We assume the majority of such highly specialised tasks are conducted either by comparatively large banks or the respective apex institutions on behalf of their local banks. Therefore, we construct in model 3 prices for labour to originate from a federal market for large banks as above. In contrast, all other banks continue to recruit locally.

Finally, the market for borrowed funds may depend to a larger extent on local conditions than one expects. We know from various monthly reports of the Bundesbank that the primary sources of funding are for most banks customer deposits. In model 2, we therefore argue that in demanding funds, for example, a local commercial bank faces competition primarily from the savings bank next door. But in contrast to small banks, large institutes rely more intensively on securitised debt and interbank lending. This lower dependence on a local customer base induces us to apply in model 3 the previous logic of one federal market for large banks.

Summing up, we treat each bank in model 1 to operate in its own market. In model 2, we specify exogenous input prices on the basis of regional markets. We use available regional identifiers for each bank and define counties as regional markets. In contrast, we assume in model 3 that large banks compete on one federal input market.

Next, we suggest how to derive input price proxies per market. We argue that the prices banks face can be approximated by the average price paid for an input factor in the market. This implies utilisation of average input prices, w_i, per county for all banks in model 2. In model 3, we use the average paid by large banks from one federal market as input price proxies. Because banks face prices, these market-specific averages are calculated for each bank excluding the bank itself. Table 1 illustrates the different calculations for the benchmark and a market model with an example for labour.

Let w_{ik} denote input prices for each bank, where $k=1,..,K$ indexes a bank and K is the total number of banks in a market.[26] We denote markets by $m=1,..,M$, where M depicts the total number of markets, i.e. counties. In the example above we have four markets m in total, as exhibited by the fourth column in table 1. Each contains three banks, respectively.

[25] Note, that the true market price for real estate, and all other inputs, also depends on the demand of other agents. Unfortunately, we do not have according data at our disposal.

Table 1: Illustration of input price calculations

Bank	**Input i = L**		**County**	**Benchmark model**			**Market model**		
k	x_i	FC_i		*m*	*Calculation*	w_{ik}	*m*	*Calculation*	w_{ik}
1	100	20	1	1	20/100	0.2	1	(0.2+0.12)/(3-1)	0.16
2	1000	250	2	2	250/1000	0.25	2		0.75
3	2000	400	1	3	400/2000	0.2	1	(0.2+0.12)/(3-1)	0.16
4	500	25	2	4	25/500	0.5	2		0.175
5	250	30	1	5	30/250	0.12	1	(0.2+0.2)/(3-1)	0.20
6	750	75	2	6	75/750	0.1	2		0.15
7	1600	120	3	7	120/1600	0.75	3		0.15
8	900	203	4	8	202.5/900	0.225	4		0.125
9	300	52.5	3	9	52.5/300	0.17	3		0.10
10	1200	180	4	10	180/1200	0.15	4		0.163
11	800	100	3	11	100/1800	0.12	3		0.125
12	400	40	4	12	40/400	0.10	4		0.188
Mean						**0.148**			**0.148**
Stdev						**0.063**			**0.036**

Note: All numbers are hypothetical for illustrative purpose.

L: Labour; x_i: Amount of labour measured in full-time equivalents; FC_i: Factor cost of labour measured in thousands of Euros.

The benchmark case can be regarded as modelling each bank as a market of its own, i.e. the number of banks equals the number of markets, *K=M*. This is illustrated in table 1 under the heading Benchmark model. Alternatively, we calculate the input price proxy as the average market price excluding the bank's own position. The optimum cost function, then, employs for each bank the average price of all other banks in the region. More formally, a bank *l* in market *m* faces a price of $w_{il} = \sum_{k=1,l\neq k}^{K} w_{ik} /(K-1)$. To illustrate this procedure for one market, consider the three rightmost columns under the heading Market model. Within market *m=1*, the first bank, *k=1*, faces an input price equal to the average input price paid by the remaining banks in the market, i.e. banks *k=3, 5*. This procedure leaves the average input price paid by all banks unaffected. However, the variation of input prices is reduced as can be seen from the last row in table 1. Intuitively, the closer the market resembles perfect competition, the lower the variation, because price differences ultimately approach zero until a single market-clearing price prevails.

We discuss next the empirical and intuitive implications of this approach with regard to the estimated level of efficiency. Recall that in model 1 we explain on the left-hand side (LHS) the observed total cost with an observed output vector and some proxies for input prices on the right-hand side (RHS), $C = \sum_i w_i^p x_i$, where the superscript *p* indicates the proxy. We assume that $w_i^p = w_i^*$, i.e. that the proxy reflects true equilibrium prices. These proxies are

[26] We calculated all prices per year and suppress time subscripts here for ease of exposition.

constructed by using different parts of total cost C_i, for example the category of interest expenses of total costs, and divide them by some input quantity, for example the volume of borrowed funds, $w_i^p = C_i / x_i$. One can argue that this amounts to employing a scaled fraction of total cost on the RHS in order to explain the total on the LHS.

Thus, we would expect a high explanatory power because the RHS variables are basically extracted from the LHS. By definition, the numerators of our ratio that should approximate true equilibrium prices sum to the LHS (or at least to the major share of it). Hence, for a bank with input prices substantially above average, employed variables still explain a lot of the observed total costs by construction. Thereby, total error would be underestimated, which is crucial for SFA. This is because efficiency is extracted from this total error.

Therefore, our suggested approach to derive proxy variables for equilibrium input prices emphasises the deviations of banks' factor costs over quantities from what other banks have paid under similar conditions. It is perceivable that a frontier using market prices is less suited to explain observed costs, as those proxies are actually not directly derived from the LHS observation. Then, the fit of the frontier might deteriorate. This would imply that the total error of our estimation increases. Given the parametric nature of SFA, we assume ex ante that a certain portion of the error distribution, i.e. deviations from the benchmark, is due to inefficiency (Coelli et al., 1998). Then, a higher total error leads us to expect higher levels of inefficiency, too.[27] Therefore, we expect higher levels of inefficiency for models 2 and 3 compared to model 1 if it is true that total error increases when using alternative input price proxies.

Apart from the level of efficiency, a major strength of SFA is the ability to rank firms. Therefore, we are particularly interested in the distribution of efficiency scores in order to identify best and worst in class banks. Note that according to our alternative input price proxies, those banks that paid the highest price within their market are now modelled to face substantially lower prices (and vice versa). This is appealing because for a bank that incurred high total costs on the left hand side we now employ a market price, which is substantially lower. Recall that for such a bank the high implicit input price emerged probably from dividing high factor payments by relatively small factor quantities employed. We could flip the problem around and state that the bank has paid for that

[27] As in Kumbhakar et al (1991), we incorporate environmental factors directly. This implies an offsetting effect. We expect that alleviating a potentially omitted variable bias improves estimation (Mester, 1997).

amount of input a too high price relative to what competitors have paid. In the benchmark model this need not imply huge inefficiency, as input prices per bank are ultimately derived from total cost of this bank only and, in turn, explain a lot of the LHS by construction. In the alternative approach, however, this bank would be labelled inefficient because it incurred the same (too high) costs, while facing a market price considerably below its own factor payment per factor quantity.

In sum, we assume in this paper that markets can be defined in terms of regions. The price a bank faces to purchase its inputs is determined in these markets. For each bank, *k*, this price is the average of prices paid by all other banks in that market excluding the bank's own price. The specification of exogenous input prices should not only affect the level of efficiency, but also the distribution of efficiency scores. Next, we introduce the empirical specification to estimate a reduced form of equation (3).

3.3 Empirical specification

We follow Kumbhakar and Lovell (2000) and use a stochastic frontier model that accommodates heterogeneity. To estimate a reduced form of equation (3), we rely largely on established procedures to measure the effects of misspecified input prices as directly as possible.[28] As in Lang and Welzel (1996), we employ a multi-output translog function and include time trend variables to capture technical change to estimate a federal cost frontier. For any bank, *k*, this cost function takes on the form[29]

$$\begin{aligned}
\ln TOC_k(w_k, y_k z_k) = \alpha_0 &+ \sum_{i=1}^{3}\alpha_i \ln w_{ik} + \sum_{i=m}^{3}\beta_m \ln y_{mk} + \delta_0 \ln z_k \\
&+ \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\alpha_{ij} \ln w_{ik} \ln w_{jk} + \frac{1}{2}\sum_{m=1}^{3}\sum_{n=1}^{3}\beta_{mn} \ln y_{mk} \ln y_{nk} \\
&+ \frac{1}{2}\sum_{i=1}^{3}\sum_{m=1}^{3}\gamma_{im} \ln w_{ik} \ln y_{mk} + \frac{1}{2}\delta_1 (\ln z_k)^2 \\
&+ \sum_{i=1}^{3}\omega_i \ln w_{ik} \ln z_k + \sum_{i=1}^{3}\zeta_m \ln y_{mk} \ln z_k + \eta_0 t + \frac{1}{2}\eta_1 (t)^2 \\
&+ \sum_{i=1}^{3}\kappa_i \ln w_{ik} t + \sum_{m=1}^{3}\tau_m \ln y_{mk} t + \delta_2 \ln z_k t + \varepsilon_k .
\end{aligned} \tag{4}$$

Thereby, we reduce the total error and the level of inefficiency. Which of the two effects, accounting for heterogeneity versus alternative price proxy measurement, dominates, remains an empirical question.

[28] Bauer et al. (1998) and Berger and Humphrey (1997) review alternative stochastic frontier models.

