Current Politics and Economics of Europe ISSN: 1057-2309
Volume 30, Issue 2

THE EDUCATIONAL PERFORMANCE OF SPANISH SECONDARY SCHOOLS IN PISA, 2003-2012*

Manuel Salas-Velasco
Department of Applied Economics, University of Granada, Spain

ABSTRACT

This chapter evaluated the performance of Spanish secondary schools whose 15-year-old students were assessed in mathematical competencies by the OECD (PISA program) in 2003 and 2012. The technique employed was the stochastic frontier analysis of panel data for a sample of schools which participated simultaneously in both waves. Our research revealed that Spanish schools increased their inefficiency over time. While typical panel data models assume a time-invariant "inefficiency effect," this assumption must be regarded as highly unrealistic for nine years. Parametric measurement of time-varying technical inefficiency was done in this chapter using three well-known models. We also introduced a new model that disentangled firm-effects from persistent (time-invariant) and residual (time-varying) technical inefficiency. Persistent technical inefficiency was a larger problem than residual technical inefficiency when evaluating the educational performance of Spanish secondary schools over time. The results shown should worry us because the average socio-economic status of the families increased significantly in this period—the expenditure per student as well. Although schools' increase of their socio-economic status led to an improved operating environment in 2012

* This is a reformatted version of a chapter previously published in: Nata, Roberta V. *Progress in Education. Volume 35.* New York: Nova Science Publishers, 2018.

compared to 2003, possible gains were not realized due to increased inefficiency in the period under analysis. Inefficiency was presumably not caused by something unexpected within each year, but rather by persistent factors such as school management decisions or regional/state regulations.

1. INTRODUCTION

Measuring the performance of educational institutions is vital for judging the degree to which resources made available to the education sector are utilized efficiently in the process of obtaining desired outcomes. Efficient use of resources occurs when the observed outputs from education (such as test results) are produced at the lowest level of resource (Johnes et al., 2017). Given the tight fiscal situation facing most governments, improvements in efficiency may be a necessary precondition for any improvements in public support for education (Grosskopf et al., 2014). For example, Spanish schools obtain their funds mostly from regional governments. Education is viewed as a responsibility of the individual regions (called *comunidades autónomas*) rather than the national government. Evaluation of performance and its public availability is also the essential characteristic of accountability. School accountability—the process of evaluating school performance on the basis of student performance measures—is increasingly prevalent around the world (Figlio & Loeb, 2011). Enhancing public accountability of the education sector in Spain is part of the constitutional mandate as well. The current Spanish Constitution of 1978 establishes in its article 31.2 that "*public expenditure will make an equitable allocation of public resources and its programming and execution will comply with criteria of efficiency and economy*."

In this context, the purpose of this chapter was to estimate the school-level technical efficiency secondary schools of Spain that participated in the OECD-PISA tests. The information available from the Spanish Ministry of Education allowed us to identify the same schools that participated in several waves in the international assessment. Specifically, the largest number of cases was for those schools that participated in 2003 and 2012. Also, as both waves focused on mathematical performance, the choice of these two years was also relevant. In our research, the concept of technical efficiency relied on Farrell (1957), meaning the capacity of schools to generate the maximum output given the inputs they used. For each school, output was measured by the average scores of 15-year-old students on the PISA mathematics literacy in the period under study. The stochastic frontier approach undertaken proved that the Spanish schools, as an industry, increased their inefficiency in the production of mathematical skills between 2003 and 2012. Consequently, there is room for efficiency improvement in the Spanish educational system. The presence of X-inefficiency in schools may mean that more efficient management of the school could lead to substantially better outcomes without increasing spending (Levin, 1997). Many countries introduced school autonomy or school-based management to improve student's performance (Dimmock, 2005; Mohrman & Wohlstetter, 1994).

Different stochastic frontier models—translog production functions—were used to estimate education production functions and the efficiency of Spanish secondary schools. In fact, in the empirical literature on frontier modeling with panel data, the issues of model specification and selection of estimation technique have not been explored in detail. The objective of this research was also to contribute to the existing literature providing a more comprehensive comparison of the most widely applied model specifications used to measure output-oriented time-varying technical inefficiency with panel data. The chapter explored the sensitivity of obtained technical inefficiency estimates to the choice of the model while maintaining an identical dataset and retaining the same assumptions about the underlying production technology. We were also able to break down the productive inefficiency of schools into two parts, one persistent and one transient. The persistent part is related to the presence of structural problems in the organization of the production process of a school or the presence of systematic shortfalls in managerial capabilities. The transient part may be due to the presence of non-systematic management problems that can be solved in the short term. The received empirical literature on the measurement of productive efficiency has paid relatively little attention to these two components of productive efficiency. As far as we know, this is the first study which measured the efficiency of Spanish schools over time and also addressed the possibility that the production efficiency can be split into two parts (i.e., transient and persistent).

The chapter was organized into five sections, apart from this introduction. Section 2 reviewed literature regarding the definition and measurement of technical efficiency in education. Section 3 introduced the methodology for the study of the in(efficiency) in education production. Section 4 presented the data and variables. Section 5 presented the results. Finally, section 6 showed the conclusion of this study.

