Gralka, Sabine

Working Paper

Persistent inefficiency in the higher education sector: Evidence from Germany

CEPIE Working Paper, No. 06/16

Provided in Cooperation with:
Technische Universität Dresden, Faculty of Business and Economics

TECHNISCHE UNIVERSITÄT DRESDEN

CEPIE Working Paper No. 06/16

Center of Public and International Economics

PERSISTENT INEFFICIENCY IN THE HIGHER EDUCATION SECTOR: EVIDENCE FROM GERMANY

September 2016

Sabine Gralka

Editors: Faculty of Business and Economics, Technische Universität Dresden.

This paper is published on the Open Access Repository Qucosa.
The complete Working Paper Series can be found at the CEPIE Homepage.

ISSN 2510-1196

Persistent Inefficiency in the Higher Education Sector: Evidence from Germany*

Sabine Gralka†
Technische Universität Dresden

September 2016

Abstract

Evaluations of the Higher Education Sector are receiving increased attention, due to the rising expenditures and the absence of efficiency enhancing market pressure. To what extent universities are able to eliminate inefficiency is a question that has only partially been answered. This paper argues that heterogeneity among universities as well as persistent inefficiency hinder the institutions to achieve full efficiency - at least in the short run. Two standard and one novel specification of the Stochastic Frontier Analysis are applied to a new, comprehensive set of panel data to show how the standard efficiency evaluation changes when both aspects are taken into account. It is the first time that the idea of persistent inefficiency is considered in the analysis of the German Higher Education Sector. The comparison reveals that the disregard of heterogeneity distorts the estimation results towards lower efficiency values. The newly introduced specification improves the accuracy of the heterogeneity assumption and exposes that inefficiency tends to be long term and persistent rather than short term and residual. This implies that increasing efficiency requires a comprehensive change of the university structure.

JEL classification: C14; C23; D61; I22; I23; H52

Keywords: Persistent Inefficiency; Stochastic Frontier Analysis; Cost Efficiency; Higher Education; Germany

---

*I thank Heike Auerswald, Alexander Kemnitz, André Seidel and Silke Übelmesser for valuable comments and intensive discussion. Moreover, I acknowledge support from the Federal Statistical Office of Germany, especially Marco Threin, for providing the data as well as helpful remarks.

†Faculty of Business and Economics, Technische Universität Dresden, 01062 Dresden, Germany. E-mail: sabine.gralka@tu-dresden.de

Introduction

As the public expenditures for the Higher Education [HE] Sector have steadily risen in industrialized countries, evaluating its efficiency receives increased attention [Baskaran and Hessami (2012)]. This is fueled by the fact that the dominant role of public fundings in many countries renders universities partially immune to the efficiency enhancing pressures of the market.

However, inefficiency can arise for different reasons and not all of them can be eliminated by the institutions themselves. The following analysis argues that heterogeneity and persistent inefficiency hinder universities to achieve full efficiency, when measured with the standard specification, at least in the short run. Heterogeneity thereby refers to permanent differences among institutions. The course of history has rendered universities heterogeneous in regard to their structures and surroundings. Hence, long term factors exist, which cannot be altered by the institutions and should therefore be ruled out from the efficiency term. While this argument recently gained popularity within the HE Sector, the distinction between residual and persistent efficiency is new to the evaluation. However the inclusion of persistent inefficiency allows a more accurate estimation because not all long term factors are fixed and can therefore be assigned to heterogeneity. The approach allows to distinguish between long term fixed factors (heterogeneity) and equally long term, but alterable persistent inefficiency. Thus, two types of efficiency are ascertainable, namely a varying short term (i.e. residual) and a stable long term (i.e. persistent). This additionally allows a more elaborate evaluation of policy implications because both components convey different types of information. While short term efficiency can be interpreted in the context of a chosen year, persistent inefficiency indicates operational problems at the institutional or state level. It is helpful in identifying whether there are groups of institutions suffering from predominantly long term (or short term) problems in spending or management strategies. Especially in Germany where education is a federal state responsibility, the varying state determined regulations could influence the efficiency of the universities lastingly to different extents. The separation of efficiency can be seen as a first step to uncover such influences. Distinguishing between influenceable short and long term efficiency, while controlling for exogenous, unchangeable factors, is thus essential to deduce appropriate policy recommendations.

The econometric opportunity to include both arguments has emerged just recently. Kumbhakar et al. (2014) were the first to separate short and long term efficiency, while controlling for heterogeneity. Hitherto, only Titus et al. (2016) have applied this approach to the HE Sector for the US, showing that cost inefficiency tends to be persistent rather than short term. However, focusing on a supplementary matter they do not perform a thorough analysis of the new specification. In addition, the employed dataset leaves room for improvement. Thus, an in-depth analysis of the new specification, including a comparison of the results to the most frequently used models, seems to be overdue. For the case of Germany, the present study shows how taking heterogeneity and persistent inefficiency into account affects the results of the standard efficiency evaluation of the HE Sector. The comparison indicates whether the new specification is advisable and policy conclusions are likely to vary by method of estimation. For this purpose, Germany provides an ideal testing ground to evaluate the efficiency of the HE Sector. Due to its distinct, unchanging and well recorded university structure, an exceptional broad and long dataset can be utilized.

The results confirm that the newly introduced specification improves the accuracy regarding the heterogeneity assumption and reveal that inefficiency tends to be long term and persistent in the German HE Sector. We additionally show that the tested models identify common sets of high and low performing institutions, but that the ranking of the remaining universities is likely to vary by method. When interpreting efficiency evaluations, policymakers should be additionally aware which method is applied and whether the specification is likely over- or underestimating efficiency.

A short literature review is given in the next section. This is followed by a look at the dataset and an exposition of the methods of analysis. A concluding section draws together the main findings and makes some suggestions for future research.

Literature Review

By now, the Stochastic Frontier Analysis [SFA] that originates from the study of Aigner et al. (1977), can be seen as a standard approach to evaluate efficiency in a variety of research areas. While the first application for the HE Sector was conducted in the nineties by Johnes (1996), focusing on the economies of scale and scope of British institutions, the method became popular for this sector only after the turn of the millennium. By now the parametric approach is applied to a broad range of countries[1] and is one of the standard methods to estimate efficiency of the HE Sector. Researchers have come to recognize the multi-product nature of HE Institutions. While initial studies were limited to cross-sectional data, the utilization of panel data sets soon became customary, starting with Flegg et al. (2004). Nonetheless the considered period typically comprises only two to four years[2]. Many reformulations of the original statistical model have emerged

[1] See for example Johnes and Johnes (2009) for the UK, Sav (2012) for the US, Zoghbi et al. (2013) for Brazil, Longlong et al. (2009) for China and Bolli et al. (2016) for a EU country comparison.

[2] See for example Johnes and Johnes (2009) and Johnes and Schwarzenberger (2011). Exceptions are the recent studies from Bolli et al. (2016), who look at a ten- or Titus et al. (2016) who look at nine-year period.

and subsequently found their way into the evaluation of the HE Sector. Among them are extensions to include heterogeneity between institutions[3]. Since universities usually evolved in a historic context, the institutions feature different locations, teaching methods, fields and extend of research as well as governance structures. To account for these structural differences, the literature usually concentrates on universities only, leaving out polytechnics as well as all specialized and private institutions. Since the heterogeneity is still severe, some authors, e.g. Johnes et al. (2005) take an additional step by estimating cost functions specific to certain pre-specified subgroups of institutions. A similar creation of sample subgroups is realized using the latent class estimation, amongst others applied by Agasisti and Johnes (2015). But, both approaches are not satisfactory due to the difficulty to define main attributes which are used for the categorization and the resulting blurry distinction between the types of institutions. The econometrical foundation to include structural differences directly into the regression was developed by Greene (2005a)[4]. Heterogeneity among institutions is thereby incorporated and measured by a university specific, time-invariant component in the estimation equation. In doing so, it is assumed that all constant influences go back to heterogeneity, which is in itself a fixed factor and cannot be altered by the institutions. The method is now widely used in the HE Context[5]. Another, adjacent argument within the econometric literature states, that part of the time-invariant component is not given and caused by heterogeneity, but can be altered by the universities. Consequently it should be included in the efficiency value. The example of management illustrates the argument, it differs between the institutions and is commonly a long term factor. But since it is adaptable, management should be part of the efficiency term. Following this argument the time-invariant component is split into heterogeneity and long term inefficiency. Two types of efficiency are ascertainable, namely a varying residual and a stable persistent term. While the residual term reflects changes, which occur in a given year, the persistent term echoes the effects of inputs like management as well as other unobserved, changeable inputs, which vary across institutions but are constant over time. Additionally they convey different types of information, which helps to derive deliberate policy implications. The first model to include this argument was presented by Kumbhakar et al. (2014) as well as Colombi et al. (2014) and allows to

[3] An extended review of these models can be found in the survey by Greene (2008).