[29] In the alternative profit model, we employ ln*PBT* as dependent variable.

As outlined in section 3.1, w_i denotes input prices, y_m denotes outputs and z depicts the level of equity capital.[30] We substitute, input prices w_i, from the three respective market definitions discussed in section 3.2 to investigate the effects on efficiency levels, rankings and robustness of estimation results. To measure inefficiency, we assume a composed error term in equation (4). The error term, ε_k, consists partly of random noise, v_k, and partly of inefficiency, u_k. For a cost frontier, inefficiency implies above frontier costs. Therefore, inefficiency enters the error term with a positive sign, leading to $\varepsilon_k = v_k + u_k$.[31]

In addition, we follow Bos et al. (2004) and specify a vector of exogenous factors, h_k. These environmental and bank-specific factors capture systematic differences between the different banks in our sample. We assume that these sources of heterogeneity influence the distribution of deviations from the efficient frontier. Kumbhakar et al. (1991) suggest a single-stage approach to allow h_k to influence the mean of the inefficiency distribution.[32]

We specify h_k to contain four variables and three dummies.[33] The first variable accounts for findings reviewed by Berger (2003) that efficiency varies across banks of different size. We therefore include the log of total assets, *TA*. Another characteristic difference between alternative banking groups refers to the risk associated with banks' balance sheet assets. To capture the regulatory risk, we therefore construct a ratio, *RISK*, that relates total risk weighted assets to gross total assets.[34] Clark and Siems (2002) point out that banks increasingly engage in off-balance sheet activities, *OBS*, and therefore we include the log of *OBS* in the inefficiency term, too. The rationale to do so is that banks increasingly attempt to manage various risks actively with these instruments.[35] Finally, we explicitly account for variations in the respective county's bank market structure. We employ a Hirschman-Herfindahl index, *HHI*, per year and county on the basis of total gross assets as to measure concentration and competitiveness.[36] The remaining terms that are included in h_k are dummy variables for the banking groups of local savings and cooperatives. Thereby, we account for systematic differences, such as e.g. previously mentioned quality

[30] Upon estimation, we impose the required restrictions of linear homogeneity in input prices and symmetry of cross-partial derivatives of the conditional factor demand functions as in Lang and Welzel (1996).
[31] For the profit frontier, inefficiency reduces profits and is therefore subtracted, leading to $\varepsilon_k = v_k - u_k$.
[32] For a discussion of the various pitfalls of first estimating efficiency and then exposing it to a set of explanatory variables see Kumbhakar and Lovell (2000).
[33] Our choice is based on results from the literature and talks to practitioners at the Bundesbank.
[34] The risk weights follow the 1988 Basle Accord.
[35] Clark and Siems (2002) specify OBS as part of the efficient frontier. Kumbhakar and Lovell (2000) call it a judgement if a variable influences the technology that determines the frontier or deviations from it. We considered both ways and decided on the basis of log-likelihood values to specify them to influence the latter.
[36] Hempell (2004) discusses virtues and pitfalls of this measure's adequacy to assess competitiveness.

differences of inputs across banking groups. The final dummy variable identifies banks located in the East. This way we accommodate less buoyant real economic conditions. We provide descriptive statistics on SFA and heterogeneity variables in the next sub-section.

As in Kumbhakar et al. (1991), we assume the random error term v_k to be i.i.d. with $v_k \sim N(0, \sigma^2_v)$ and independent of the explanatory variables. The inefficiency term is $u_k \sim |N[(a_0+\mathbf{d'}\ h_k), \sigma^2_u]|$, where **d'** is a vector of parameters to be estimated. We use OLS estimates as starting values when maximising the likelihood function, derived by Aigner et al. (1977). We employ their re-parameterisation of $\lambda=\sigma_u\ /\sigma_v$ and $\sigma^2=\sigma^2_u+\sigma^2_v$. Consequently, λ indicates the ratio of standard deviation attributable to inefficiency relative to the standard deviation due to random noise. An insignificant estimate of λ means that no inefficiency prevails. All of the error is due to random noise and specification of a stochastic frontier model is inappropriate.

We estimate firm-specific efficiency according to Jondrow et al. (1982) and use the conditional expectation of u_k given ε_k to calculate a measure of CE as[37]

$$CE_k = \{E[\exp(u_k) \mid \varepsilon_k]\}^{-1}. \tag{5}$$

Note an important implication of this procedure to obtain efficiency estimates with regard to the interpretation of individual coefficients estimated for h_k. A positive coefficient for a heterogeneity variable h_i for bank k implies a higher mean value for the truncated inefficiency distribution, $f(u)$. It, however, need not imply a higher inefficiency score for bank k. This is because the latter is conditional on the total error, ε_k. While the effect of an increase of h_i implies a higher mean of $f(u)$, total error for this bank, ε_k, might be sufficiently small to still result in a lower than mean efficiency for the individual firm. Put differently, the effect of an exogenous variable's influence on efficiency depends not only on the influence on the truncated distribution, but also on the position (and shape) of the distribution of total error. To evaluate the effect of an exogenous variable, h_{ik}, on efficiency, Kumbhakar and Lovell (2000) suggest to use the derivative of the conditional mean of the inefficiency term with respect to the exogenous factor, i.e. $[\partial E(u_k|\varepsilon_k)/\partial h_{ik}]$. In our result section, we are primarily interested in whether our choice of h is significant and how the entirety of exogenous factors influence efficiency. Therefore, we do not investigate marginal effects but focus on differences of efficiency scores across models.

Efficiency measures take on values between 0 and 1. The latter indicates a fully efficient bank. The value indicates which percentage of observed cost would have been enough to produce the observed output if the bank was fully efficient. Firm-specific efficiency estimates further allow us to analyse if rankings are sensitive to measurement error. We turn next to a brief description of the employed data.

4 Data

For the years 1993 to 2003, we use data from unconsolidated, annual balance sheets, profit and loss accounts and audit reports either reported to or compiled by the supervision department of the Bundesbank. The data comprise 32,211 observations on commercial, savings and cooperative banks. Table 2 depicts descriptive statistics of both variables used in the deterministic kernel and those accounting for heterogeneity in the error.

Table 2: Mean production and heterogeneity variables 1993-2003

Variable	Name	Mean	Sd	Min	Max	N
			SFA			
y_1	**Interbank loans** [1)]	361	4,200	0.001	154,000	32,211
y_2	**Commercial loans** [1)]	723	6,640	0.002	315,000	32,211
y_3	**Securities** [1)]	339	3,500	0.003	218,000	32,211
w_1	**Price of fixed assets** [2)]	18.25	139.23	0.219	17,700.6	32,211
w_2	**Price of labour** [3)]	49.62	17.46	0.377	1,133.1	32,211
w_3	**Price of borrowed funds** [2)]	3.98	25.68	0.047	4,585.8	32,211
z	**Equity** [1)]	55.7	496.0	0.065	21,600	32,211
TOC	**Total operating cost** [1)]	81.7	756.0	0.137	40,500	32,211
PBT	**Profit before tax** [1)]	9.9	66.3	-989	3,210	32,211
			Heterogeneity			
TA	**Total assets** [1)]	1,540	15,400	0.678	742,000	32,211
OBS	**Off-balance sheet** [1)]	207	2,970	0.000	128,000	32,211
HHI	**Market concentration** [4)]	3,316	1,885	566	10,000	32,211
RISK	**Average asset risk**	0.604	0.131	0.000	0.996	32,211

Note: Reported statistics refer to data pooled over the whole period.
[1)] In millions of Euro; [2)] In percent; [3)] In thousands of Euro; [4)] In points per county; between 1 (perfect competition) and 10,000 (monopoly).

Two issues in table 2 deserve attention. First, the distribution of both SFA and heterogeneity variables clearly illustrate the differences across firms in our sample. For example, some banks only had outstanding loans on the order of some thousand Euro. In contrast, the largest bank recorded a position of outstanding commercial loans worth € 315 billion in a single year. With respect to input prices the standard deviation and extreme

[37] Similarly, a measure of PE is calculated as $PE_k=E[\exp(-u_k)|\varepsilon_k]$. This measure indicates the percentage of actual profits relative to what the bank could have realised given its input quantity and output price mix.

values underline potential measurement problems. It is hard to believe that the observed maxima of input prices, w_i, accurately reflect the competitive position of a bank. Hence, an alternative to approximate input market prices seems warranted.

Second, the "implausible" values for input prices may be due to faulty data. Some studies exclude extreme observations. But Maudos et al. (2002) caution that this cure may be worse than the disease. They argue that any truncation point is ultimately arbitrary. Given the exceptional quality of our data, our concerns of this kind are limited. Only around 150 out of 32,211 observations exhibit input prices for fixed assets and/or borrowed funds above a 100 percent.[38] As a check we exclude extreme observations for a range of cut-off points. Results are qualitatively not affected. This underpins the ability of stochastic frontier analysis to accommodate random error appropriately for this sample.

In sum, a comparison of bank performance should consider measurement error, random noise and heterogeneity in a more explicit manner. Consideration of random noise is at the very heart of stochastic frontier analysis. We take heterogeneity into account by specifying a range of bank-specific factors in addition to the standard production plan used in most studies. To address the issue of appropriate input prices we employ our alternatives discussed above. Therefore, we turn next to our empirical results to assess the impact of this measurement error on efficiency measures.