2. BACKGROUND

The microeconomic theory of production looks at the activity of an organization mainly as a production process that transforms inputs (such as capital and labor) into output(s). In this context, the concepts of productivity and efficiency, although related, are different. The concept of productivity is more general and it can be measured by the ratio of the product obtained and the consumption of resources it is necessary to achieve that product (ratio between system outputs and inputs in the system), while efficiency indicates how well the resources are used in the transformation process of inputs into output(s). Technical efficiency reflects the ability of the firm to produce the maximum amount of output from a set of resource inputs, given the technology (output-orientation).[1] This technical efficiency measure assumes that the production function of the fully efficient firm is known.[2] As this is usually not the case, the efficient frontier must be estimated using sample data. Two alternative techniques have mainly been used in the applied research: (i) a non-parametric approach to identify the efficient production frontier; and (ii) estimation of a stochastic production frontier.

When dealing with multiple inputs yielding multiple outputs, efficiency literature usually makes use of data envelopment analysis (DEA) frontier methods (Cooper et al., 2011). DEA is a non-parametric technique that, through linear programming, approximates the true but unknown production function

[1] Technical efficiency in production is an empirical concept that involves the comparison of the maximum potential output, given a combination of inputs, with the observed output value. Koopmans (1951) already defined technical efficiency as the capability of a firm to maximize output for given inputs.

[2] In Microeconomics, a production function is a mathematical relationship that indicates the highest output that a firm can produce for every specified combination of inputs (the physical relationship between inputs and output) while holding technology constant at some predetermined state. The microeconomics analysis assumes that the firm operates on its production possibility frontier. On the other hand, allocative efficiency, also called price efficiency, exhibits the ability of this economic unit of production to choose an optimal set of resources, given the input prices.

without imposing any restriction on the sample distribution.[3] Nevertheless, the non-parametric DEA suffers from being highly vulnerable to potential outliers and measurement error, because every unit is related to the most efficient units. As an alternative analytical technique, stochastic frontier analysis (SFA), which uses regression analysis, is used to construct relative efficient production frontiers from performance data on output and inputs for samples of firms, organizations, institutions and economic agents of different kinds that are pursuing similar activities. The parametric SFA was criticized for relying on restrictive assumptions concerning the functional form and the distribution of the inefficiency term.

Applied efficiency analysis in education has used extensively both approaches to estimate the efficiency of educational institutions.[4] Identifying how efficiently education is provided has challenged researchers over decades. Some early studies already considered the possibility that technical inefficiency exists in educational production (e.g., Levin, 1974). For example, Bessent and Bessent (1980), Bessent et al. (1982), Färe et al., (1989), and McCarty & Yaisawarng, 1993) made one of the first attempts to account for the technical inefficiency of schools using DEA. We also found many examples of empirical studies which applied DEA (and related non-parametric techniques) to measuring the productive efficiency of universities (e.g., Abbott & Doucouliagos, 2003; Agasisti & Dal Bianco, 2009; Agasisti & Wolszczak-Derlacz, 2015; Halkos et al., 2012; Johnes, 2006; Tyagi et al., 2009; Wolszczak-Derlacz & Parteka, 2011).[5]

More recent papers also investigated educational productivity and efficiency via non-parametric methods as well. For example, Johnson and Ruggiero (2014) extended the non-parametric approach to decompose the Malmquist Productivity Index s into efficiency, technological and environmental changes. The approach was applied to analyze the educational

[3] It does not require a specification a priori of the functional form for the production function. DEA requires neither that output nor input prices be available for this analysis, which is a significant advantage for public sector applications where such prices are not available. See Simar and Wilson (2015) for a recent review of the development of various non-parametric approaches since the early work of Farrell.

[4] See Agasisti and Munda (2017), Emrouznejad and Yang (2018), and De Witte and López-Torres (2017), among others, for a recent review on efficiency in education.

[5] See Thanassoulis et al., (2016) for a comprehensive review of applications of DEA in secondary and tertiary education.

production of Ohio school districts. They found that changes in environmental harshness were the primary drivers in productivity changes of underperforming school districts. In another application in school assessment, Harrison and Rouse (2014) investigated the effect of competition on public high school performance using data from New Zealand public high schools. For this purpose, a categorical DEA model was used to calculate efficiency scores for schools. A regression model was then used to analyze differences in these efficiency scores and their relationship to different levels of competition. The study found that average school performance tended to be higher when schools were located in areas of high competition.

Thanks to the participation of a wide group of countries in large-scale international assessments, researchers have now access to rich databases that they use to assess the performance of educational systems in an international context. For example, using OECD PISA 2012 data, and DEA as a method of analysis, Agasisti, and Zoido (2018) derived efficiency scores for around 8,500 schools in 30 countries. Their work showed an important inefficiency of the schools in the countries studied. On average, achievement scores of schools could be increased by 27%, holding inputs constant. Using the same dataset, Aparicio et al. (2018) showed that the majority of schools in OECD countries tend to be less efficient in reading than in mathematics. DEA non-radial measures allowed them identifying different levels of inefficiency for each output considered (reading and maths).

The parametric frontier approach to efficiency measurement involving the specification and estimation of a parametric representation of the production technology has also been applied extensively in education to analyze whether educational institutions operate at an efficient level and to explain any inefficiencies. For example, this methodology was used by Cooper and Cohn (1997) and Wyckoff and Lavinge (1991) to estimate technical efficiency using school district data from South Carolina and New York, respectively. Recent papers also used SFA for studying the technical efficiency of higher education institutions. For example, Johnes's results (2014) showed that average efficiency in the English university sector over the period 1996/97 to 2008/9 was around 75% to 83%. In the analysis, the author estimated an output distance function using both SFA and DEA.