[4] An alternative concept to include heterogeneity, relaxes the assumption that all units must face the same underlying costfunction. The "Random-Parameter" model, developed by Tsionas (2002) and Greene (2005b) allows the parameters of the function to vary across institutions, while the institution-specific parameters are constrained to be constant over time. Using panel data, it has been applied for the HE Sector of selected countries [see for example Johnes and Johnes (2009) for the UK and Johnes and Schwarzenberger (2011) for Germany]. It has to be noted that the model entails a complex econometrical implementation, especially due to the possibility of flat likelihood functions [Kumbhakar et al. (2015)].

[5] See for example Johnes and Johnes (2009) for the English, Agasisti and Johnes (2015) for the Italian and Johnes and Schwarzenberger (2011) for the German HE Sector.

separate short and long term efficiency, while controlling for heterogeneity[6]. One part of the time-invariant component, which each university exhibits is persistent inefficiency, the other is heterogeneity. This novel specification was only recently applied for the HE Sector by Titus et al. (2016), who looked at master institutions in the US and showed that cost inefficiency tends to be persistent rather than short term. The authors primarily focused on the supplementary thought of spatial interdependency and missed the opportunity to analyze the method thoroughly. To deducted recommendations regarding the further usage, it is necessary to examine, if the results are in line with the familiar models and how sensitive the specification is to the underlying variables. For the subsequent localization of efficiency potential it could be helpful to distinguish between short and long term effects and to compare groups of universities which lack (or exceed at) the same type of efficiency. The present paper follows up directly from that state of research and benefits additionally from an improved database.

The case of German Universities

Due to its distinct and well recorded university structure, Germany is an ideal object to study the efficiency of the HE Sector. The historical development entails institutions in diverse locations, which in turn could be an obvious cause for heterogeneity effects. The federal sovereignty on the other side represents a strong argument for potential persistent inefficiency induced for example by state determined regulations. The literature on German universities is surprisingly sparse. Kempkes and Pohl (2010), utilizing panel data and applying the SFA, showed that the German universities work at high level of efficiency. Subsequent studies confirmed this result and extended the analysis by taking heterogeneity into account. The finds of Johnes and Schwarzenberger (2011) suggest that heterogeneity in fixed costs and in research surroundings accounts for interinstitutional differences in cost structures. The most recent analysis by Olivares and Wetzel (2014) particularly focused on the economies of scale and scope of the institutions. Like the other studies mentioned, the authors do not separate efficiency into a short and long term.

Data

The panel data set for the present study covers the years from 2001 to 2013 and represents 73 of the 76 German public universities[7], providing a comprehensive view of the German HE Landscape. To our knowledge this is the most recent and longest time

[6] Kumbhakar and Heshmati (1995) proposed the first estimation specification which includes the idea of persistent inefficiency. They assume that the time-invariant component is entirely due to long term inefficiency, therefore neglecting the idea of heterogeneity.

[7] Due to substantial merger within the period the universities “U Duisburg-Essen”, “Brand. TU Cottbus-Senftenberg”, and “HafenCity U Hamburg” are omitted from the sample. It has to be noted that almost all universities have undergone smaller restructurings, which are not explicitly commented.

frame utilized in efficiency evaluations. Information relating to student numbers cover the academic years 2001/2002 through 2013/14 and statements about financial variables cover the business years 2001 until 2013. Institutions specializing in some fields only, like fine arts and medicine, are dropped from the sample. Universities of applied sciences and distance universities are also excluded, as they are more oriented towards teaching instead of research. The data were provided by the Federal Statistical Office of Germany. All monetary variables are deflated to the year 2013. All medicine related factors are excluded from the sample. The inclusion of their (inflated) costs could lead to a severely bias of the efficiency results as they are part of the general health provision.

In order to asses both familiar and novel methods we follow the literature closely and choose the most frequently used setting[8]. We consider teaching and research as the primary activities [Abbott and Doucouliagos (2009)]. These two outputs are evaluated with respect to the main input, the expenses of the institution. The first output variable teaching is represented by the total number of students[9] from bachelor and master courses (or equivalent)[10], differentiated across the two subject groups science and non-science subjects[11]. The research output is measured by third-party funding ("Drittmittel"), divided along the same line. Novel to the estimation of HE Efficiency, this separation allows to control for possible differences in the research output between the subject groups. The approximation of research through third party funding is common in the literature. In fact, the amount of acquired third-party funding is one of the most important performance measures used by the German states. Alternatives like publications or citations are only rarely included in resource allocation mechanisms, since e.g. publication-based measures are highly retrospective [Broemel et al. (2010)]. Moreover, one could argue that the funding provides a quality adjusted measure, since it reflects the market value of research [Johnes (1997), Worthington (2001)]. The dependent variable is the sum of annual personnel and other current expenditures of institutions, deducted by research grants. Wages, approximated by the total personnel expenditures divided by the number of occupied full-time equivalents, are included as an input-price [Stevens (2005)]. Through the wage level, differences in the structure of staff across universities can be captured. While some universities might prefer to employ a higher density of expensive research personnel and

[8] A comprehensive view of possible in- and output factors, with special regard to the German HE Sector, can be found in Warning (2005).

[9] Alternative to the amount of students the number of graduates can be used in the estimation. We follow the argumentation from Olivares and Wetzel (2014) and reason that students are the cost drivers and increase their human capital already before completing their degree.

[10] Ph.D. Students are not implemented as an output variable to avoid bias from double counting. Within the German HE Sector the majority of Ph.D. students work as research associates and are hence considered in the wage rate.

[11] General science contains mathematics, natural sciences, veterinary medicine, agricultural, forest and nutritional sciences and engineering. Non-Science subjects are courses related to art, economics, law, sport and culture.

a smaller number of less costly technical staff, other universities might have a more technical staff and less academic researchers. Since Kempkes and Pohl (2010) illustrated that there are significant differences between East and West German universities, a dummy for East German universities is additionally implemented. Costs as well as third-party funds and the number of students are normalized by the number of graduates, following Kempkes and Pohl (2010)[12].

Employing similar variables as Titus et al. (2016), the present dataset features a wider scope and contains more detailed information. Whereas Titus et al. (2016) look at selected master institutions within the US, the study at hand covers almost all German universities. This is possible through the clear distinction, permanence and well documentation of universities in Germany. The records moreover allow to exclude explicitly all medicine related factors. Titus et al. (2016) merely control for the existence of a medical degree program or an affiliated hospital. A general improvement is realized through the allocation of research funding to the two subject groups. The separation allows to control for different research structures between categories.

Descriptive statistics are reported in Table 1. The values are similar to Kempkes and Pohl (2010) and Johnes and Schwarzenberger (2011), looking at Germany, as well as Bolli et al. (2016), considering selected European countries. The student numbers amount to around 6,200 on average in sciences and 10,000 in non-science areas. The number of graduates are accordingly smaller for science (around 800) than for non-science (around 1,300). Research income, the second output variable, amounts to 45 million euro on average. All three output figures are lower for East than for West Germany. Current expenditures amount to annual 130 million and are also higher for the West German states. The comparison of the average wage rate shows that some part of the cost variance could be due to different average salaries, which in turn could be caused by different personnel structures. While West German universities employ substantially more technical than scientific employees, numbers are almost uniform in East Germany. A quite prominent characteristic of the descriptive statistics, which is in line with the literature, is that for each variable, the standard deviation is close to the mean. This indicates a considerable degree of heterogeneity among institutions.