5 Results

In this section, we present estimation results for CE and PE from models 1 through 3. First, we discuss cost and, second, profit frontier estimations. For each, we compare the benchmark model 1 and alternative input price models 2 and 3, respectively. Subsequently, we compare efficiency results across region and banking groups. Third, we compare efficiency to traditional KPI, identify extreme performers and describe their characteristics.

5.1 Cost frontier

We start by discussing cost estimation results, continue with an inspection of regional differences of CE and close with a CE comparison between banks from alternative sectors.

[38] Note also that around 75 percent of these extreme observations are banks located in the East. These banks may enjoy special tax breaks meant to subsidise structurally weaker regions. For example, if a tax-minimising firm books high depreciation, this can lead to "implausible" prices for fixed assets because the numerator of w_3 is inflated. But such a bank is certainly still in operation in the market. Therefore, it should be considered in a benchmark. Exclusion may in fact bias efficiency measures as well.

Table 3. Cost frontier estimates 1993 - 2003

Model	CE model 1		CE model 2		CE model 3	
N	32,211		32,211		32,211	
Iterations	84		83		87	
Log likelihood	18,058		14,712		14,594	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Constant	-5.108	0.000	-4.701	0.000	-4.626	0.000
lny_1	0.287	0.000	0.362	0.000	0.344	0.000
lny_2	0.405	0.000	0.138	0.000	0.187	0.000
lny_3	0.338	0.000	0.422	0.000	0.435	0.000
lnw_1	0.277	0.000	0.632	0.000	0.557	0.000
lnw_2	0.722	0.000	0.571	0.000	0.624	0.000
lnz	0.035	0.107	0.105	0.000	0.037	0.103
½ lny_1lny_1	0.032	0.000	0.026	0.000	0.026	0.000
½ lny_1lny_2	-0.114	0.000	-0.067	0.000	-0.068	0.000
½ lny_1lny_3	-0.035	0.000	-0.045	0.000	-0.046	0.000
½ lny_2lny_2	0.085	0.000	0.111	0.000	0.111	0.000
½ lny_2lny_3	-0.116	0.000	-0.095	0.000	-0.100	0.000
½ lny_3lny_3	0.048	0.000	0.042	0.000	0.042	0.000
½ lnw_1lnw_1	-0.036	0.000	-0.031	0.000	-0.032	0.000
½ lnw_1lnw_2	0.039	0.000	0.196	0.000	0.270	0.000
½ lnw_2lnw_2	-0.174	0.000	-0.403	0.000	-0.448	0.000
½ lnz^2	-0.142	0.000	-0.052	0.000	-0.054	0.000
lny_1lnw_1	0.001	0.475	-0.010	0.000	-0.013	0.000
lny_1lnw_1	0.020	0.000	-0.004	0.265	0.001	0.823
lny_1lnw_1	-0.003	0.123	0.016	0.000	0.024	0.000
lny_1lnw_1	-0.058	0.000	0.009	0.033	0.001	0.786
lny_1lnw_1	0.006	0.000	0.043	0.000	0.043	0.000
lny_1lnw_1	-0.035	0.000	-0.067	0.000	-0.066	0.000
lny_1lnz	0.037	0.000	0.019	0.000	0.021	0.000
lny_2lnz	0.050	0.000	-0.011	0.000	-0.010	0.000
lny_3lnz	0.030	0.000	0.025	0.000	0.028	0.000
lnw_1lnz	-0.019	0.000	-0.112	0.000	-0.120	0.000
lnw_2lnz	0.097	0.000	0.127	0.000	0.134	0.000
T	0.091	0.000	0.048	0.000	0.054	0.000
T^2	-0.002	0.000	0.001	0.000	0.001	0.003
lny_1T	0.003	0.000	0.001	0.000	0.002	0.000
lny_2T	-0.005	0.000	0.000	0.734	0.000	0.317
lny_3T	0.003	0.000	0.001	0.008	0.001	0.002
lnw_1T	0.002	0.000	0.015	0.000	0.017	0.000
lnw_2T	-0.032	0.000	-0.037	0.000	-0.041	0.000
$lnNPD$	0.012	0.000	0.016	0.000	0.016	0.000
			Heterogeneity in inefficiency			
Constant	-8.394	0.000	-0.530	0.000	-1.432	0.000
lnTA	0.369	0.000	-0.022	0.001	0.023	0.001
RISK	0.339	0.000	0.133	0.000	0.100	0.000
lnOBS	0.149	0.000	0.075	0.000	0.080	0.000
HHI	-0.137	0.000	0.084	0.000	0.082	0.000
Savings banks	5.966	0.000	3.892	0.000	3.935	0.000
Cooperative banks	6.135	0.000	4.050	0.000	4.185	0.000
East	-5.172	0.000	-3.542	0.000	-3.677	0.000
λ	7.822	0.000	4.607	0.000	4.701	0.000
σ	0.755	0.000	0.552	0.000	0.564	0.000

5.1.1 Estimation results and cost efficiency

Table 3 depicts parameter estimates for cost frontiers according to models 1 through 3. We employ the cost of borrowed funds, w_3, to impose homogeneity as in Lang and Welzel (1996). Most parameter estimates are significant. Due to interaction terms, the interpretation of single coefficients of the deterministic kernel is not straightforward. We therefore abstain from any conclusions of this kind.

Individual coefficients that account for heterogeneity in the inefficiency distribution, h, are also significantly different from zero. Therefore, not all non-random deviations from the frontier are due to the inability to employ resources in optimal proportions. Specification of a set of explanatory variables that influence the truncated distribution of inefficiency improves the fit significantly.[39] But as the effect on efficiency depends ultimately on total error we abstain from the interpretation of single coefficients. Instead, we concentrate on two issues. First, whether accounting for heterogeneity leaves any further systematic deviations that are identified as inefficiency. Second, to assess the change in efficiency due to the use of alternative input prices.

Regarding the former consider the estimate of λ, the ratio of the standard deviation attributable to inefficiency to the standard deviation due to random noise. The parameter is significantly different from zero in both the benchmark and all alternative models. We therefore conclude that inefficiency prevails even after accounting for heterogeneity and no matter how input prices are proxied. Parameter estimates of λ and σ in model 1 are substantially larger compared to both alternative price models. This indicates that the distribution of estimated inefficiencies is influenced by the use of alternative input prices.

To quantify the differences in efficiency scores, table 4 provides descriptive statistics of CE estimates according to the three models.

Table 4: Mean cost efficiency estimates 1993 - 2003

CE	Mean	SD	Skew	Kurt	Min	Max	N
Model 1	0.915	0.077	-2.752	10.859	0.571	0.985	32,211
Model 2	0.862	0.093	-2.074	7.175	0.509	0.982	32,211
Model 3	0.877	0.090	-2.324	8.425	0.508	0.982	32,211

[39] We reject a log-likelihood ratio test that all heterogeneity parameters are simultaneously equal to zero at the 1% level.

Mean cost inefficiency according to model 1 is on the order of 8 percent. This is less than the 20 percent often found in the literature. However, it is in line with studies that also account for heterogeneity and thereby disentangle CE from random noise on the one hand, and other exogenous factors on the other. The standard deviation and skew in the benchmark model indicate that many banks are located fairly close to the frontier. However, the performance is diverse as underlined by a difference between best- and worst-performing banks of 42.4 percent. We hypothesised earlier that alternative input price measures affect the level and distribution of CE.

Mean CE in model 2 confirms this notion. It is five percent below the efficiency score in model 1. The difference between best- and worst-performing banks increases. This indicates that model 2 identifies extreme performing banks with more emphasis. At the same time, alternative market price proxies locate the majority of banks close to full efficiency. But worst performing banks are further away from full efficiency as compared to the benchmark model. This explains lower industry average CE compared to the benchmark case. The use of exogenous input prices seems to highlight extreme performers. A reason could be that in perfect markets healthy banks gravitate to market averages while those banks close to market exit stand out more clearly as poor performers.

Other studies report that large banks are less efficient than small banks. This result could partly be due to measurement error if these banks operate on one federal rather than on multiple regional markets. In model 3, we account for this possibility and calculate input price proxies accordingly. Qualitatively, the results between models 2 and 3 do not differ substantially. Mean CE slightly improves to 87.7 percent, thus being around four percent below the benchmark case.

In addition to altered mean efficiency levels, we find that CE rankings are affected by the use of alternative input prices. Rank-order correlations between model 1 and the two alternative models 2 and 3 are 80.4% and 82%, respectively.[40] As rankings are not perfectly correlated, we conclude that our alternative price proxies do not identify the same banks as best and worst performers.[41] We will address the issue in section 5.3. Beforehand, we identify CE differences between banks of different regions and banking type.

[40] Measured by Spearman's ρ, significant at the 1% level.
[41] Differences of CE ranks between models 2 and 3 are negligible given a rank order correlation of 98.9%.

5.1.2 Cost efficiency across regions

Table 5 depicts differences of mean CE for each model per state. We compare CE to the ratio of general administrative expenses to the raw result, the CI ratio. The last line reports a rank-sum test statistic if efficiency measures are significantly different from each other.