In recent years, an increasing number of papers used more advanced econometric frontier efficiency measurement techniques. For example, Kumbhakar et al., (2014) were the first authors to decompose (in)efficiency into short-term/residual (in)efficiency and long-term/persistent (in)efficiency, while controlling for heterogeneity.[6] While short-term inefficiency can be interpreted in the context of a chosen year, persistent inefficiency indicates operational problems at the institutional level. Titus et al., (2017) applied this approach to the US higher education sector, showing that cost inefficiency tends to be persistent rather than short-term. Recently, in the analysis of the German higher education sector, Gralka (2018) showed that heterogeneity among universities as well as persistent inefficiency hindered the institutions in achieving full efficiency. In any case, recent applications demonstrated that research findings are highly sensitive to modeling choices (Johnes & Tone, 2016), with parametric methods providing lower estimates of efficiency than non-parametric methods (Johnes, 2014). This suggests that a mixed-methods approach is most likely to lead to effective education reforms (Grosskopf et al., 2014).

3. Methodology

In this research, the measurement of efficiency—and its evolution over time—was based on the stochastic frontier methodology. Stochastic production frontier models were introduced by Aigner, Lovell, and Schmidt (1977), and Meeusen and van den Broeck (1977). The parametric frontier approach to efficiency measurement has been refined over the last decades with significant improvements in panel data models opening some new directions for empirical analysis. Important contributions to this refinement include those of Battese and Coelli (1995), Kumbhakar (1990), and Kumbhakar and Lovell (2000), among others. The use of panel data, repeated observations on each production unit, considerably enriched the econometric analysis of stochastic production frontier models with important advantages over cross-section data (Karagiannis & Tzouvelekas, 2009). First, it offered a more efficient econometric estimation of the production frontier model. Second, it provided consistent estimators of firm inefficiency, as long as the time dimension of the data set was sufficiently large. Third, it removed the necessity to make specific distributional assumptions

[6] The effect of unobserved firm heterogeneity on output.

regarding the one-sided error term associated with technical inefficiencies in the sample.

While typical panel data models assumed a time-invariant "inefficiency effect," this assumption must be regarded as highly unrealistic, as it excluded by assumption the possibility for firms to react to inefficiencies. Therefore, a natural extension of the model was to allow inefficiency to change over time. To accommodate the notion of productivity and efficiency improvement, different models were introduced in the applied research in which the inefficiency effects were time-varying.[7] Our goal was not to investigate all existing panel data models. We selected four of these models to evaluate the internal efficiency (technical efficiency) of Spanish secondary schools in the period 2003-2012.

We used the following generic formulation to discuss the models in a unifying framework. A stochastic frontier production function can be defined for panel data as follows

$$y_{it} = f(\boldsymbol{x}_{it}; \boldsymbol{\beta}) + \epsilon_{it} \tag{1}$$

where: y_{it} is the log of output for producer *i* at time *t*; $f(x_{it}; \beta)$ is the production technology; x_{it} is the vector of inputs (in logs); β is the associated vector of technology parameters to be estimated; and ϵ_{it} is the error term which is assumed to have two components in a stochastic frontier model: the standard (idiosyncratic) error term and a one-sided term that represents inefficiency

$$\epsilon_{it} = v_{it} - u_{it} \tag{2}$$

The v term takes care of the stochastic nature of the production process and possible measurement errors of output and inputs, and the u term is the possible inefficiency of the producer. The idiosyncratic error term is assumed to have a normal distribution. For the inefficiency term, it is assumed generally either a half-normal or a truncated-normal distribution. In long panels, it is desirable to allow technical inefficiency to vary over time. We introduced models in which the inefficiency effects were time-varying.

[7] See Kumbhakar et al., (2014, 2015) for an extensive review of these models and applications.

3.1. Model 1. The Battese and Coelli (1992) Time Decay Model

Battese and Coelli (1992) proposed a ML random-effects time-varying efficiency decay model [hereafter BC (1992) model]. The model was defined by the expressions (1) and (2).[8] In the latter, the idiosyncratic error term was assumed to have a normal distribution. And it was assumed a truncated-normal distribution for the inefficiency term (Battese & Coelli, 1992).[9] The inefficiency (u_{it}) was not fixed for a given individual; instead, it changed over time and also across individuals. Inefficiency in this model was composed of two distinct components (Battese & Coelli, 1992): (i) a non-stochastic time component; and (ii) a stochastic individual component

$$u_{it} = \eta_{it}\, u_i = \{\exp[-\eta(t - T_i)]\}\, u_i \tag{3}$$

where: T_i is the last period of the sample; η is the decay parameter; and $u_i \sim iid\ N^+(\mu, \sigma_u^2)\ (truncated - normal)$. When $\eta > 0$, the degree of inefficiency decreases over time. When $\eta < 0$, the degree of inefficiency increases over time. When $\eta = 0$, the time-varying decay model reduces to the time-invariant model (Battese & Coelli, 1988).

3.2. Model 2. The True Random-Effects Model (Greene, 2005)

A drawback of Model 1 was that individual heterogeneity—the effect of unobserved firm heterogeneity on output—could not be distinguished from inefficiency. In other words, all the time-invariant heterogeneity was confounded into inefficiency, and therefore $\widehat{u_i}$ might be picking up heterogeneity in addition to (or even instead of) inefficiency (Greene, 2004).[10] In an effort to address the limitations of the stochastic frontier analysis models with the random-effects specification, Greene (2005) proposed and developed

---

[8] ML stands for maximum likelihood estimation.

[9] The truncated-normal frontier model was due to Stevenson (1980).