[12] Amongst other things, this allows to include the fact that HE Institutions experience strongly varying non-completion rates [Johnes (2014)].

Table 1 - Descriptive statistics

2001-2013	**Germany** (n=962)		**East Germany** (n=182)		**West Germany** (n=780)	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Students, Science	6243.62	5041.12	4768.54	4098.93	6593.64	5180.85
Students, Non-Science	10274.70	7610.05	7693.78	4893.63	10887.12	8003.50
Graduates, Science	840.50	742.11	622.71	543.51	892.17	773.19
Graduates, Non-Science	1330.37	1011.37	977.30	779.63	1414.15	1041.93
Third-party funding, Science[a]	31.60	34.52	24.98	29.22	33.17	35.49
Third-party funding, Non-Science[a]	14.29	12.73	9.84	7.41	15.35	13.49
Costs [a]	130.28	78.94	92.39	51.10	139.26	81.69
Wages [b]	75.39	14.73	71.80	11.42	76.24	15.30
Scientific Employees	912.05	543.87	707.64	372.48	960.56	566.57
Technical Employees	998.47	632.49	708.90	347.28	1067.18	664.79

Source: Federal Statistical Office of Germany; own calculations.
[a] In € million, 2013 prices, [b] In 1000 €, 2013 prices.

Methodology Review

Since it is predictable that different specifications will give different results, our objective is not to investigate all existing models. Instead, we evaluate the classic model by Battese and Coelli (1992) and compare it to the approaches by Greene (2005a) and Kumbhakar et al. (2014)[13]. Table 2 briefly summarizes the characteristics of these three models. Because of the use of the time-invariant dummy for East Germany, a random-effects model is employed. Before discussing the empirical specifications, a short debate concerning the underlying function is necessary. Within the HE Literature a cost function is customarily used to estimate efficiency [Eagan and Titus (2016)][14]. Derived from

Table 2 - Some main characteristics of the models investigated

	Battese and Coelli	**Greene**	**Kumbhakar et al.**
Time varying efficiency	Yes	Yes	Yes
Heterogeneity	*No*	Yes	Yes
Persistent inefficiency	*No*	*No*	Yes

Source: Own representation.

[13] Not all applied models are available within on statistical package, hence model by Greene (2005a) is carried out in LIMDEP while the others are executed in STATA.

[14] Recently the (multi-) input/output distance functions are gaining popularity within the efficiency estimation [see for example Abbott and Doucouliagos (2009), Bolli et al. (2016)]. In the present study we follow the argument from Kempkes and Pohl (2010), who choose to estimate a cost function compared to a distance function, as the interpretation of the coefficients are more intuitive in the context of a cost function.

microeconomic cost theory, the cost function is the mathematical representation of the relationship between the total costs of producing a given level of outputs from a specific set of inputs. In other words, a cost function is a boundary describing the lowest cost at which an institution can produce a set of outputs[15]. Since a sufficient dataset is at hand a scaled translog function is assumed for the present analysis, which allows the usage of variables with zero values[16]. A further benefit of the specification is the valuable information offered by the included cross-terms [Coelli et al. (2005)]. This choice is in line with a variety of studies, including the earliest and most recent analysis of university costs by Koshal and Koshal (1999), Stevens (2005) and Bolli et al. (2016). The efficiency distribution is assumed to be half-normal[17]. Orientating at Christensen and Greene (1976) and Kumbhakar (1997) the translog cost function has the following form:

$$ln\, C_{it} = \alpha_0 + f(Q_{it}, w_{it}, EAST_{it}) + v_{it} + u_{it} \qquad \text{with} \qquad (1)$$

$$f(Q_{it}, w_{it}, EAST_{it}) = \sum_{j=1}^{4} \beta_j \, ln\, Q_{jit} + \frac{1}{2}\sum_{j=1}^{4}\sum_{k=1}^{4} \beta_{jk}(ln\, Q_{jit}\, ln\, Q_{kit}) + \kappa_1\, ln\, w_{it} + \kappa_2 \frac{1}{2}(ln\, w_{it})^2 + \sum_{j=1}^{4} \kappa_3 \big(ln\, w_{it}\, ln\, Q_{jit}\big) + \varphi_1 EAST_{it} \qquad (2)$$

where i denotes universities and t the time period, covering the years 2001 to 2013. C_{it} represents the normalized costs of university i in time period t. The function $f(Q_{it}, w_{it}, EAST_{it})$ describes the output technology. Here Q_{jit} represents the normalized science and non-science graduates as well as the third-party fundings per year and university. The annual average wage rate is denoted by w_{it} and EAST indicates all East German universities. The term α_0 captures the constant and β, κ, φ are unknown parameter vectors to be estimated. The composed error term ε_{it} consists of v_{it} and u_{it}. The former accounts for statistical noise and follows a normal distribution. The latter

[15] To verify the assumption for the present dataset, a skewness test on the OLS Residuals was conducted and found to be significant providing support for the cost frontier specification of the model.

[16] The literature emphasized the difficulty of choosing a cost function and highlighted three that make sense in the general multiproduct context. Baumol et al. (1982) where the first to determine the requirements and propose the constant elasticity of substitution, the quadratic and the hybrid translog specification. The first of these is known to present some conceptual difficulties [Johnes (2004), Titus et al. (2016)]. The second, which is used by most studies implementing the aspect of heterogeneity, has the disadvantage of depending on numerous assumptions. The last is demanding both in terms of data and its highly non-linear specification, but has the advantage of having a sufficiently flexible form.

[17] The efficiency could in principle follow any non-normal distribution, so that it can be separated out from the other residual term, but a common assumption is that it follows a half-normal distribution. For a comparison of the most frequently used distributions and their impact see Eagan and Titus (2016).

represents the non-negative random error term, which is independently distributed from the v_{it} term and captures efficiency.

The SFA is commonly based on two sequential steps. First, the estimates of the model parameters are obtained by maximization of a log-likelihood function. Second, university specific efficiency values are calculated. This is necessary because the first estimation allows the computation of the error term ε_{it}, but not of the individual efficiency values. The most well-known strategy for disentangling this unobserved component, proposed by Battese and Coelli (1988) [BC], exploits the conditional distribution of u_{it} given ε_{it}. The resulting BC term denotes efficiency. Ranging between 0 and 1, a higher value indicates higher efficiency.

After this short introduction into the underlying function, the three empirical specifications are discussed in more detail. Model 1 thereby represents the standard efficiency evaluation which is used in the literature most frequently. Each of the following models introduces an additional aspect, as illustrated in Table 2. While Model 2 takes account of heterogeneity among institutions, Model 3 controls for heterogeneity and allows to separate short and long term efficiency.

Model 1 - Time-Variant Efficiency

Extending the panel data approach from Pitt and Lee (1981) to accommodate the notion of efficiency improvement, Battese and Coelli (1992) proposed a time-varying SFA using Maximum Likelihood [ML] estimation[18]. In this model, efficiency is not fixed, instead it changes over time and also across institutions. The model is specified as:

$$ln\, C_{it} = \alpha_0 + f(Q_{it}, w_{it}, EAST_{it}) + v_{it} + u_{it} \quad \text{with} \tag{3}$$

$$u_{it} = u_i G(t) \tag{3.1}$$

$$G(t) = exp[\gamma(t - T)] \tag{3.2}$$

The term v_{it} denotes the normally distributed noise term and u_i captures efficiency differences across observations. Inefficiency (u_{it}) is composed of two distinct components, one is a stochastic individual component u_i, which is constant over time. The other is a non-stochastic time varying component $G(t)$, common for all institutions. Heterogeneity and persistency are not considered within the model. Therefore it can be seen as a lower boundary of efficiency, where all time-invariant effects are categorized as inefficiency.