Table 5: Mean cost efficiency per state

State	Model 1	Model 2	Model 3	CI [1)]	N
Baden-Wuerttemberg	0.943	0.900	0.913	70.7	5,949
Bavaria	0.935	0.885	0.899	71.5	7,593
Berlin+	0.824	0.755	0.778	75.5	200
Bremen	0.898	0.853	0.867	71.9	149
Hamburg	0.849	0.746	0.766	59.3	295
Hesse	0.904	0.858	0.871	73.1	3,209
Lower Saxony	0.911	0.858	0.875	72.0	3,344
North Rhine-Westphalia	0.932	0.884	0.898	50.2	5,166
Rhineland-Palatinate	0.926	0.873	0.889	71.9	2,022
Saarland	0.933	0.878	0.895	72.8	434
Schleswig-Holstein	0.894	0.846	0.862	72.8	1,111
Mecklenburg-W. P.+	0.808	0.718	0.740	76.3	417
Brandenburg+	0.780	0.682	0.702	77.0	417
Saxony+	0.783	0.691	0.712	75.1	557
Thuringia+	0.823	0.725	0.746	74.5	593
Saxony-Anhalt+	0.801	0.693	0.708	75.4	755
Federal average	0.915	0.862	0.877	68.5	32,211
Kruskall Wallis	7231.2*	7702.1*	7833.5*	843.1*	

Note: A '+' indicates eastern states.

[1)] General administrative expenses to raw result, measured in percent.

* Significant at the 1% level.

Two major conclusions arise. First, the CI ratio yields by and large similar rankings as CE measures. However, some differences prevail. For example, those states hosting most of the small cooperative banks, Baden-Wuerttemberg and Bavaria, are according to all CE measures ranked first and second. In turn, based on the CI ratio the former state is ranked 3rd and the latter is ranked 4th. Two potential implications arise. On the one hand, CE measures might simply be wrong and therefore useless. The ability to demand input quantities in correct proportions subject to prevailing prices does not matter for the economic success of banks. We find this interpretation hard to believe. Alternatively, Bauer et al. (1998) suggest that efficiency measures contain additional information compared to CI ratios. For example, if increasing competition squeezes interest margins for all banks in the market, CI ratios naturally rise. However, how well a single bank converted inputs into outputs cannot be assessed from CI ratios alone.

The second inference from table 5 is confirmation of an "East" effect. All new states yield mean CE significantly below the federal average. Whether this result is due to systematically lower managerial skills or due to structural deficiencies not captured by a simple dummy variable remains at this stage unclear.

In sum, western states host many top performing banks in terms of CE. Banks in eastern states perform below average. CE scores may contain additional information that complement traditional KPI, such as the CI ratio.

5.1.3 Cost efficiency across banking groups

We analyse next to what extent CE scores differ across banking groups. Mean CE differences between different banking groups in our three models are reported in table 6.

Table 6: Mean cost efficiency per banking group

Group	Model 1	Model 2	Model 3	CI [3)]	N
Large banks [1)]	0.872	0.663	0.741	57.8	231
Regional commercial [2)]	0.826	0.759	0.772	28.7	2,117
Regional savings	0.921	0.856	0.871	66.4	6,459
Regional cooperatives	0.922	0.875	0.889	72.8	23,404
Federal average	0.915	0.862	0.877	68.5	32,211
Kruskall Wallis	476.0*	1137.9*	1008.3*	2218.6*	

[1)] Including large commercial banks, Postbank, Landbanks, Central cooperatives; [2)] Including regional commercial and branches of foreign banks; [3)] Cost-Income ratio general administrative expenses to raw result, measured in percent.
* Significant at the 1% level.

For all three models the difference in means is significant. In model 1, large banks are located far away from the frontier, while for small banks the opposite holds. Average potential savings per bank and year are € 8.9 million. Total average annual savings amount to € 25.5 billion for all banks in our sample.[42] For large banks, average CE amounts to 87.2 percent according to model 1. Consequently, the share of foregone Euro amounts in terms of potential savings accrues to € 15.4 billion. This implies that 60 percent of foregone producer surplus originates from large banks. Note that in relative terms, i.e. by means of efficiency scores alone, large banks perform only 4 percent worse than the overall industry

[42] Potential cost savings per bank are calculated according to $TOCSav_{kt} = \Sigma_k[(1-CE_{kt}) * TOC_{kt}]$.

average. But assessing the lost Euro value reveals that the costs to society due to cost inefficiencies of large banks are greatest.[43]

This result is amplified when using alternative input price proxies. In model 2, large bank CE differences relative to savings and cooperatives widens from 15 percent in model 1 to 19 and 21 percent, respectively. Compared to regional commercial banks, the difference is even inverted. In contrast to model 1, large banks suffer the most from inefficiency on the input side. Lower mean industry efficiency with an over-proportionate waste identified with big banks implies that annual average savings per bank increase almost by a factor 2.5 to € 22 million. Accordingly, the total of potential annual savings increase to € 61.6 billion. The share of potential savings for the group of large banks increases to 72 percent, equal to € 44.5 billion. Results therefore confirm earlier findings in the literature (e.g. Altunbas et al., 2001) that especially large banks suffer from inefficiencies.

Model 3 accounts for the possibility that demanded inputs differ substantially between banks, e.g. due to different quality of labour demanded, by assuming single federal input markets for this group. This model alleviates the shortcomings of large commercial banks found previously. While still identified as the worst performing banking group in Germany, CE differences relative to the federal average are reduced from 20 percent in model 2 to 13 percent in model 3. Annual average savings across all bank types are on the order of € 17.8 million. Excessive total costs incurred compared to fully efficient use of inputs sum to € 50.9 billion. While lower than identified inefficiencies in model 2, it is worth noting that the share of large banks still scores an impressive 68 percent. Therefore, the conclusion that large banks are most cost inefficient is robust across all three models. Likewise, we find that small banks exhibit superior cost management skills.[44]

In sum, the use of alternative input prices leads to lower mean CE. From a regional angle, all models yield CE that (i) is significantly different across states, (ii) is lowest in eastern states (and to a lesser extent city states) and (iii) yields additional information compared to

[43] Mean TOC for large banks is € 6,170 million but only around € 82.9 million and € 13.6 million for savings and cooperative banks, respectively. A 10 basis point improvement in CE at an average large bank thus requires approximately a 45 percent improvement in CE for an average cooperative bank.

[44] Note that some activities of large banks, e.g. custody or advisory services, may not be included sufficiently in the specified output vector. This could imply comparably high costs for large banks and hence to low CE. But if omitted outputs cause low CE for large banks, we would expect that they exhibit above average PE as profits would be comparably high without the burden of corresponding products specified in the frontier. The mean PE estimates across banking groups discussed in section 5.2.3 show that this is not the case. Given this result in addition to our approach to adjust the frontier for heterogeneity and the consistency with the literature, we are confident that our results are robust. Ideally, we are able to obtain additional data on output proxies to empirically test alternative production specifications in future research.

traditional KPI. From a bank sector angle, the evidence of all models suggests that large banks suffer most from cost inefficiency. It simply matters if a small cooperative wastes 20 percent of actual cost or if a large bank does. Note, however, that out of 231 large bank observations, only 39 refer to commercial banks. The remainder are either cooperative or saving banks' apex institutions or Postbank. Large commercial banks certainly did not excel in terms of efficiency. But the largest share of "wasteful" banking business on the input side occurred with large apex institutions of cooperative banks. We turn next to PE.

5.2 Alternative profit frontier

Alternative PE indicates the percentage of realised profits relative to profits that could have been realised. First, we provide estimation results and mean PE per model. Second, we discuss regional PE differences and, third, between banking groups.

5.2.1 Estimation results and alternative profit efficiency

Table 7 presents parameter estimates for the three efficient profit frontiers, respectively. No estimation reached the maximum number of iterations when maximising the likelihood. Parameter estimates of the deterministic kernel are by and large significantly different from zero. This implies that a stochastic profit frontier is superior to OLS. We conclude that profit inefficiencies prevail, but abstain as before from the interpretation of single coefficients. While substantially higher compared to the cost case, both σ and λ are significantly different from zero, too. We conclude that profit inefficiencies prevail and expect PE to be more dispersed than CE.