[10] The panel data model of Battese and Coelli (1992) was also somewhat restrictive because it only allowed inefficiency to change over time exponentially.

the true random-effects stochastic frontier analysis [hereafter Greene (2005) model]. His 'true' random effect frontier model could be specified as

$$y_{it} = (\alpha + \omega_i) + f(\boldsymbol{x}_{it}; \boldsymbol{\beta}) + v_{it} - u_{it} \tag{4}$$

where: ω_i captures unobserved heterogeneity; u_{it} represents time-varying inefficiency; and v_{it} is a random shock with the following distributions.

$$u_{it} \sim N^{+}\left(0, \sigma_{it}^2\right) \tag{5}$$

$$v_{it} \sim N(0, \sigma_v^2) \tag{6}$$

$$\omega_i \sim N^{+}(0, \sigma_\omega^2) \tag{7}$$

The time-invariant/random constant term embodies the cross-country heterogeneity in the production function (not the mean of inefficiency). The one-sided inefficiency component now varies freely across time and country. Because of its complexity, the Greene's model is estimated by the maximum simulated likelihood method. We expect that models that allow time-invariant effects but do not treat them as inefficiency (as in Model 2) will give lower estimates of inefficiency (Kumbhakar et al., 2014).

3.3. Model 3. The Kumbhakar and Heshmati (1995) Model

Although Model 2 could separate firm-heterogeneity from time-varying inefficiency, none of the models discussed earlier considered persistent technical inefficiency. For example, Greene (2005) viewed firm-effects (unobservable individual effects or unobserved heterogeneity) as something other than inefficiency. Thus, this model failed to capture persistent inefficiency, which was confounded within firm-effects. Consequently, the model is likely to produce a downward bias in the estimate of overall inefficiency (Kumbhakar et al., 2015).[11]

Kumbhakar and Heshmati (1995) proposed a model [hereafter KH (1995) model] in which technical inefficiency was assumed to have a persistent firm-specific (time-invariant) component and a time-varying residual component. They separated persistent inefficiency from time-varying inefficiency as follows

$$y_{it} = \alpha_0 + f(\boldsymbol{x}_{it};\boldsymbol{\beta}) + v_{it} - \eta_i - u_{it} \tag{8}$$

where: $u_{it} \geq 0$ is time-varying inefficiency; $\eta_i \geq 0$ captures persistent technical inefficiency; and v_{it} is a random shock. In this model, overall technical inefficiency is $\eta_i + u_{it}$.[12]

Such a decomposition is desirable from a policy point of view because the persistent component is unlikely to change over time without any change in government policy or ownership of the firm, whereas the residual component changes both across firms and over time. However, the KH (1995) model did not take into account the firm-effects (unobservable individual effects) which were treated as long-term (persistent) inefficiency.[13] Consequently, the model is likely to produce an upward bias in inefficiency by treating firm-effects as inefficiency (Kumbhakar et al., 2015).

3.4. Model 4. The Kumbhakar, Lien, and Hardaker (2014) Model

Kumbhakar et al. (2014) proposed a model— hereafter KLH (2014) model—in which the error term was split into four components in order to take into account the various factors affecting inputs, given outputs

$$y_{it} = \alpha_0 + f(x_{it};\beta) + \mu_i + v_{it} - \eta_i - u_{it} \tag{9}$$

---

[11] Greene's model considered only time-varying inefficiency but not persistent inefficiency. Firm-effect was separated from time-varying inefficiency without taking into account the possibility of persistent inefficiency.

[12] For estimation purposes, see Kumbhakar et al., (2014, 2015).

[13] This model confounded firm-effects—that are not part of inefficiency—with persistent inefficiency.

where μ_i captures unobserved heterogeneity which is disentangled form persistent or time-invariant inefficiency (η_i) whereas u_{it} and v_{it} are time-varying technical inefficiency and random shocks respectively. The model therefore disentangled firm-effects from persistent (time-invariant) and residual (time-varying) technical inefficiency. Until recently, panel data stochastic frontier models were not able to distinguish among these components.[14]

4. DATA

Researchers in education economics seldom have access to panel data on individual schools in sufficient detail to estimate the models that we introduced before. Fortunately, in this research, we were able to use an own database which included a representative sample of Spanish secondary schools. The information available from the Spanish Ministry of Education allowed us to identify the schools that participated simultaneously in the OECD PISA tests in 2003 and 2012. How did we build the PISA panel data 2003 and 2012? The construction of the panel was somewhat complicated. The Spanish Ministry of Education, on its website, provided a panel of all the (anonymous) schools that participated in PISA in 2000, 2003, 2006, 2009, and 2012. Only one school had participated in all the waves. After doing several tests, the best option where there were more schools was in 2003 and 2012. We were able to identify 95 schools—in all schools, there are at least ten students—.[15] With their unique *schoolid*, we were able to plug in information in our panel data from the original PISA databases of schools and students.

The frontier production function was defined by the output and input variables summarized in Table 1. In Table 2 we showed the descriptive statistics. For each school, output (PV_MATH) was measured by the average scores of 15-year-old students on the PISA mathematics literacy in 2003 and 2012.[16] We took the original student database of each year, and we used *repest*

[14] For estimation purposes, see Kumbhakar et al., (2014, 2015).
[15] In both years, PISA focused on mathematics.
[16] Schools produce a variety of outputs, some of which are not easily quantifiable such as socialization. However, achievement test scores are normally the only output measure available.

in Stata to calculate the average per school that we then plugged into our school panel. *Repest* estimates statistics using replicate weights, thus accounting for complex survey designs in the estimation of sampling variances. It also allows for analyses with multiply imputed variables (plausible values) (Avvisati, 2016). In PISA, students' scores were calculated using an imputation method referred to as plausible values (PVs) which are a selection of likely proficiencies for students who attained each score. For each scale and subscale, five plausible values per student were included in the international database in 2003 and 2012.[17] The use in our analysis of the five plausible values is an improvement compared to recent studies that have used only a plausible value such as pv1math—the average score in mathematics obtained by the school—in the study of the efficiency of schools in various OECD countries (e.g., Agasisti & Zoido, 2018).