[18] A similar model was proposed by Kumbhakar (1990). The specifications differ only in the specific form of the time-varying component, where the mentioned model has one more parameter. The estimation of both models showed that the extra parameter is not warranted and hence the Battese and Coelli (1992) model is selected (see Appendix A). This is confirmed by a Likelihood Ratio Test, which displays that the specification from Battese and Coelli (1992) is preferred to the Kumbhakar (1990) design.

Model 2 - Heterogeneity

Since the HE Sector usually evolved out of an historic context, the institutions are heterogeneous concerning their locations and structures. When evaluating the efficiency, one should account for these unchangeable institution specific effects. The econometric basis to include heterogeneity directly into the regression was developed by Greene (2005a):

$$ln\, C_{it} = \alpha_0 + f(Q_{it}, w_{it}, EAST_{it}) + \rho_i + v_{it} + u_{it} \tag{4}$$

The additional time-invariant component ρ_i is individual specific and covers heterogeneity. Compared to previous specifications, this model allows to disentangle time-varying efficiency from institution specific time-invariant, unobservable heterogeneity. For this reason, the specification is known as the "True Random Effects" [TRE] model. Although the model may appear to be the most appropriate, it can be argued that part of the time-invariant unobserved heterogeneity does belong to efficiency and should be considered as such. The example of management illustrates the argument, it differs between the institutions and is commonly a long term factor. Therefore it would most likely be classified as heterogeneity in Model 2. But, as it is in fact adaptable in the long run, it should be included in the efficiency term. The model fails to distinguish between fixed and adjustable long term factors and therefore neglects persistent inefficiency. In a way, Model 1 and Model 2 constitute opposites. In the first model, all time-invariant effects are considered as inefficiency, whereas they are ruled out from the inefficiency component in the second model. While Model 1 is likely to produce a downward bias in efficiency, because institution specific effects are treated as inefficiency, Model 2 is likely to produce an upward bias, since the persistent inefficiency is compounded in heterogeneity. As Greene (2005a) points out, neither formulation is fully satisfactory.

Model 3 - Heterogeneity and Persistent Efficiency

To avoid the shortcomings of the positions above it is necessary to distinguish between influenceable short and long term efficiency, while controlling for exogenous, unchangeable factors. Kumbhakar et al. (2014) developed a model that allows to separate heterogeneity as well as residual and persistent efficiency. The model is specified as:

$$\ln C_{it} = \alpha_0 + f(Q_{it}, w_{it}, EAST_{it}) + \rho_i + v_{it} + u_{it} \quad \text{with} \tag{5}$$

$$u_{it} = \mu_i + \tau_{it} \tag{5.1}$$

The term ρ_i is again a random institution effect that captures unobserved time-invariant factors (heterogeneity). The overall inefficiency term u_{it} is divided into the persistent (long term, constant) part μ_i and the residual (short term, changing) component τ_{it}. The inclusion of two time-invariant factors allows to separate between institution specific heterogeneity (ρ_i) and persistent inefficiency (μ_i). Hence the model has four components, two of which are institution effects and random noise (ρ_i and v_{it}) and the other two (μ_i and τ_{it}) are inefficiency. Thus, Model 3 falls in between the aforesaid boundaries, as heterogeneity and persistency are included.

A multistep procedure is used to estimate efficiency and the model in (5) is rewritten to:

$$\ln C_{it} = \alpha_0^* + f(Q_{it}, w_{it}, EAST_{it}) + \alpha_i + \varepsilon_{it} \quad \text{with} \tag{6}$$

$$\alpha_0^* = \alpha_0 - E(\mu_i) - E(\tau_{it}) \tag{6.1}$$

$$\alpha_i = \rho_i + \mu_i + E(\mu_i) \tag{6.2}$$

$$\varepsilon_{it} = v_{it} + \tau_{it} + E(\tau_{it}) \tag{6.3}$$

The model can be estimated in three steps [see Kumbhakar et al. (2014)]. In step 1 the standard random effect panel regression is used to estimate the coefficients β as well as the predicted values α_i and ε_{it}. In step 2, the prediction of ε_{it} is exploited to estimate the time-varying efficiency τ_{it} using the standard SFA. In step 3, following a similar procedure, α_i is used to obtain estimates of the persistent efficiency. Lastly, the overall efficiency u_{it} is acquired from the product of residual and persistent efficiency. While the strategy is complex and greatly dependent on the underlying distributional assumptions, its advantages lie in the improved accuracy regarding the time-invariant component and the additional information that can be gained.

Results

The comparison of the three models demonstrates how the results of the standard efficiency evaluation of the HE Sector change when taking heterogeneity and persistent inefficiency into account. The estimated cost equations are reported in Appendix B. In all three cases they have been calculated using a SFA in which efficiency is modelled as a half-normal residual. The coefficients of the outputs and inputs behave well in the sense that the values and the significance levels stay alike throughout all three specifications. A further interpretation of the results in the table is not advisable, owing to the presence of quadratic and interaction terms[19]. Table 3 presents the mean efficiency values for the estimated models[20]. The results of the standard Model 1, without heterogeneity and

[19] The implications of the cost function for economies of scale and scope in university production are not the main thrust of this study and are therefore not considered in any depth.

[20] The mean efficiency values for each university can be found in Appendix C.

persistent effects, show that universities operate moderately on the upper level of efficiency with estimates around 0.593[21]. This value is slightly lower than commonly observed within the literature of higher education using this approach. However, Johnes (2014) casts some doubt, whether estimates of efficiency, varying between 85 percent and 95 percent offer an accurate reflection of the current efficiency. Figure 1 depicts the mean efficiency score over time. The results of Model 1 suggest that universities improved their efficiency performance considerably between 2001 and 2013, which confirms the findings of Kempkes and Pohl (2010). The comparison of the mean efficiency value of Model 1 and Model 2 yields the expected upward shift of the efficiency value. According to Model 2, universities operate distinctly on the upper level of efficiency. The difference is caused by the allowance of heterogeneity among institutions in Model 2. All time-invariant factors are seen as aspects, which are not alterable by the universities. The higher value suggests that there is considerable heterogeneity among institutions. A general high mean value is in line with the literature using the TRE model to estimate efficiency [Olivares and Wetzel (2014)]. Figure 1 shows an increase of the efficiency value for Model 2 over time. However, this increase is smaller than for Model 1. This can be regarded as a first indication that long term factors causing inefficiency or heterogeneity have diminished.

With an estimated overall efficiency of 0.730, the mean efficiency of Model 3 lies in between the other two values, as anticipated. The residual efficiency corresponds to the efficiency of Model 2, both in absolute terms and over time. This is to be expected since both values are cleared of all long term factors (institution specific effects and persistent inefficiency). Table 3 also illustrates that persistent efficiency is lower than residual efficiency. Hence, inefficiency is presumably not caused by something unexpected within each year, but rather by persistent factors, as management decisions or state regulations. This outcome confirms the above mentioned results by Titus et al. (2016). Figure 1