Table 7: Alternate profit frontier estimates 1993-2003

Model	PE model 1		PE model 2		PE model 3	
N	32,211		32,211		32,211	
Iterations	72		68		69	
Log likelihood	-18,705		-18,199		-18,207	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Constant	-9.374	0.000	-10.979	0.000	-12.226	0.000
lny_1	0.230	0.000	0.393	0.000	0.505	0.000
lny_2	-0.286	0.000	-0.426	0.000	-0.482	0.000
lny_3	0.790	0.000	0.465	0.000	0.466	0.000
lnw_1	0.482	0.000	1.209	0.000	1.155	0.000
lnw_2	-0.387	0.000	0.438	0.009	0.897	0.000
lnz	0.899	0.000	1.187	0.000	1.201	0.000
½ lny_1lny_1	0.015	0.000	0.008	0.000	0.008	0.000
½ lny_1lny_2	-0.104	0.000	-0.132	0.000	-0.135	0.000
½ lny_1lny_3	0.001	0.857	-0.003	0.377	-0.005	0.191
½ lny_2lny_2	0.127	0.000	0.204	0.000	0.206	0.000
½ lny_2lny_3	-0.091	0.000	-0.061	0.000	-0.063	0.000
½ lny_3lny_3	0.097	0.000	0.103	0.000	0.103	0.000
½ lnw_1lnw_1	0.046	0.000	-0.020	0.032	-0.017	0.051
½ lnw_1lnw_2	-0.283	0.000	-0.328	0.000	-0.316	0.000
½ lnw_2lnw_2	0.454	0.000	0.199	0.000	0.173	0.000
½ lnz^2	0.033	0.001	0.111	0.000	0.107	0.000
lny_1lnw_1	0.031	0.000	0.033	0.000	0.032	0.000
lny_1lnw_1	-0.031	0.000	-0.080	0.000	-0.111	0.000
lny_1lnw_1	-0.099	0.000	0.030	0.001	0.037	0.000
lny_1lnw_1	0.103	0.000	0.027	0.060	0.041	0.005
lny_1lnw_1	0.002	0.643	0.066	0.000	0.065	0.000
lny_1lnw_1	-0.107	0.000	-0.045	0.000	-0.039	0.001
lny_1lnz	0.036	0.000	0.060	0.000	0.062	0.000
lny_2lnz	-0.024	0.000	-0.104	0.000	-0.102	0.000
lny_3lnz	-0.067	0.000	-0.087	0.000	-0.086	0.000
lnw_1lnz	0.091	0.000	-0.162	0.000	-0.165	0.000
lnw_2lnz	0.012	0.225	0.097	0.000	0.085	0.000
T	-0.038	0.026	-0.029	0.231	0.014	0.574
T^2	0.008	0.000	0.012	0.000	0.011	0.000
lny_1T	-0.003	0.001	-0.007	0.000	-0.010	0.000
lny_2T	0.014	0.000	0.013	0.000	0.014	0.000
lny_3T	-0.013	0.000	-0.007	0.000	-0.007	0.000
lnw_1T	-0.027	0.000	-0.032	0.000	-0.032	0.000
lnw_2T	0.034	0.000	0.027	0.000	0.022	0.000
$lnNPD$	-1.001	0.000	-0.994	0.000	-0.992	0.000
			Heterogeneity in inefficiency			
Constant	4.459	0.136	6.177	0.010	8.030	0.001
lnTA	-0.271	0.121	-0.165	0.212	-0.272	0.042
RISK	-10.887	0.000	-11.009	0.000	-11.308	0.000
lnOBS	-0.182	0.002	-0.294	0.000	-0.259	0.000
HHI	-0.531	0.000	-0.658	0.000	-0.592	0.000
Savings banks	-60.508	0.000	-55.713	0.000	-54.667	0.000
Cooperative banks	-34.003	0.000	-28.591	0.000	-28.256	0.000
East	14.455	0.000	11.985	0.000	11.739	0.000
λ	15.853	0.000	14.555	0.000	14.453	0.000
σ	4.133	0.000	3.801	0.000	3.768	0.000

Most parameters of the specified heterogeneity variables are also significant. The direction and magnitude of these effects is stable across models. As in section 5.1, we limit ourselves to note that accounting for heterogeneity is necessary to disentangle "managerial inefficiency" from other influences. Contrary to the cost case, however, the specification of alternative input prices leaves the magnitude of both error and heterogeneity parameters largely unaffected.[45] Consequently, PE appears to be relatively insensitive to alternative input price measurement.

We turn therefore next to an assessment of efficiency differences across models, states and sectors. Table 8 illustrates the similarity of PE scores for the three models employed.

Table 8: Mean profit efficiency estimates 1993 - 2003

PE	Mean	SD	Skew	Kurt	Min	Max	N
Model 1	0.647	0.177	-0.902	3.081	0.216	0.959	32,211
Model 2	0.647	0.174	-0.926	3.183	0.216	0.959	32,211
Model 3	0.643	0.174	-0.897	3.120	0.216	0.959	32,211

Mean alternative PE is substantially lower compared to CE, around 65 percent. An average bank could have realised a third more profits relative to actually accrued profits. In relative terms, profit outweighs cost inefficiency. Regarding the distribution of PE most banks are no longer located close to the frontier. Many banks are far below best practice behaviour. Also, the distribution of PE hardly changes across models. Consequently, sub-optimal behaviour on the input side is of lesser importance for PE. Most of the estimated profit inefficiencies seem to arise from sub-optimal output choices.

Standard deviations and extreme values further indicate that the dispersion of PE is larger compared to CE. Best and worst performing firms differ by as much as 70 percent. Potentially, this larger difference is due to a higher volatility of earnings compared to costs.[46] To this end, we investigated whether poor performers in one year tended to be good performers in another. This is not the case – banks exhibiting a poor ability to realise their chances in output markets do so consistently over time. Thus, the fundamental ability of banks to realise profits is more diverse compared to CE.

[45] For example, neither σ nor λ vary much. We observe only a slight improvement of σ in models 2 and 3. Therefore, the explanatory power benefits only marginally from our alternative input price proxies.

[46] Consider as an example two banks that engage in riskier production plans. Let the two banks speculate on foreign exchange appreciation and depreciation, respectively. Depending on what "bet" materialises, this yields one very profit-efficient and one markedly inefficient bank. In contrast, on the expense side of the income statement, interest rates, wages and rents are specified in contracts and are therefore less volatile.

5.2.2 Alternative profit efficiency across regions

To examine the geographical pattern of PE, consider table 9 below. Alternative mean PE levels are more stable across specifications compared to CE. Also, all states are ranked almost identically. The distortion of PE due to measurement error seems negligible. At the same time, regional differences of both PE measures and KPI ratios are significant.[47]

Table 9: Mean alternative profit efficiency per state

State	Model 1	Model 2	Model 3	ROE [1)]	ROA [2)]	Margin [3)]	N
Baden-Wuert.	0.656	0.663	0.659	12.4	0.64	2.77	5,949
Bavaria	0.628	0.632	0.628	13.2	0.63	2.77	7,593
Berlin+	0.497	0.488	0.477	4.8	0.33	2.12	200
Bremen	0.578	0.578	0.573	9.9	0.49	2.51	149
Hamburg	0.523	0.524	0.521	13.5	1.42	2.25	295
Hesse	0.590	0.579	0.577	13.8	0.70	2.66	3,209
Lower Saxony	0.674	0.668	0.663	14.4	0.78	3.18	3,344
North Rhine-Westph.	0.674	0.667	0.664	15.2	0.70	2.97	5,166
Rhineland-Palatinate	0.671	0.671	0.667	14.8	0.71	3.04	2,022
Saarland	0.631	0.636	0.629	13.0	0.57	2.92	434
Schleswig-Holstein	0.678	0.672	0.668	14.2	0.73	3.09	1,111
Mecklenburg-W. P.+	0.662	0.664	0.661	10.4	0.41	3.12	417
Brandenburg+	0.669	0.679	0.675	8.4	0.35	3.12	417
Saxony+	0.684	0.689	0.686	11.2	0.41	3.08	557
Thuringia+	0.686	0.698	0.695	10.2	0.38	3.00	593
Saxony-Anhalt+	0.628	0.647	0.645	6.8	0.25	2.85	755
Federal average	0.647	0.647	0.643	13.3	0.65	2.87	32,211
Kruskall Wallis	912.0*	991.0*	992.3*	1236.6*	2292.9*	3263.9*	

Note: A '+' indicates eastern states.
[1)] Profit before tax over equity, measured in percent; [2)] Profit before tax over gross total assets, measured in percent; [3)] Interest income less interest expenses over gross total assets, measured in percent.

* Significant at the 1% level.

The negative "East" effect identified for the cost frontier is inverted. By and large, estimated PE of banks located in eastern states is higher compared to western competitors.[48] However, a more differentiated view is warranted. Of the five new states (excluding Berlin for now), Brandenburg, Saxony and Thuringia are ranked the top three in terms of alternative PE. The two remaining eastern states, Mecklenburg-Western Pomerania and Saxony-Anhalt, perform only mediocrely with a mean PE of around 65 percent. Thus, it may be overly simplistic to refer to “the East”.

[47] As in section 5.1, we provide here the comparison to traditional accounting-based performance indicators.
[48] This contradicts results obtained from a predecessor study based on Bankscope data. There, we concluded that further studies need to incorporate systematic differences, e.g. macroeconomic conditions. In this study, we adjust the benchmark frontier for factors causing heterogeneity. As inference changes substantially, we once more underline the necessity to account for exogenous factors in efficiency studies.

With respect to the additional information contained in PE estimates relative to KPI we find that a number of states are assigned similar rankings by all measures in table 9.[49] But as in the cost case some differences exist. For example, ROE and ROA rank eastern states differently. Thus, our previous conclusion that efficiency measures contain additional information is underlined.

Next, we discuss systematic PE differences across banking groups and link inefficiency scores to Euro amounts of foregone profits.

5.2.3 Alternative profit efficiency across banking groups

Mean PE estimates suggested that alternative input price proxies do not matter much. We argue in table 10 that it might indeed matter especially if large banks are affected.

Table 10: Mean alternative profit efficiency across banking groups

Banking Group	Model 1	Model 2	Model 3	ROE [3]	ROA [4]	Margin [5]	N
Large banks [1]	0.416	0.460	0.431	9.5	0.25	0.84	231
Regional commercial [2]	0.458	0.467	0.466	8.9	0.56	2.35	2,117
Regional savings	0.707	0.711	0.708	14.7	0.64	2.73	6,459
Regional cooperatives	0.650	0.647	0.644	13.3	0.67	2.98	23,404
Federal average	0.647	0.647	0.643	13.3	0.65	2.87	32,211
Kruskall Wallis	2176.0*	2142.0*	2194.4*	782.2*	358.6*	3190.8*	

[1] Including large commercial banks, Postbank, Landbanks, central cooperatives; [2] Including regional commercial and branches of foreign banks; [3] Profit before tax over equity, measured in percent; [4] Profit before tax over gross total assets, measured in percent; [5] Interest income less interest expenses over gross total assets, measured in percent

* Significant at the 1% level.