In relation to the inputs, we used the aggregate information at the school level provided directly from PISA databases: SCMATBUI (quality of schools' physical infrastructure), SCMATEDU (quality of schools' educational resources), and MACTIV (mathematics extracurricular activities at school). In addition, we took the original databases of students and got with *repest* the average socioeconomic status per school (ESCS) that we then plugged into our school panel. In the selection of the variables, we took into account, fundamentally, the availability in the database of factors that most of the studies have identified as inputs of educational production processes, such as buildings, school material, etc.[18] But since we are interested in studying the efficiency with which schools transform their resources into results, we must ensure that we have a sample of "homogeneous" producers where nondiscretionary factors are important contributors to output (Johnson & Ruggiero, 2014). We paid particular attention to the role of inputs over which schools have little or no control, such as schools' socio-economic status. The non-incorporation of inputs beyond schools' control may create the appearance of technical inefficiency. In words of Blackburn et al., (2014, p. 4): "failure to properly control for the socio-economic environment leads to inappropriate comparisons

---

[17] A full description of the scoring method can be found in the PISA 2012 technical report (OECD, 2014).

[18] We should caution the reader that (in a technical sense) we do not have a perfect mapping of inputs. For example, we also used the student-teacher ratio and the number of certified mathematics teachers, but the different models did not converge.

and biased efficiency estimates." As such, our models took this uncontrollable input directly into account in computing efficiency scores. Therefore, each school was compared to another, public or private, with similar resources in quantity and quality. In addition, the problem of multicollinearity was carefully addressed in the selection of these variables.

Table 1. School-level variables on output and inputs

Variable	***Type***	***Definition***
PV_MATH	Output	Schools' performance in mathematics. Students' mathematics scores in PISA 2003 and 2012, by school. To compute school-level averages, based on five PVs (plausible values), we used the Stata program and a user-contributed package *repest* (Avvisati, 2016).
ESCS	Input	Schools' socio-economic status. PISA index of economic, social and cultural status (ESCS), by school using *repest*. Higher scores indicated higher levels of socio-economic status. The PISA index of economic, social and cultural status was created on the basis of the following variables: the International Socio-Economic Index of Occupational Status (ISEI); the highest level of education of the student's parents, converted into years of schooling; the PISA index of family wealth; the PISA index of home educational resources; and the PISA index of possessions related to "classical" culture in the family home.
SCMATBUI	Input	Quality of schools' physical infrastructure. The index of quality of physicals' infrastructure was derived from three items measuring school principals' perceptions of potential factors hindering instruction at their school. These factors were: i) shortage or inadequacy of school buildings and grounds; ii) shortage or inadequacy of heating/cooling and lighting systems; and iii) shortage or inadequacy of instructional space (e.g., classrooms). Higher values on this index indicated better quality of educational resources.
SCMATEDU	Input	Quality of schools' educational resources. The index was computed on the basis of six items measuring the school principals' perceptions of potential factors hindering instruction at school such as shortage or inadequacy of instructional materials, shortage or inadequacy of computers for instruction, etc. Higher values on this index indicate better quality of educational resources.
MACTIV	Input	Mathematics extracurricular activities at school. The index was derived from school principals' reports on whether their schools offered the following activities to 15-year-olds: i) mathematics club; ii) mathematics competition; iii) club with a focus on computers/Information, Communication Technology; and iv) additional mathematics lessons. This index was developed by summing up the number of activities that a school offers.

Table 2. Descriptive statistics for variables included in the stochastic frontier production models§

	Schools	**Mean**	**Std. Dev.**
PV_MATH	182	504.574	41.847
ESCS	182	0.015	0.501
SCMATBUI	182	0.180	1.016
SCMATEDU	182	0.105	0.998
MACTIV	182	0.819	0.844
year = 2003	Obs.	Mean	Std. Dev.
PV_MATH	91	505.764	40.852
ESCS_M_B	91	-0.027	0.468
SCMATBUI	91	0.273	1.030
SCMATEDU	91	0.139	1.089
MACTIV	91	0.516	0.565
year = 2012	Obs.	Mean	Std. Dev.
PV_MATH	91	503.384	43.011
ESCS	91	0.056	0.532
SCMATBUI	91	0.087	0.999
SCMATEDU	91	0.070	0.902
MACTIV	91	1.121	0.964

§ Descriptive statistics using the original information in the PISA databases. All indices in PISA were scaled so that they had a mean 0 and a standard deviation of 1 for OECD countries. See OECD (2014) for more details.

5. RESULTS

This section presents the parameter estimates and efficiency scores generated from stochastic frontier analyses of panel data spanning the 2003 to 2012 time period. We choose a translog specification of the $f(x_{it};\beta)$ function in our empirical analysis in Models 1–4 because of its flexibility (Christensen et al., 1973). Despite the possible problems associated with estimating a translog functional form, analysts should (when possible) utilize it to estimate technical efficiency (Titus & Eagan, 2016).