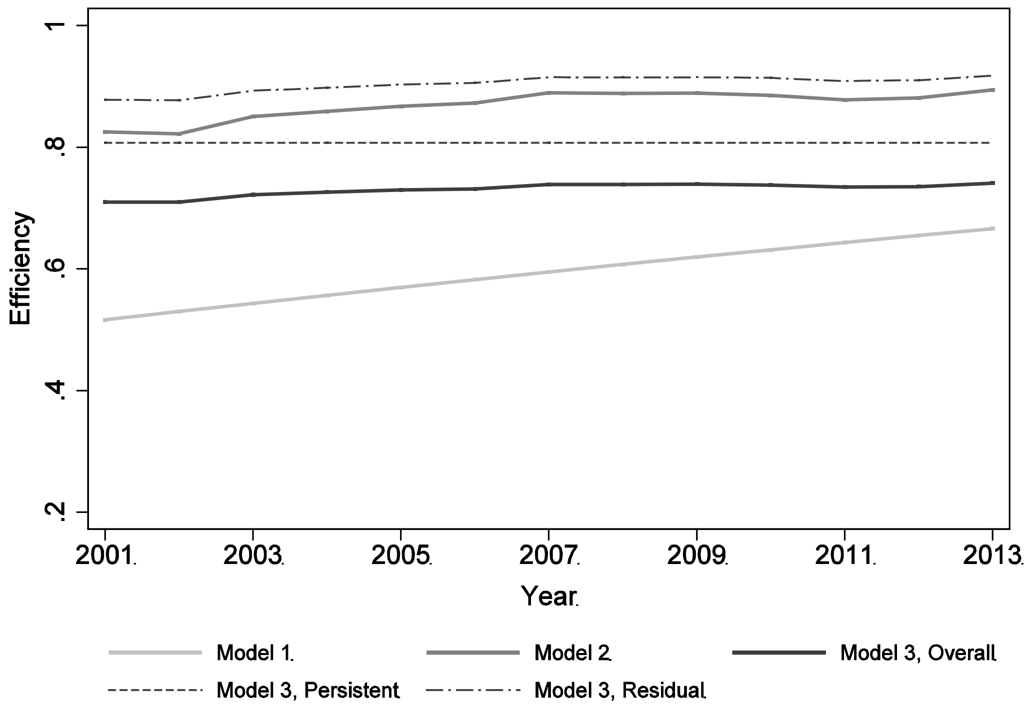
Table 3 - Efficiency values

	Mean efficiency	**Std. Dev.**	**Minimum**	**Maximum**
Model 1	0.593	0.152	0.159	0.987
Model 2	0.869	0.070	0.475	0.974
Model 3, Overall	0.730	0.104	0.256	0.913
Model 3, *Persistent*	0.807	0.106	0.395	0.955
Model 3, *Residual*	0.904	0.040	0.648	0.972

Source: Own calculations.

[21] The value implies that universities could decrease their cost by around 66 % ($\frac{1}{0.6}$ - 1) without reducing their output.

Figure 1 - Efficiency Score over time



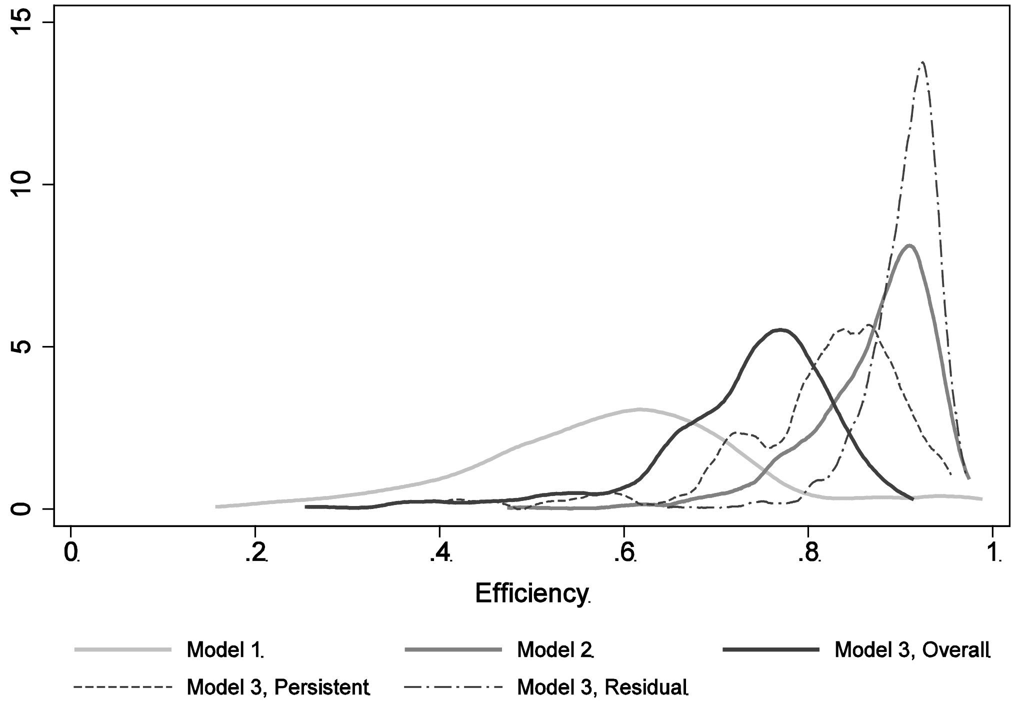
Source: Own calculations.

displays only a slight increase of the efficiency value of Model 3 over time. This increase is again smaller than for Model 1, indicating a diminishing heterogeneity among universities.

What can policymakers derive from the results so far? Evidently, both heterogeneity and persistent effects of the institutions should be incorporated in the efficiency evaluation. If the standard model is used for the estimation, a downward bias in efficiency is observable and universities cannot attain full efficiency according to the approach - at least not in the short run. While it is crucial to consider heterogeneity, only the further distinction of fixed and adjustable long term factors delivers accurate results. The calculations also suggest that increasing efficiency requires a comprehensive change in the structural framework, that is a change in policy.

Subsequent to the short analysis of the absolute values and development of efficiency over time, a more thorough assessment of the results of all models is necessary to see if the comparison is appropriate. Therefore Figure 2 gives the kernel distribution of the estimated efficiency values for all three models. The picture confirms the findings that the estimated values of Model 3 are in between Model 1 and Model 2 and that the efficiency of Model 2 and the residual efficiency of Model 3 are similar. Across all three models, distributions are comparably well shaped. Another important issue concerns the ranking

Figure 2 - Kernel Density



Source: Own calculations.

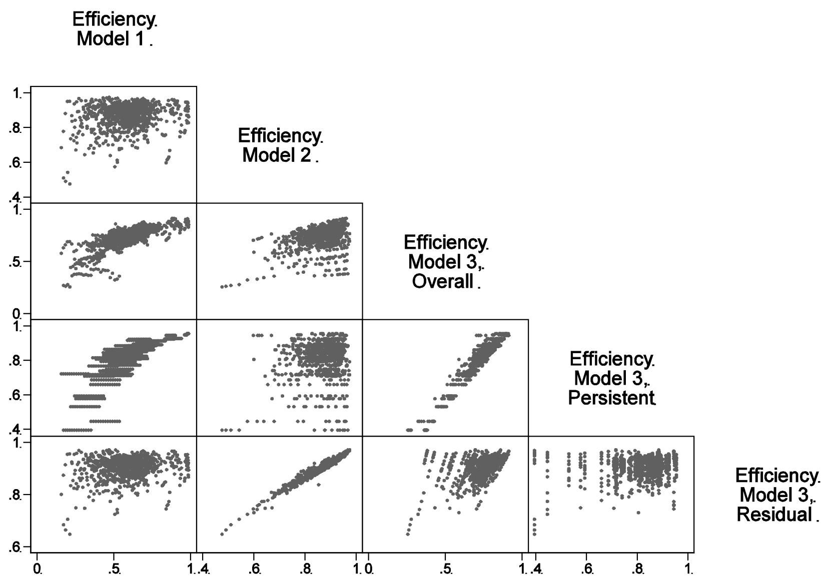
of institutions. For policy purposes it is relevant to know, if and how the university specific efficiency varies by estimation method. Table 4 shows that the rank correlations between the methods are significantly positive but varying (between 0.14 and 0.79). Merely the persistent efficiency of Model 3 exhibits no significant correlation to the residual efficiency and the efficiency values of Model 2. The high correlation between overall and persistent efficiency in Model 3 demonstrates the strong contribution of the persistent term for the overall result. The term is also responsible for the high correlation of Model 1 and Model 3. To examine the correlations further, Figure 3 shows the plotted efficiency values of each model within a matrix. While there are obvious disparities between Model 1 and Model 2, which confirm the argument of opposites, Model 3 shows similarities to both models. Particularly interesting is the illustration of the efficiency of the first two models with the separated values of Model 3. While Model 2 is highly correlated to the residual component, Model 1 is more compliant to the persistent term. University rankings should therefore be handled with caution, keeping in mind the specifications of the applied model. While they all identify common sets of high and low performing institutions, the ranking of the remaining universities should be carefully looked at.