Table 10 illustrates that mean PE of large banks improves when alternative input prices are used in models 2 and 3 compared to the benchmark model. It is possible that especially the latter model grasps quality differences of inputs between large, nationally operating banks and small regional institutes more appropriately. Nonetheless, despite improvements of 4.4 and 1.5 percent, respectively, large banks exhibit the worst PE measures in all models. Opportunity costs of foregone profits are highest for large banks. These account for 57 percent of total annual potential savings of € 22.7 billion (i.e. € 13 billion).[50] Annual profits foregone for the average bank are € 6.8 million. In comparison, the equivalent value for large banks stood at € 623 million. As with wasted cost savings, improvements in large

[49] Examples are the states Saarland, Schleswig-Holstein and Berlin.
[50] We calculate total annual profits forgone according to the sum of optimal profits for each bank less the sum of actually realised profits, i.e. Σ_{kt} (PBT_{kt} / PE_{kt} - PBT_{kt}).

bank PE thus entail the largest gains due to (i) the highest relative inefficiency and (ii) the largest volume of potentially foregone producer surplus.[51]

In sum, the differences between relative inefficiency scores (8.5% cost versus 35.3% profit inefficiency) are put into perspective when absolute amounts of potential savings are examined. Then, the absolute savings potential in terms of foregone profits is of a similar order compared to excessive costs (€ 25.5 billion cost vs. € 22.7 billion profit savings). To obtain a complete picture of performance, we should therefore assess absolute savings, too.

A remaining matter of interest is to gauge similarities and differences of identified extreme performing banks according to the various alternative indicators discussed so far.

5.3 Comparing efficiency

First, we compare CE and PE with each other and to traditional KPI. Second, we investigate the degree to which the three efficiency models identify similar (or different) extreme performers. Third, we turn our attention to these groups' respective characteristics.

5.3.1 Efficiency and accounting based KPI

Table 11 below depicts correlation coefficients between our six efficiency measures and the various KPI displayed throughout. In line with Bauer et al. (1998), correlation coefficients between KPI and both CE and PE estimates are low and exhibit only miniscule changes across the three different SFA models.[52]

Table 11: Correlation between efficiency and KPI

	CI [1)]	**ROE [2)]**	**ROA [3)]**	**Margin [4)]**
CE Model 1	0.02	0.13	0.08	-0.14
CE Model 2	0.02	0.13	0.11	-0.02
CE Model 3	0.02	0.13	0.11	-0.03
PE Model 1	-0.02	0.21	0.15	0.35
PE Model 2	-0.02	0.22	0.15	0.33
PE Model 3	-0.03	0.22	0.16	0.33
CI		0.20	0.25	0.02
ROE			0.54	0.14
ROA				0.20

Note: All correlation coefficients significant at the 1% level.
1) General administrative expenses to raw result; 2) Profit before tax over equity; 3) Profit before tax over gross assets; 4) Interest income less interest expenses over gross total assets.

[51] Note in line with footnote 44 that large banks exhibit the lowest PE despite earning profits with products and services that may not be explicitly accounted for in the production plan, e.g. custody and advisory services. In line with the argument raised earlier, we therefore regard our CE and PE results as robust.
[52] It is also well known that CE and PE are weakly correlated as they measure different sorts of efficiency. Correlation coefficients are significant at 12.4%, 15.4% and 15.1% in models 1, 2 and 3, respectively.

In the upper panel of table 11, correlation coefficients between CE and KPI ratios are low. Potentially, this reflects that traditional KPI can suffer from influences due to accounting and reporting rules.[53] This is in line with our interpretation that CE conveys additional information compared to traditional KPI.

The middle panel displays correlation coefficients between PE measures and KPI. Magnitudes are low, especially between the CI ratio and PE. With respect to profitability measures, correlations are higher. Yet, they are far from perfect. We conclude therefore that PE measures also contain additional information.

The lower panel underpins that already within the group of traditional KPI correlation is low. For example, the correlation coefficient between the CI ratio and the net interest margin is close to zero. Also note that the former KPI is positively correlated with e.g. ROA and ROE. But intuitively we would expect a profitable bank to be cost efficient, too. Apparently, alternative accounting based KPI already contain different information.

In line with Bauer et al (1998) we conclude that multiple indicators convey different information. Not only the correlation between efficiency measures and KPI is low. Also, correlation between various traditional KPI is low. Therefore, low correlation between CE, PE and KPI is per se no reason to discard efficiency measures as an indicator. Instead, efficiency may be useful as a complementary KPI.[54] We turn next to the influence of alternative input prices on the identification of extreme performers.

5.3.2 Identification of extreme performers

Usually, we are less interested in identifying a single "winner" or "loser". Rather, we like to identify groups of top and flop performers. Hence, we investigate how alternative input price proxies affect the composition of the highest and lowest performing deciles of banks.

To this end, consider table 12. We compare on the basis of CE and PE how banks in the best and worst deciles of model 1 are re-ranked when input prices are measured differently. These two cohorts are referred to as top and flop deciles, respectively. We report the decile distribution of both cohorts according to model 2 and 3, respectively.

---

[53] For example, special depreciation rules to subsidise regions or industries may result in "extreme" KPI.

[54] It is beyond the scope of this study to qualify the importance of this additional information. However, Koetter et al. (2004) find that efficiency is a significant determinant of distressed and non-distressed mergers.

Table 12: Frequency distribution of top and flop banks across models

Decile	**Ranking according to model 2**				**Ranking according to model 3**			
	Top model 1		*Flop model 1*		*Top model 1*		*Flop model 1*	
	CE	PE	CE	PE	CE	PE	CE	PE
1 (Flop)	1.0	0.0	72.2	86.3	1.0	0.0	72.8	86.4
2	0.9	0.0	19.7	11.4	0.8	0.1	19.5	11.4
3	1.6	0.1	3.0	1.1	1.5	0.0	3.1	1.2
4	2.0	0.2	1.1	0.5	2.1	0.2	0.9	0.4
5	2.1	0.2	0.8	0.3	2.1	0.2	0.8	0.3
6	4.1	0.6	0.6	0.2	4.0	0.6	0.5	0.1
7	5.0	1.0	0.7	0.1	5.1	1.1	0.7	0.1
8	9.8	3.9	0.7	0.0	9.7	3.8	0.5	0.1
9	20.5	16.3	0.5	0.1	20.5	16.4	0.5	0.0
10 (Top)	53.1	77.8	0.7	0.1	53.1	77.6	0.7	0.1
Total N	3,218	3,218	3,223	3,223	3,218	3,218	3,223	3,223

Note: Frequencies in percent.

The left panel in table 12 compares the distribution of top and flop performers in model 1 across efficiency rank deciles of model 2. The best performing banks according to model 1 are depicted in the first two columns. Between 74 (CE) and 94 (PE) percent of banks identified as role models in the traditional model are identically identified as top performers when alternative input prices are used. This result indicates that efficiency studies using implicit prices arrive at similar conclusions for the majority of banks. However, note that a quarter of those banks formerly regarded as potential role models regarding CE are now re-ranked in decile eight or below. Most importantly, approximately 2 percent of all identified CE top performers are ranked diametrically opposed in deciles one and two. Such drastic re-classifications are absent for PE rankings. This confirms that PE is hardly affected by alternative input prices.

Banks with the worst efficiency scores under traditional SFA analysis are compared to model 2 in the third and fourth columns of the left-hand side panel. The cumulative share of identically identified flop performers in deciles one and two is even higher, ranging between 92 (CE) and 98 (PE) percent. One percent of banks switch into the best performing CE deciles nine and ten after adjusting for implicit input prices. Changed identification of top and flop performers is again limited to CE rankings.

In the right-hand panel, we conduct the same comparison for models 1 and 3. The implications are qualitatively identical. For both top and flop performers, the cumulative share of banks located in deciles nine and ten under the two models ranges between 74 and 98 percent. We can confirm that cost efficiency rankings are most sensitive to input price specification. Approximately two percent of former role model banks are re-classified under model 3 and the cumulative share for former trouble banks is again one percent.

In sum, the cumulative share of banks that are re-classified diametrically different as PE top and flop performers across models is very low. But based on CE, around 1-2 percent of former top (flop) performers are now worst (best) in class.

5.3.3 Characteristics of extreme performers

Because the use of alternative input prices affects CE especially, we limit our attention in table 13 to top and flop performers according to cost models.