A translog production model for one school output and four school inputs would be specified as

$$\ln y_{it} = \beta_0 + \sum_{k=1}^{4} \beta_k \ln x_{kit} + \frac{1}{2}\sum_{k=1}^{4}\sum_{m=1}^{4} \beta_{km} \ln x_{kit} \ln x_{mit} + v_{it} - u_{it} \quad (10)$$

with $\beta_{km} = \beta_{mk}$ between two inputs *k* and *m*.
where: y_{it} denotes the production of the *i*-th school in period *t*; x_{it} is a vector of input quantities of the *i*-th school in period *t*; β is a vector of unknown parameters to be estimated; and v_{it} and u_{it} are the random variables defined above.

Parameter estimates of the translog production frontier arising from the econometric estimation for each of the alternative models presented before are shown in Tables 3 through 6. We use log values for the input variables in the translog production function.[19] Within each model type, we provide parameter estimates with different error distributions for the inefficiency term: half-normal in Models 1, 2, and 3; and truncated-normal in Model 4.[20] In this regard, we would like to point out that the BC (1992) model with a truncated-normal distribution for the inefficiency term proved that *mu* was not statistically significant (results not shown here), so the BC (1992) model was estimated with a half-normal distribution. And the KLH (2014) model did not converge with a half-normal distribution; hence, a fixed-effect model with a truncated-normal distribution was estimated.[21]

[19] Prior to taking logs, the x-variables were rescaled to work with positive values greater than zero and divided by their geometric means. Consequently, the first-order coefficients can be interpreted as elasticities of output evaluated at the means of the data.

[20] Heteroscedastic problems may also exist in time-varying models. The heteroscedasticity could be introduced into u_{it} or v_{it} or both. However, this aspect was not taken into account in our analysis due to the difficulty of finding exogenous Z variables. In any case, the literature on stochastic frontiers indicates that the size of the bias is small when no heteroscedasticity is included in either error term (Conroy & Arguea, 2008).

[21] The complexity combined with the accompanying log-likelihood function means that the translog functional form of an SFA model may sometimes result in non-convergence (Titus & Eagan, 2016). For example, Greene's model (2005) of fixed effects with half-normal did not converge in our study. The KH (1995) model of random effects with half-normal did not converge. The KLH (2014) model with half-normal did not converge either with fixed effects or with random effects. The latter did not converge with a truncated-normal.

Table 3. A four-input translog production function: BC (1992) model

	Coef.		**Std. Err.**
ln(escs)	0.081	**	0.0090
ln(scmatbui)	-0.018		0.0144
ln(scmatedu)	0.026	**	0.0108
ln(mactiv)	0.013		0.0099
½[ln(escs)]^2	0.010	**	0.0018
½[ln(scmatbui)]^2	-0.039	**	0.0153
½[ln(scmatedu)]^2	0.020	**	0.0063
½[ln(mactiv)]^2	-0.002		0.0397
ln(escs)*ln(scmatbui)	0.053	**	0.0187
ln(escs)*ln(scmatedu)	-0.009		0.0153
ln(escs)*ln(mactiv)	0.029	*	0.0158
ln(scmatbui)*ln(scmatedu)	0.016		0.0148
ln(scmatbui)*ln(mactiv)	0.039	*	0.0233
ln(scmatedu)*ln(mactiv)	-0.039	*	0.0209
Constant	6.259	**	0.0092
/lnsigma2	-4.925	**	0.1886
/ilgtgamma	1.123	**	0.3247
/mu	(omitted)		
/eta	-0.216	*	0.1227
sigma2	0.007	**	0.0014
gamma	0.755	**	0.0601
sigma_u2	0.005	**	0.0014
sigma_v2	0.002	**	0.0003
Log-likelihood	272.087		
Wald chi2(14)	215.420		
Prob > chi2	0.000		
Number of obs.	182		
Number of groups	91		
Distribution for the inefficiency term	half-normal		
Overall mean efficiency	0.950		
Mean efficiency 2003	0.955		
Mean efficiency 2012	0.945		
output = ln(PV_MATH)			

*, ** represent 10%, 5% levels of significance, respectively.

Table 4. A four-input translog production function: Greene (2005) model

	Coef.		**Std. Err.**
ln(escs)	0.083	**	0.0098
ln(scmatbui)	-0.008		0.0152
ln(scmatedu)	0.019	*	0.0114
ln(mactiv)	0.003		0.0102
½[ln(escs)]^2	0.010	**	0.0019
½[ln(scmatbui)]^2	-0.031	*	0.0167
½[ln(scmatedu)]^2	0.016	**	0.0068
½[ln(mactiv)]^2	-0.013		0.0416
ln(escs)*ln(scmatbui)	0.045	**	0.0198
ln(escs)*ln(scmatedu)	-0.004		0.0159
ln(escs)*ln(mactiv)	0.027		0.0177
ln(scmatbui)*ln(scmatedu)	0.017		0.0154
ln(scmatbui)*ln(mactiv)	0.031		0.0245
ln(scmatedu)*ln(mactiv)	-0.036		0.0221
Constant	6.259	**	0.0120
Usigma			
Constant	-5.442	**	0.3191
Vsigma			
Constant	-7.548	**	0.8626
Theta			
Constant	-0.038	**	0.0070
sigma_u	0.066	**	0.0105
sigma_v	0.023	**	0.0099
lambda	2.866	**	0.0187
Log simulated-likelihood	266.643		
Number of pseudo random draws	250		
Wald chi2(14)	208.720		
Prob > chi2	0.000		
Number of obs.	182		
Number of groups	91		
Distribution for the inefficiency term	half-normal		
Overall mean efficiency	0.950		
output = ln(PV_MATH)			

*, ** represent 10%, 5% levels of significance, respectively.