Table 4 - Spearman Rank Order Correlation

	Model 1	Model 2	Model 3, Overall	Model 3, *Persistent*	Model 3, *Residual*
Model 1	1.000				
Model 2	0.137*	1.000			
Model 3, Overall	0.790*	0.370*	1.000		
Model 3, *Persistent*	0.818*	0.049	0.916*	1.000	
Model 3, *Residual*	0.146*	0.978*	0.383*	0.061	1.000

Source: Own calculations.

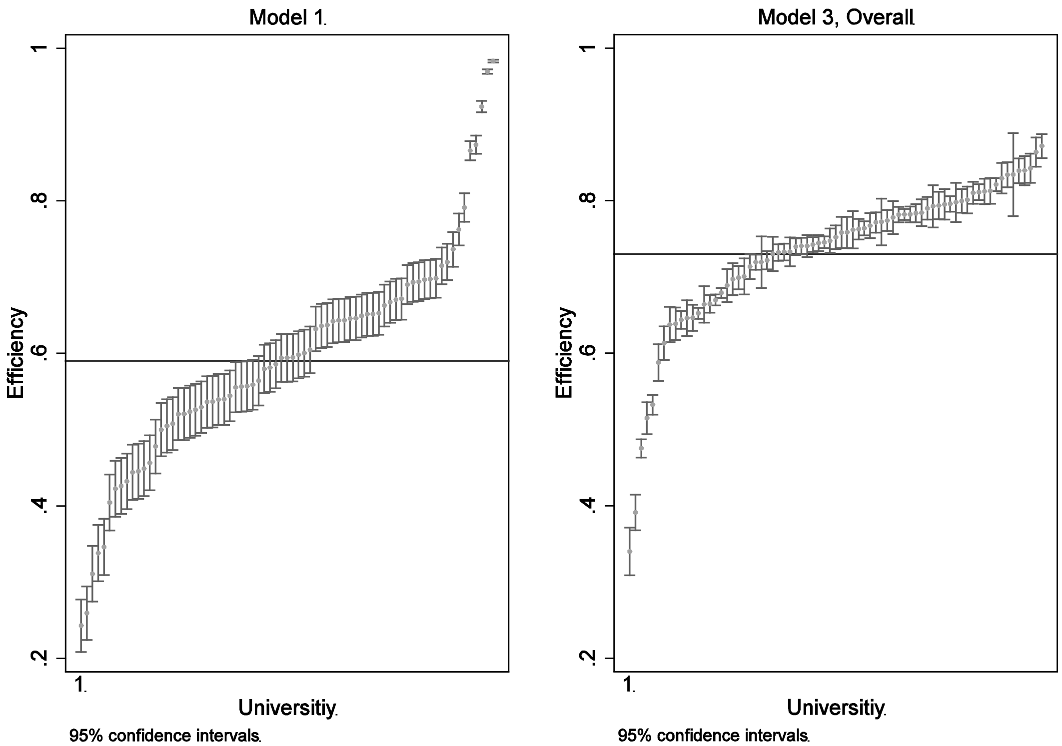
Figure 3 - Efficiency Distribution Matrix



Source: Own calculations.

Last but not least, it is important to check if the efficiency results are definite. Especially Model 3 is not extensively tested yet and should therefore be examined thoroughly. The plot of the confidence interval of standard Model 1 and novel Model 3 in Figure 4 shows that both methods can clearly discriminate between the highest and lowest performing universities. The intervals for each university are even smaller for Model 3 than for the new standard model. Therefore there are no objections in this context regarding the usage of the model. The closer study of the estimations showed that all models are generally suitable for the estimation, but deliver slightly deviating results. Especially the efficiency values of universities which are neither particularly good nor bad must be interpreted carefully. The differences should be kept in mind, when utilizing the results for further policy implications.

Figure 4 - Efficiency Score and associated 95 percent Confidence Interval by University



Source: Own calculations.

Conclusion

The present study demonstrated how the results of the standard efficiency evaluation of the HE Sector change when heterogeneity and persistent inefficiency are taken into account for the case of Germany. We show that the standard specification by Battese and Coelli (1992) produces low efficiency values because institution specific effects are treated as inefficiency. Controlling for heterogeneity, applying the model by Greene (2005a), improves the accuracy but displays high efficiency values, since persistent inefficiency is compounded in heterogeneity. Distinguishing between short and long term efficiency while controlling for heterogeneity, using the analysis by Kumbhakar et al. (2014), reveals values which are in between the afore mentioned boundaries. We expose that persistent efficiency is lower than residual efficiency. This indicates operational problems at the institutional or state level. An increase in the efficiency level could therefore only be generated through a comprehensive change in policy. Applying the novel specification is advisable. While it is crucial to consider heterogeneity, only the further distinction of fixed and adjustable long term factors delivers accurate results. The separation also permits more detailed policy implications. While short term efficiency can be interpreted in the context of a chosen year, long term inefficiency indicates operational problems at the institutional or state level. The separation is additionally helpful to identify groups of institutions which are suffering from predominant long term (or short term) problems in spending or management strategies.

The comparison of the three models shows that the university specific conclusions are likely to vary by method, since the correlation of the university ranking is significantly positive but partially small. At best, the specifications identify common sets of high and low performing institutions, but the order of the remaining universities should be carefully looked at. When interpreting efficiency evaluations policymakers should be aware which specification is applied and if the model is likely over- or underestimating efficiency.

Some further advancements are conceivable. An additional enhancement could be achieved by including a vector with time-invariant covariates z_i. This would allow to examine the marginal effect of these variables on inefficiency, which in turn gives the opportunity to deduce deliberate policy implications. Varying the timeframe and the variables could show how sensitive the results are to the underlying cost function.

We showed that the specification from Kumbhakar et al. (2014) improves the accuracy of the estimation and delivers valuable additional information. While the estimation results of all three models show similarities, the university specific efficiency values are likely to vary by method. The novel result that the overall inefficiency of the universities does only slightly change in the context of a chosen year and is mostly persistent, can be used as a starting point to new research questions. In a subsequent step one can examine whether the persistent inefficiency is caused by the management of the university itself or - especially in the German case - by long term regulations of the respective state.

References

Abbott, M. and C. Doucouliagos (2009): “Competition and Efficiency: Overseas Students and Technical Efficiency in Australian and New Zealand Universities.” *Education Economics* 17 (1): 31–57.

Agasisti, T. and G. Johnes (2015): “Efficiency, Costs, Rankings and Heterogeneity: The Case of US Higher Education.” *Studies in Higher Education* 40 (1): 60–82.

Aigner, D.; C.K. Lovell and P. Schmidt (1977): “Formulation and Estimation of Stochastic Frontier Production Function Models.” *Journal of Econometrics* 6 (1): 21–37.

Baskaran, T. and Z. Hessami (2012): “Public Education Spending in a Globalized World.” *International Tax and Public Finance* 19 (5): 677–707.

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.

Battese, G.E. and T.J. Coelli (1992): *Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India*. Springer.

Baumol, W.J.; J.C. Panzar; R.D. Willig; E.E. Bailey; D. Fischer and D. Fischer (1982): *Contestable Markets and the Theory of Industry Structure*. Harcourt Brace Jovanovich, Inc.

Bolli, T.; M. Olivares; A. Bonaccorsi; C. Daraio; A.G. Aracil and B. Lepori (2016): “The Differential Effects of Competitive Funding on the Production Frontier and the Efficiency of Universities.” *Economics of Education Review* 52: 91–104.

Broemel, R.; A. Pilniok; S. Sieweke and H.-H. Trute (2010): “Disciplinary Differences from a Legal Perspective.” In *Governance and Performance in the German Public Research Sector*, 19–41. Springer.

Christensen, L.R. and W.H. Greene (1976): “Economies of Scale in US Electric Power Generation.” *The Journal of Political Economy*, 655–676.

Coelli, T.J.; D.S.P. Rao; C.J. O’Donnell and G.E. Battese (2005): *An Introduction to Efficiency and Productivity Analysis*. Springer Science & Business Media.

Colombi, R.; S.C. Kumbhakar; G. Martini and G. Vittadini (2014): “Closed-Skew Normality in Stochastic Frontiers with Individual Effects and Long/short-Run Efficiency.” *Journal of Productivity Analysis* 42 (2): 123–136.