Table 13: Characteristics of top and flop performers based on CE across models

		Model 1		Model 2		Model 3		Population
		Flop *	**Top***	**Flop ***	**Top***	**Flop ***	**Top***	
Group	**Large banks** [1)]	2.0	0.4	4.9	0.0	4.1	0.0	0.7
	Commercial [2)]	24.8	9.1	26.9	7.6	26.7	7.7	6.6
	Savings	20.6	16.9	24.6	12.7	24.8	13.1	20.1
	Cooperatives	52.7	73.7	43.6	79.7	44.4	79.2	72.7
State	**Baden-Wuert.**	3.5	33.5	3.7	32.6	3.7	32.9	18.5
	Bavaria	7.6	26.7	7.8	27.2	7.7	27.0	23.6
	Berlin[+]	2.6	1.0	2.5	0.8	2.3	0.8	0.6
	Bremen	0.7	0.4	0.6	0.4	0.7	0.3	0.5
	Hamburg	2.6	0.7	4.2	1.0	4.0	1.0	0.9
	Hesse	12.9	12.4	12.0	12.7	12.1	13.0	10.0
	Lower Saxony	9.1	5.3	6.3	5.2	6.1	5.1	10.4
	North Rhine-West.	5.4	12.0	6.3	13.8	6.2	13.8	16.0
	Rhineland-Palat.	2.1	4.0	1.4	3.2	1.2	3.1	6.3
	Saarland	0.1	0.7	0.2	0.9	0.1	0.8	1.4
	Schleswig-Holstein	6.0	3.1	3.9	2.3	3.9	2.0	3.5
	Mecklenburg-W. P.[+]	6.6	0.0	7.5	0.0	7.4	0.0	1.3
	Brandenburg[+]	8.6	0.0	8.4	0.0	8.6	0.0	1.3
	Saxony[+]	10.8	0.0	11.5	0.0	11.5	0.0	1.7
	Thuringia[+]	8.6	0.0	9.2	0.0	9.4	0.0	1.8
	Saxony-Anhalt[+]	12.9	0.1	14.5	0.0	15.1	0.0	2.3
Het.	**Total assets** [a)]	3,540	470	7,640	402	7,090	421	3,316
	HHI [b)]	4,839	2,677	5,012	2673	5026	2,655	1,540
	Risk [c)]	51.9	58.9	52.3	59.3	52.0	59.4	61.1
	OBS [a)]	449.0	40.0	1,230	31.5	1,150	36.1	207
KPI	**CI**	49.5	63.0	49.2	63.3	49.4	63.1	68.5
	ROE	8.3	18.6	8.0	19.3	8.2	19.3	13.3
	ROA	0.5	0.9	0.4	1.0	0.4	1.0	0.7
	Margin	3.0	2.8	2.8	2.9	2.9	2.9	2.9

Note: Based on cost efficiency ranking for models 1 through 3, respectively.
[1)] Including large commercial banks, Postbank, Landbanks, central cooperatives; [2)] Including regional commercial and branches of foreign banks; [+)] Formerly East Germany; [a)] Measured in millions of Euro; [b)] Between 1 for perfect competition and 10,000 for monopoly; [c)] Average Basel risk weight.

* Frequencies in percentages.

In the first panel, we compare relative frequencies of top and flop performers across models and banking groups. The rightmost column depicts throughout population representation. We find evidence that large banks are relatively often in the group of worst performers and never a member of the top performer group according to our alternative models. To a lesser extent, savings banks exhibit the same over- and under-representation in the top and flop performer groups, respectively. Both alternative models amplify the indication that especially cooperative banks are role models regarding CE.

In the second panel, we inspect the geographical representation. Only the states of Baden-Wuerttemberg and Bavaria simultaneously host relatively many top and relatively few flop banks, respectively. All eastern states suffer from a relative abundance of poorly performing banks and an absence of top performers. Average cost inefficiency in the East is thus higher even after accounting for less favourable economic conditions. Small city states host only marginally more good banks compared to their population representation. At the same time, they are home to too many worst performers. Potentially, a more limited geographical scope to demand inputs restricts the cost management abilities of these banks.

Characteristics in terms of heterogeneity variables reveal that top decile banks are small. Worst performers are larger than the average bank. These differences are amplified when alternative input price models are used. Both groups operate in local markets with a higher concentration than the average county in Germany. The mean HHI for top performers in models 2 and 3, though, is almost twice as high as for banks in the top decile. Consequently, some concentration is not per se an impediment to high CE. But beyond a certain level, the lack of market discipline entails that managers forego cost savings. Our measure of risk-taking, RISK, indicates that top performers choose an asset composition close to average risk. This implies that on average the successful bank did not venture into particularly risky business. In turn, too little risk-taking did not yield a successful strategy in terms of CE, either. The flop decile average shows that poorly performing banks exhibit an average risk weight around 9 percentage points below the population mean for all three models. Finally, our proxy for the degree of active risk management, the volume of OBS, indicates that top banks engaged less in OBS activities. The opposite holds for poor performers. The results regarding a bank’s approach to risk suggest that a passive strategy of choosing assets close to average risk is most beneficial for banks that aim at high CE.

In the fourth panel, we compare mean KPI of the two groups. A CI I ratio that is higher for top compared to flop performers is counterintuitive. However, it is in line with the positive correlation coefficient between CI and ROE and ROA found previously. Profitability measures ROA and ROE lead to clearer conclusions. In all three models top performers exhibit substantially lower returns compared to both top performers and the overall mean. The divergence between tops and flops further increases when alternative input prices are used. Profitability KPI and CE therefore do not necessarily contradict each other when identifying best and worst in class banks. Furthermore, alternative input prices help to carve out best and worst in class performers in a clearer fashion.

6 Conclusion

In this paper, we analyse the direction and magnitude of a measurement error inherent in most bank efficiency studies. Despite the assumption of perfect markets and, thus, exogenous input prices, most analyses employ bank-specific prices for labour, fixed assets and borrowed funds when estimating cost and profit efficiency. Based on regional factor markets, we suggest two approaches to obtain alternative input price proxies.

Our sample covers all German savings, cooperative and commercial banks between 1993 and 2003. To take the substantial degree of heterogeneity across banks into consideration, we distinguish explicitly between deviations from optimal costs and profits due to sub-optimal production and those caused by additional factors, such as bank sector, region of operation, asset risk, size and local bank market conditions. Our main findings are six-fold.

First, after accounting for heterogeneity the use of alternative input price proxies affects mean CE especially. On average, cost inefficiency is around 5 percent higher compared to traditional input prices. In turn, PE is hardly affected. Only the group of large banks yields improvements of PE on the order of 1.5 to 4.4 percent.

Second, our findings confirm results from the literature that in relative terms mean CE dominates mean PE. The former is around 92 percent while the latter is around 65 percent. However, this seemingly stark difference is put into perspective when the associated absolute Euro amounts of foregone cost savings and profits are investigated. These are fairly equal to each other at € 25.5 billion and € 22.7 billion, respectively. Therefore, the analysis of unrealised profits and forfeited cost savings yields important insights.

Third, in terms of CE small banks operating in large western states constitute top performers in German banking. Large banks and those located in eastern states are worst in class, the latter exhibiting CE 15 to 18 percent below the federal average. We find that larger size, more risky assets, active risk management and higher market concentration all lead to above frontier cost. These characteristics of top and flop performers are amplified when using alternative input prices.

Fourth, in terms of PE small banks are also among the best performing banks while their large competitors continue to be worst in class. In contrast to CE, however, those banks located in the East are on average more profit efficient than banks operating in the West in general and in city states in particular.

Fifth, we confirm previous findings in the literature that both efficiency measures are only weakly correlated with traditional KPI. We interpret this result accordingly to imply that efficiency measures contain information complementary to accounting based indicators.

Finally, the use of alternative input prices does not only affect efficiency levels but also, perhaps even more importantly, efficiency rankings. PE rankings are fairly stable across models. But around 2 percent of traditionally identified CE top and flop performers are reclassified diametrically opposed as flop and top performers, respectively, when employing alternative input prices.

In sum, the use of alternative input prices complements and strengthens previous German efficiency studies. Our results suggest that regional information is an important ingredient in SFA. Explicitly accounting for heterogeneity is key for future studies. Despite doing so we still find systematic deviations that seem to arise due to sub-optimal input use. Future extensions may therefore follow either of the following two approaches. First, to model input markets explicitly, estimate equilibrium prices directly and combine such results with efficiency research. Second, to further detail the production technology of banks as to include, for example, custody and advisory services. This would enable us to analyse if large bank inefficiency is primarily the result of fundamentally different production processes or whether it is simply more difficult to manage large organizations efficiently.

References

Aigner, D. (1974). MSE dominance of least squares with errors of observation. *Journal of Econometrics* 2, 365–372.

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

Altunbas, Y., L. Evans and P. Molyneux (2001). Bank ownership and efficiency. *Journal of Money, Credit, and Banking* 33(4), 926–54.

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

Battese, G. E. and T. J. Coelli (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* 38(3), 387–399.

Bauer, P. W., A. N. Berger, G. D. Ferrier and D. B. Humphrey (1998). Consistency conditions for regulatory analysis of financial institutions: A comparison of frontier efficiency methods. *Journal of Economics and Business* 50(2), 85–114.

Berger, A. N. (2003). The efficiency effects of a single market for financial services in Europe. *Journal of Operational Research* 150, 466–481.

Berger, A. N. and D. B. Humphrey (1997). Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research* 98(2), 175–212.

Berger, A. N. and L. J. Mester (2003). Explaining the dramatic changes in performance of US banks: Technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation* 12(1), 57–95.

Bikker, J. A. and K. Haaf (2002). Competition, concentration and their relationship: An empirical analysis of the banking industry. *Journal of Banking & Finance* 26, 2191–2214.

Bos, J. W. B., F. Heid, J. W. Kolari, M. Koetter and C. J. M. Kool (2004). The stability and sensitivity of stochastic frontier scores when controlling for heterogeneity. Mimeo.

Clark, J. A. and T. F. Siems (2002). X-efficiency in banking: Looking beyond the balance sheet. *Journal of Money, Credit, and Banking* 34(4), 987–1013.

Coelli, T., D. P. Rao and G. E. Battese (1998). *An Introduction to Efficiency Analysis*. Boston/Dordrecht/London: Kluwer Academic Publishers.

European Central Bank (2002). Structural analysis of the EU banking market.

Freixas, X. and J.-C. Rochet (1997). *Microeconomics of Banking*. Cambridge: MIT Press.