Table 5. A four-input translog production function: KH (1995) model[§]

	Coef.		**Std. Err.**
ln(escs)	0.056	**	0.0190
ln(scmatbui)	-0.017		0.0208
ln(scmatedu)	0.041	**	0.0153
ln(mactiv)	0.004		0.0130
½[ln(escs)]^2	0.006	*	0.0033
½[ln(scmatbui)]^2	-0.021		0.0204
½[ln(scmatedu)]^2	0.029	**	0.0086
½[ln(mactiv)]^2	-0.052		0.0502
ln(escs)*ln(scmatbui)	0.025		0.0234
ln(escs)*ln(scmatedu)	-0.013		0.0207
ln(escs)*ln(mactiv)	0.032		0.0227
ln(scmatbui)*ln(scmatedu)	0.022		0.0189
ln(scmatbui)*ln(mactiv)	0.059	**	0.0295
ln(scmatedu)*ln(mactiv)	-0.056	**	0.0264
Constant	6.213	**	0.0078
sigma_u	0.059		
sigma_e	0.045		
rho	0.626	(fraction of variance due to u_i)	
F test that all u_i = 0			
F(90, 77)	2.55		
Prob > F	0.0000		
F(14,77)	6.53		
Prob > F	0.0000		
Number of obs.	182		
Number of groups	91		
Distribution for the inefficiency term	half-normal		
Persistent efficiency	0.893		
Residual efficiency	0.987		
Total efficiency	0.882		
[§] Which corresponds to the fixed-effects model			
output = ln(PV_MATH)			

*, ** represent 10%, 5% levels of significance, respectively.

Table 6. A four-input translog production function: KLH (2014) model[§]

	Coef.		**Std. Err.**
ln(escs)	0.056	**	0.0190
ln(scmatbui)	-0.017		0.0208
ln(scmatedu)	0.041	**	0.0153
ln(mactiv)	0.004		0.0130
½[ln(escs)]^2	0.006	*	0.0033
½[ln(scmatbui)]^2	-0.021		0.0204
½[ln(scmatedu)]^2	0.029	**	0.0086
½[ln(mactiv)]^2	-0.052		0.0502
ln(escs)*ln(scmatbui)	0.025		0.0234
ln(escs)*ln(scmatedu)	-0.013		0.0207
ln(escs)*ln(mactiv)	0.032		0.0227
ln(scmatbui)*ln(scmatedu)	0.022		0.0189
ln(scmatbui)*ln(mactiv)	0.059	**	0.0295
ln(scmatedu)*ln(mactiv)	-0.056	**	0.0264
Constant	6.213	**	0.0078
sigma_u	0.059		
sigma_e	0.045		
rho	0.626	(fraction of variance due to u_i)	
F test that all u_i = 0:			
F(90, 77)	2.55		
Prob > F	0.0000		
F(14,77)	6.53		
Prob > F	0.0000		
Number of obs.	182		
Number of groups	91		
Distribution for the inefficiency term	truncated-		
Persistent efficiency	0.954		
Residual efficiency	0.989		
Total efficiency	0.944		
[§] Which corresponds to the fixed-effects			
output = ln(PV_MATH)			

*, ** represent 10%, 5% levels of significance, respectively.

We estimated the BC (1992) and Greene (2005) models with Stata *sfpanel* command (Stata module for panel data stochastic frontier models estimation). On the other hand, the KH (1995) and KLH (2014) models were run using the programming and the routines contained in the book of Kumbhakar et al., (2015) and the Stata files available on the website: https://sites.google.com/site/sfbook 2014/. These commands are not part of the official Stata package but instead are commands that authors wrote themselves in the form of Stata ado-files.[22]

For all models the estimated first-order parameters (β) are having the anticipated (positive) sign and magnitude (being between zero and one), confirming that, all else being equal, a more significant amount of one input has a positive effect on aggregate output. The stochastic frontier estimates suggest a significant role played by schools' socio-economic background in determining performance in the PISA tests: the higher socio-economic status of the schools, the higher the academic performance in mathematics of the schools. Student performance also increases with a better quality of educational resources. However, indicators for higher quality physical resources and extracurricular activities of mathematics appear to not be correlated with better performance at the school level. The estimates coefficients in the stochastic frontier analysis are statistically insignificant. In any case, according to researchers (e.g., Johnes & Schwarzenberger, 2011; Laband & Lentz, 2003), the non-linear nature of a transcendental function does not allow for easy interpretation of the estimated coefficients. Instead, Greene (2005) suggests examining the efficiency scores generated by different specifications for stochastic frontier production models. The means for the efficiency scores are presented in Table 7.

In addition to testing hypotheses concerning β, stochastic frontier researchers are often interested in testing for the absence of inefficiency effects. In other words, we want to properly test whether technical inefficiency is present and, thus, whether frontier estimation is appropriate. Otherwise, ordinary least squares estimates would be more correct. In our research, the stochastic frontier analysis is appropriate for the description of the production technology. The null hypothesis of absence of random technical inefficiency (gamma = 0) is rejected and thus the stochastic frontier model seems quite appropriate for the data

[22] The ''generalized true random effects'' model also distinguishes between persistent and transient levels of efficiency (Filippini & Greene, 2016). However, the computation is not possible with Stata.

(Table 3).[23] Also, Table 4 provides values for lambda (λ). Lambda is the ratio of the non-random error component to the random error component. A lambda value greater than 1 suggests that the total error is composed of more inefficiency than of random noise. A significant lambda value suggests the existence of inefficiency. There is, therefore, inefficiency that is captured with the u term. In this regard, and as we will see in the next subsection, the results suggest that Spanish schools are moderately inefficient. However secondary schools showed a worsening of efficiency over time: when $\eta < 0$ (and statistically significant) the degree of inefficiency increases over time (Table 3). In any case, the comparison of the different models demonstrates how the results of the standard efficiency evaluation of the secondary education sector—BC (1992) model—change when taking heterogeneity and persistent inefficiency into account.