Eagan, K. and M.A. Titus (2016): “Examining Production Efficiency in Higher Education: The Utility of Stochastic Frontier Analysis.” In *Higher Education: Handbook of Theory and Research*, 441–512. Springer.

Flegg, A.T.; D.O. Allen; K. Field and T.W. Thurlow (2004): “Measuring the Efficiency of British Universities: A Multi-Period Data Envelopment Analysis.” *Education Economics* 12 (3): 231–249.

Greene, W.H. (2005a): “Fixed and Random Effects in Stochastic Frontier Models.” *Journal of Productivity Analysis* 23 (1): 7–32.

Greene, W.H. (2005b): "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model." *Journal of Econometrics* 126 (2): 269–303.

Greene, W.H. (2008): "The Econometric Approach to Efficiency Analysis." *The Measurement of Productive Efficiency and Productivity Growth*, 92–250.

Johnes, G. (1996): "Multi-Product Cost Functions and the Funding of Tuition in UK Universities." *Applied Economics Letters* 3 (9): 557–561.

Johnes, G. (1997): "Costs and Industrial Structure in Contemporary British Higher Education." *The Economic Journal* 107 (442): 727–737.

Johnes, G. (2004): "A Fourth Desideratum: The CES Cost Function and the Sustainable Configuration of Multiproduct Firms." *Bulletin of Economic Research* 56 (4): 329–332.

Johnes, G. and J. Johnes (2009): "Higher Education Institutions' Costs and Efficiency: Taking the Decomposition a Further Step." *Economics of Education Review* 28 (1): 107–113.

Johnes, G.; J. Johnes; E. Thanassoulis; P. Lenton and A. Emrouznejad (2005): *An Exploratory Analysis of the Cost Structure of Higher Education in England*. UK Department for Education and Skills.

Johnes, G. and A. Schwarzenberger (2011): "Differences in Cost Structure and the Evaluation of Efficiency: The Case of German Universities." *Education Economics* 19 (5): 487–499.

Johnes, J. (2014): "Efficiency and Mergers in English Higher Education 1996/97 to 2008/9: Parametric and Non-Parametric Estimation of the Multi-Input Multi-Output Distance Function." *The Manchester School* 82 (4): 465–487.

Kempkes, G. and C. Pohl (2010): "The Efficiency of German Universities–some Evidence from Nonparametric and Parametric Methods." *Applied Economics* 42 (16): 2063–2079.

Koshal, R.K. and M. Koshal (1999): "Economies of Scale and Scope in Higher Education: A Case of Comprehensive Universities." *Economics of Education Review* 18 (2): 269–277.

Kumbhakar, S.C. (1990): "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency." *Journal of Econometrics* 46 (1): 201–211.

Kumbhakar, S.C. (1997): "Modeling Allocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship." *Journal of Econometrics* 76 (1): 351–356.

Kumbhakar, S.C. and A. Heshmati (1995): "Efficiency Measurement in Swedish Dairy Farms: An Application of Rotating Panel Data, 1976–88." *American Journal of Agricultural Economics* 77 (3): 660–674.

Kumbhakar, S.C.; G. Lien and J.B. Hardaker (2014): "Technical Efficiency in Competing Panel Data Models: A Study of Norwegian Grain Farming." *Journal of Productivity Analysis* 41 (2): 321–337.

Kumbhakar, S.C.; H. Wang and A.P. Horncastle (2015): *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press.

Longlong, H.; L. Fengliang and M. Weifang (2009): "Multi-Product Total Cost Functions for Higher Education: The Case of Chinese Research Universities." *Economics of Education Review* 28 (4): 505–511.

Olivares, M. and H. Wetzel (2014): "Competing in the Higher Education Market: Empirical Evidence for Economies of Scale and Scope in German Higher Education Institutions." *CESifo Economic Studies* 60 (4): 653–680.

Pitt, M.M. and L.-F. Lee (1981): "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry." *Journal of Development Economics* 9 (1): 43–64.

Sav, G.T. (2012): "For-Profit College Entry and Cost Efficiency: Stochastic Frontier Estimates vs Two-Year Public and Non-Profit Colleges." *International Business Research* 5 (3): 26.

Stevens, P.A. (2005): "A Stochastic Frontier Analysis of English and Welsh Universities." *Education Economics* 13 (4): 355–374.

Titus, M.A.; A. Vamosiu and K.R. McClure (2016): "Are Public Master's Institutions Cost Efficient? A Stochastic Frontier and Spatial Analysis." *Research in Higher Education*, 1–28.

Tsionas, E.G. (2002): "Stochastic Frontier Models with Random Coefficients." *Journal of Applied Econometrics* 17 (2): 127–147.

Worthington, A.C. (2001): "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education." *Education Economics* 9 (3): 245–268.

Zoghbi, A.C.; F. Rocha and E. Mattos (2013): "Education Production Efficiency: Evidence from Brazilian Universities." *Economic Modelling* 31: 94–103.

Appendix

A. Regression Results, Battese and Coelli (1992) and Kumbhakar (1990)

		Battese and Coelli (1992)		Kumbhakar (1990)	
TPF, Sc		0.529*	(1.9)	0.538**	(1.9)
TPF, NSc		0.329	(1.1)	0.331	(1.1)
Stud, Sc		0.569**	(2.1)	0.555**	(2.1)
Stud, NSc		0.504**	(2.0)	0.467*	(1.8)
Wage		2.681***	(2.9)	2.678***	(2.9)
TPF2, Sc		0.059***	(3.5)	0.059***	(3.5)
TPF2, NSc		0.008	(0.3)	0.008	(0.3)
Stud2, Sc		0.102***	(6.8)	0.102***	(6.7)
Stud2, NSc		0.075***	(6.3)	0.074***	(6.2)
Wage2		-0.668***	(-3.6)	-0.673***	(-3.6)
Stud, Sc	*x* TPF, Sc	-0.080***	(-2.6)	-0.079**	(-2.5)
Stud, Sc	*x* TPF, NSc	-0.057	(-1.4)	-0.058	(-1.4)
Stud, NSc	*x* TPF, Sc	-0.010	(-0.2)	-0.012	(-0.3)
Stud, NSc	*x* TPF, NSc	-0.071***	(-2.8)	-0.073***	(-2.9)
TPF, Sc	*x* TPF, NSc	-0.051	(-1.4)	-0.051	(-1.4)
Stud, Sc	*x* Stud, NSc	-0.124**	(-2.5)	-0.125**	(-2.5)
Wage	*x* TPF, Sc	-0.049	(-0.9)	-0.05	(-1.0)
Wage	*x* TPF, NSc	-0.067	(-1.2)	-0.067	(-1.2)
Wage	*x* Stud, Sc	-0.148***	(-2.7)	-0.144***	(-2.6)
Wage	*x* Stud, NSc	-0.087*	(-1.9)	-0.080*	(-1.7)
East		-0.226***	(-3.0)	-0.228***	(-3.0)
Constant		-7.927***	(-3.0)	-0.104***	(-2.9)
$\gamma 1$		-0.043***	(-9.6)	0.089***	(6.5)
$\gamma 2$				-0.001	(-1.5)
σ_u^2		0.224***			
σ_v^2		0.017***			
No. of observations		949		949	
No. of institutions		73		73	
Log likelihood		422.52		422.92	

Source: Own calculations.
Note: *p=0.10, **p=0.005, ***p=0.001; t-statistics in parentheses.
Abbreviations: TPF = Third Party Funding.