Greene, W. H. (1993). *Econometric Analysis* (4th ed.). New York: Macmillan.

Hempell, H. S. (2004). Testing for competition among German banks. *Bundesbank Discussion Paper* 04/02, 1–47.

Hughes, J. and L. J. Mester (1993). A quality and risk adjusted cost function for banks: Evidence on the ‘too-big-to-fail-doctrine’. *Journal of Productivity Analysis* 4, 292–315.

Humphrey, D. B. and L. B. Pulley (1997). Bank’s response to deregulation: Profits, technology and efficiency. *Journal of Money, Credit, and Banking* 29 (No. 1), 73–93.

Jondrow, J., C. A. K. Lovell, S. Van Materov and P. Schmidt (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19(2-3), 233–38.

Koetter, M., J. W. B. Bos, F. Heid, J. W. Kolari, C. J. M. Kool and D. Porath (2004). Determinants of German Bank Consolidation. Mimeo.

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

Kumbhakar, S. C., S. Ghosh and J. T. McGuckin (1991). A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms. *Journal of Business and Economic Statistics* 9 (3), 279–86.

Lang, G. and P. Welzel (1996). Efficiency and technical progress in banking: Empirical results for a panel of German cooperative banks. *Journal of Banking & Finance* 20(6), 1003–23.

Lang, G. and P. Welzel (1998). Mergers Among German Cooperative Banks: A Panel-Based Stochastic Frontier Analysis. *Small Business Economics* 13(4), 273–86.

Lang, G. and P. Welzel (1998). Technology and Cost Efficiency in Universal Banking a "Thick Frontier"-Analysis of the German Banking Industry. *Journal of Productivity Analysis* 10(1), 63-84.

Lozano-Vivas, A., J.-T. Pastor and J. M. Pastor (2002). An efficiency comparison of European banking systems operating under different environmental conditions. *Journal of Productivity Analysis* 18, 59–77.

Maudos, J., J. M. Pastor, F. Perez, and J. Quesada (2002). Cost and Profit Efficiency in European Banks. *Journal of International Financial Markets, Institutions and Money* 12 (1), 33–58.

Meeusen, W. and J. V. D. Broeck (1977). Efficiency estimation for Cobb-Douglas production functions with composed error. *International Economic Review* 18 (2), 435–44.

Mester, L. J. (1997). Measuring efficiency at U.S. banks: Accounting for heterogeneity matters. *Journal of Operational Research* 98(2), 230–42.

Mountain, D. C. and H. Thomas (1999). Factor price misspecification in bank cost function estimation. *Journal of International Financial Markets, Institutions and Money* 9, 163–82.

Waldmann, D. (1982). A stationary point for the frontier likelihood. *Journal of Econometrics* 18 (2), 275–79.

The following Discussion Papers have been published since 2004:

Series 1: Economic Studies

1	2004	Foreign Bank Entry into Emerging Economies: An Empirical Assessment of the Determinants and Risks Predicated on German FDI Data	Torsten Wezel
2	2004	Does Co-Financing by Multilateral Development Banks Increase "Risky" Direct Investment in Emerging Markets? – Evidence for German Banking FDI	Torsten Wezel
3	2004	Policy Instrument Choice and Non-Coordinated Monetary Policy in Interdependent Economies	Giovanni Lombardo Alan Sutherland
4	2004	Inflation Targeting Rules and Welfare in an Asymmetric Currency Area	Giovanni Lombardo
5	2004	FDI versus cross-border financial services: The globalisation of German banks	Claudia M. Buch Alexander Lipponer
6	2004	Clustering or competition? The foreign investment behaviour of German banks	Claudia M. Buch Alexander Lipponer
7	2004	PPP: a Disaggregated View	Christoph Fischer
8	2004	A rental-equivalence index for owner-occupied housing in West Germany 1985 to 1998	Claudia Kurz Johannes Hoffmann
9	2004	The Inventory Cycle of the German Economy	Thomas A. Knetsch
10	2004	Evaluating the German Inventory Cycle Using Data from the Ifo Business Survey	Thomas A. Knetsch
11	2004	Real-time data and business cycle analysis in Germany	Jörg Döpke

12	2004	Business Cycle Transmission from the US to Germany – a Structural Factor Approach	Sandra Eickmeier
13	2004	Consumption Smoothing Across States and Time: International Insurance vs. Foreign Loans	George M. von Furstenberg
14	2004	Real-Time Estimation of the Output Gap in Japan and its Usefulness for Inflation Forecasting and Policymaking	Koichiro Kamada
15	2004	Welfare Implications of the Design of a Currency Union in Case of Member Countries of Different Sizes and Output Persistence	Rainer Frey
16	2004	On the decision to go public: Evidence from privately-held firms	Ekkehart Boehmer Alexander Ljungqvist
17	2004	Who do you trust while bubbles grow and blow? A comparative analysis of the explanatory power of accounting and patent information for the market values of German firms	Fred Ramb Markus Reitzig
18	2004	The Economic Impact of Venture Capital	Astrid Romain, Bruno van Pottelsberghe
19	2004	The Determinants of Venture Capital: Additional Evidence	Astrid Romain, Bruno van Pottelsberghe
20	2004	Financial constraints for investors and the speed of adaption: Are innovators special?	Ulf von Kalckreuth
21	2004	How effective are automatic stabilisers? Theory and results for Germany and other OECD countries	Michael Scharnagl Karl-Heinz Tödter

22	2004	Asset Prices in Taylor Rules: Specification, Estimation, and Policy Implications for the ECB	Pierre L. Siklos Thomas Werner Martin T. Bohl
23	2004	Financial Liberalization and Business Cycles: The Experience of Countries in the Baltics and Central Eastern Europe	Lúcio Vinhas de Souza
24	2004	Towards a Joint Characterization of Monetary Policy and the Dynamics of the Term Structure of Interest Rates	Ralf Fendel
25	2004	How the Bundesbank really conducted monetary policy: An analysis based on real-time data	Christina Gerberding Andreas Worms Franz Seitz
26	2004	Real-time Data for Norway: Challenges for Monetary Policy	T. Bernhardsen, Ø. Eitrheim, A.S. Jore, Ø. Røisland
27	2004	Do Consumer Confidence Indexes Help Forecast Consumer Spending in Real Time?	Dean Croushore
28	2004	The use of real time information in Phillips curve relationships for the euro area	Maritta Paloviita David Mayes
29	2004	The reliability of Canadian output gap estimates	Jean-Philippe Cayen Simon van Norden
30	2004	Forecast quality and simple instrument rules - a real-time data approach	Heinz Glück Stefan P. Schleicher
31	2004	Measurement errors in GDP and forward-looking monetary policy: The Swiss case	Peter Kugler Thomas J. Jordan Carlos Lenz Marcel R. Savioz

32	2004	Estimating Equilibrium Real Interest Rates in Real Time	Todd E. Clark Sharon Kozicki
33	2004	Interest rate reaction functions for the euro area Evidence from panel data analysis	Karsten Ruth
34	2004	The Contribution of Rapid Financial Development to Asymmetric Growth of Manufacturing Industries: Common Claims vs. Evidence for Poland	George M. von Furstenberg
35	2004	Fiscal rules and monetary policy in a dynamic stochastic general equilibrium model	Jana Kremer
36	2004	Inflation and core money growth in the euro area	Manfred J.M. Neumann Claus Greiber
37	2004	Taylor rules for the euro area: the issue of real-time data	Dieter Gerdesmeier Barbara Roffia
38	2004	What do deficits tell us about debt? Empirical evidence on creative accounting with fiscal rules in the EU	Jürgen von Hagen Guntram B. Wolff
39	2004	Optimal lender of last resort policy in different financial systems	Falko Fecht Marcel Tyrell
40	2004	Expected budget deficits and interest rate swap spreads - Evidence for France, Germany and Italy	Kirsten Heppke-Falk Felix Hüfner
41	2004	Testing for business cycle asymmetries based on autoregressions with a Markov-switching intercept	Malte Knüppel
1	2005	Financial constraints and capacity adjustment in the United Kingdom – Evidence from a large panel of survey data	Ulf von Kalckreuth Emma Murphy

2	2005	Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model	Sandra Eickmeier
3	2005	Financial intermediaries, markets, and growth	F. Fecht, K. Huang, A. Martin
4	2005	The New Keynesian Phillips Curve in Europe: does it fit or does it fail?	Peter Tillmann
5	2005	Taxes and the financial structure of German inward FDI	Fred Ramb A. J. Weichenrieder
6	2005	International diversification at home and abroad	Fang Cai Francis E. Warnock

Series 2: Banking and Financial Studies

1	2004	Forecasting Credit Portfolio Risk	A. Hamerle, T. Liebig, H. Scheule
2	2004	Systematic Risk in Recovery Rates – An Empirical Analysis of US Corporate Credit Exposures	Klaus Düllmann Monika Trapp
3	2004	Does capital regulation matter for bank behaviour? Evidence for German savings banks	Frank Heid Daniel Porath Stéphanie Stolz
4	2004	German bank lending during emerging market crises: A bank level analysis	F. Heid, T. Nestmann, B. Weder di Mauro, N. von Westernhagen
5	2004	How will Basel II affect bank lending to emerging markets? An analysis based on German bank level data	T. Liebig, D. Porath, B. Weder di Mauro, M. Wedow
6	2004	Estimating probabilities of default for German savings banks and credit cooperatives	Daniel Porath
1	2005	Measurement matters – Input price proxies and bank efficiency in Germany	Michael Koetter