B. Regression Results, Model 1 - 3

		Model 1		Model 2		Model 3	
TPF, Sc		0.529*	(1.9)	0.594**	(2.3)	0.607*	(1.9)
TPF, NSc		0.329	(1.1)	-0.027	(-0.1)	-0.060	(-0.2)
Stud, Sc		0.569**	(2.1)	0.740***	(2.6)	0.630**	(2.0)
Stud, NSc		0.504**	(2.0)	0.860***	(5.5)	0.714***	(2.7)
Wage		2.681***	(2.9)	2.70***	(3.8)	2.789***	(2.6)
TPF2, Sc		0.059***	(3.5)	0.047***	(3.8)	0.050***	(2.8)
TPF2, NSc		0.008	(0.3)	-0.021	(-1.0)	0.007	(0.3)
Stud2, Sc		0.102***	(6.8)	0.108***	(10.7)	0.113***	(7.4)
Stud2, NSc		0.075***	(6.3)	0.094***	(9.7)	0.094***	(7.6)
Wage2		-0.668***	(-3.6)	-0.770***	(-5.9)	-0.764***	(-3.5)
Stud, Sc	*x* TPF, Sc	-0.080***	(-2.6)	-0.026	(-1.2)	0.000	(-0.0)
Stud, Sc	*x* TPF, NSc	-0.057	(-1.4)	-0.001	(0.0)	-0.036	(-0.7)
Stud, NSc	*x* TPF, Sc	-0.010	(-0.2)	0.087***	(2.8)	0.070	(1.4)
Stud, NSc	*x* TPF, NSc	-0.071***	(-2.8)	-0.078***	(-3.4)	-0.090***	(-3.2)
TPF, Sc	*x* TPF, NSc	-0.051	(-1.4)	-0.081***	(-3.0)	-0.103**	(-2.4)
Stud, Sc	*x* Stud, NSc	-0.124**	(-2.5)	-0.302***	(-7.9)	-0.283***	(-5.5)
Wage	*x* TPF, Sc	-0.049	(-0.9)	-0.126**	(-2.3)	-0.131**	(-2.2)
Wage	*x* TPF, NSc	-0.067	(-1.2)	-0.044	(-0.8)	-0.002	(-0.0)
Wage	*x* Stud, Sc	-0.148***	(-2.7)	-0.065	(-1.1)	-0.048	(-0.8)
Wage	*x* Stud, NSc	-0.087*	(-1.9)	-0.0872***	(-2.9)	-0.079	(-1.6)
East		-0.226***	(-3.0)	-0.130***	(-10.5)	-0.086	(-1.5)
Constant		-7.927***	(-3.0)	-8.663***	(-3.8)	-8.407***	(-2.7)
γ		-0.043***	(-9.6)				
σ_u^2		0.224***		0.184***			
σ_v^2		0.017***		0.093***			
σ_u^2 *from step 2*						0.017***	
σ_v^2 *from step 2*						0.013***	
σ_u^2 *from step 3*						0.088***	
σ_v^2 *from step 3*						0.018***	
No. of observations		949		949		949	
No. of institutions		73		73		73	
Log likelihood		422.92		356.29		520.08	

Source: Own calculations.

Note: *p=0.10, **p=0.005, ***p=0.001; t-statistics in parentheses.
The estimation results from Model 3 are from the baseline model, first step.

Abbreviations: TPF = Third Party Funding.

C. Mean Efficiency Values, Model 1 - 3

University	**Model 1**	**Model 2**	**Model 3,** Overall	**Model 3,** *Persistent*	**Model 3,** *Residual*
Bauhaus-U Weimar	0.426	0.868	0.647	0.717	0.901
Europa-U Viadrina	0.594	0.874	0.801	0.885	0.905
FU Berlin	0.653	0.892	0.782	0.858	0.911
H Vechta	0.646	0.865	0.733	0.812	0.903
H Speyer	0.445	0.811	0.391	0.446	0.878
Helmut-Schmidt-Universität	0.259	0.788	0.340	0.395	0.862
HU Berlin	0.873	0.814	0.834	0.944	0.883
TH Aachen	0.508	0.883	0.719	0.793	0.907
TU Bergakademie Freiberg	0.432	0.854	0.689	0.767	0.898
TU Berlin	0.581	0.889	0.782	0.859	0.910
TU Braunschweig	0.500	0.884	0.670	0.739	0.907
TU Chemnitz	0.544	0.878	0.774	0.852	0.908
TU Clausthal	0.338	0.860	0.532	0.593	0.898
TU Darmstadt	0.559	0.876	0.767	0.846	0.907
TU Dresden	0.671	0.876	0.839	0.922	0.910
TU Hamburg-Harburg	0.422	0.829	0.638	0.716	0.891
TU Ilmenau	0.456	0.887	0.740	0.814	0.909
TU Kaiserslautern	0.595	0.888	0.764	0.840	0.910
TU München	0.449	0.839	0.613	0.686	0.894
U Augsburg	0.866	0.873	0.839	0.924	0.909
U Bamberg	0.720	0.865	0.798	0.886	0.901
U Bayreuth	0.540	0.889	0.679	0.748	0.908
U Bielefeld	0.598	0.884	0.732	0.806	0.908
U Bochum	0.580	0.883	0.741	0.816	0.908
U Bonn	0.520	0.849	0.664	0.740	0.897
U Bremen	0.651	0.863	0.730	0.810	0.901
U Dortmund	0.649	0.879	0.800	0.880	0.908
U Düsseldorf	0.557	0.855	0.713	0.792	0.901
U Erfurt	0.346	0.861	0.515	0.577	0.893
U Erlangen-Nürnberg	0.651	0.874	0.794	0.875	0.907
U Flensburg	0.983	0.887	0.872	0.955	0.913
U Frankfurt a.M.	0.697	0.851	0.795	0.880	0.904
U Freiburg i.Br.	0.632	0.887	0.742	0.817	0.909
U Gießen	0.505	0.881	0.665	0.734	0.905
U Greifswald	0.521	0.865	0.762	0.844	0.902
U Göttingen	0.564	0.864	0.701	0.778	0.900
U Halle	0.404	0.876	0.639	0.708	0.902
U Hamburg	0.643	0.880	0.747	0.824	0.907
U Hannover	0.555	0.864	0.699	0.774	0.903
U Heidelberg	0.690	0.869	0.778	0.861	0.904
U Hildesheim	0.698	0.846	0.772	0.861	0.897
U Hohenheim	0.444	0.837	0.588	0.659	0.892

University	Model 1	Model 2	Model 3, Overall	Model 3, *Persistent*	Model 3, *Residual*
U Jena	0.523	0.874	0.752	0.830	0.906
U Karlsruhe	0.586	0.870	0.758	0.839	0.904
U Kassel	0.694	0.880	0.810	0.891	0.909
U Kiel	0.646	0.880	0.790	0.870	0.908
U Koblenz-Landau	0.969	0.874	0.864	0.948	0.911
U Konstanz	0.594	0.875	0.697	0.772	0.903
U Köln	0.791	0.876	0.829	0.913	0.908
U Leipzig	0.540	0.891	0.745	0.820	0.910
U Lübeck	0.243	0.849	0.646	0.722	0.895
U Lüneburg	0.923	0.869	0.842	0.927	0.909
U Magdeburg	0.529	0.876	0.784	0.864	0.907
U Mainz	0.637	0.883	0.784	0.862	0.909
U Mannheim	0.642	0.875	0.741	0.817	0.906
U Marburg	0.556	0.877	0.722	0.796	0.906
U München	0.693	0.890	0.796	0.874	0.911
U Münster	0.671	0.882	0.782	0.859	0.910
U Oldenburg	0.526	0.883	0.644	0.711	0.905
U Osnabrück	0.537	0.887	0.652	0.719	0.907
U Paderborn	0.715	0.890	0.821	0.900	0.912
U Passau	0.762	0.888	0.811	0.890	0.911
U Potsdam	0.663	0.882	0.834	0.916	0.910
U Regensburg	0.643	0.882	0.771	0.849	0.908
U Rostock	0.478	0.881	0.733	0.807	0.908
U Siegen	0.605	0.868	0.758	0.839	0.904
U Stuttgart	0.536	0.838	0.719	0.805	0.894
U Trier	0.736	0.877	0.812	0.894	0.908
U Tübingen	0.636	0.887	0.744	0.819	0.909
U Ulm	0.311	0.858	0.475	0.531	0.896
U Wuppertal	0.600	0.878	0.763	0.840	0.908
U Würzburg	0.697	0.878	0.813	0.894	0.909
U des Saarlandes	0.667	0.826	0.793	0.885	0.896

Source: Own calculations.