5.1. Technical Efficiency

This subsection shows an analysis of the efficiency scores generated from each model. Estimates of the mean output-oriented technical efficiency over schools for each model specification are presented in Table 7. The results indicate a significant variation in estimated mean technical efficiency scores with the mean values ranging from a low of 88.17% (Model 3) to a high of 95.01% (Model 1). In general, all the models reveal high mean technical efficiency values. On average, achievement scores of schools can be increased by 12%-5%, holding inputs constant.

As illustrated in Table 7, efficiency is estimated to be 95.01% on average in the BC (1992) model. However, it doesn't consider the unobserved heterogeneity.[24] As we said, this model confounds firm effects or unobserved heterogeneity (that are not part of inefficiency) with inefficiency. On the contrary, the Greene's model separates firm heterogeneity from inefficiency. As

[23] The appropriateness of the stochastic frontier approach can be tested by calculating the value of the parameter γ, which contains a value between 0 and 1 and depends on two variance parameters of the stochastic frontier function (Battese and Corra 1977; Coelli et al., 2005). If the value γ is close to 0, deviations from the frontier are attributed to noise, whereas a value close to 1 indicates that deviations are ascribed to technical inefficiency (Coelli et al., 2005; Tran et al., 2008).

[24] Neither the persistent inefficiency.

illustrated, efficiency is now estimated to be 94.97% on average. However, Greene's model considers only time-varying inefficiency but not persistent inefficiency: the firm effect is separated from time-varying inefficiency without taking into account the possibility of persistent inefficiency.

Table 7. The efficiency of secondary schools generated from stochastic frontier analyses of panel data

	Overall mean efficiency	**Persistent efficiency**	**Residual efficiency**
The Battese and Coelli model (1992)	0.95011		
The True Random-Effects model (Greene, 2005)	0.94974		
The Kumbhakar and Heshmati model (1995)	0.88168	0.89289	0.98744
The Kumbhakar, Lien, and Hardaker model (2014)	0.94388	0.95422	0.98913

Table 8. Spearman's rank-order correlation coefficients

		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
BC (1992) model	(1)	1							
Greene (2005) model	(2)	0.5317*	1						
KH (1995) model	(3)	0.9288*	0.5525*	1					
Persistent efficiency	(4)	0.9303*	0.5047*	0.9973*	1				
Residual efficiency	(5)	-0.0002	0.7507*	0.0589	0	1			
KLH (2014) model	(6)	0.9116*	0.6297*	0.9899*	0.9798*	0.1554*	1		
Persistent efficiency	(7)	0.9303*	0.5047*	0.9973*	1.0000*	0	0.9798*	1	
Residual efficiency	(8)	-0.0002	0.7507*	0.0589	0	1.0000*	0.1554*	0	1

* 5%-level of significance.

According to the KH (1995) model, as shown in Table 7, persistent efficiency is estimated to be 89.29% on average, residual efficiency 98.74%, and total efficiency of 88.17%. Finally, according to the KLH (2014) model, as illustrated also in Table 7, persistent efficiency is 95.42% (not 89.29% as predicted by the Kumbhakar-Heshmati formulation), residual efficiency is 98.91% (which was 98.74% in Kumbhakar-Heshmati formulation), and the overall efficiency was 94.39% (much higher than 88.17% predicted by the

Kumbhakar-Heshmati formulation). In contrast, the overall efficiency is lower compared to true random-effects model (95%).[25]

What can policymakers derive from the results so far? For policy purposes, it is relevant to know whether, and how, the institution-specific efficiency varies by estimation method. As the above results illustrate, the efficiency scores are sensitive to model specification. How then are the technical efficiency ranking of secondary schools affected by different model specifications used? In Table 8, pairwise rank order correlations for Models 1–4 illustrate the differences between the models in technical efficiency ranking of the sample. Table 8 shows that the rank correlations between the four models are significantly positive but varying (between 0.53 and 0.99) for the overall efficiency. First, there is a high correlation among BC (1992), KH (1995), and KLH (2014) models (overall efficiency). Second, the high correlation between overall and persistent efficiency in BC (1992), KH (1995), and KLH (2014) models demonstrates the strong contribution of the persistent term for the overall result. In short, our results suggest that persistent technical inefficiency is a larger problem than residual technical inefficiency when evaluating the educational performance of Spanish secondary schools over time. A high degree of persistent technical efficiency could reflect long-run problems. Finally, the residual efficiency generated by KH (1995) model and KLH (2014) model is highly correlated only with the overall efficiency of the Greene model. This result is consistent considering that Greene's model considers only time-varying inefficiency but not persistent inefficiency.

6. Conclusion

Research on educational production functions attempts to model the relationship between resource inputs and school outcomes such as educational achievement. How to improve school performance and achieve better outcomes is a question that has faced the public sector for some time and one that has received ever-increasing attention for the sake of enhancing accountability. In this context, this chapter provides an empirical comparison of time-varying

---

[25] As mentioned earlier, failure to separate firm-effects from persistent inefficiency in the true random-effects model is likely to produce biased estimates of overall efficiency.

technical inefficiency measures obtained from the econometric estimation of different specifications of the stochastic production frontier model. Specifically, four different frontier model specifications, which are most widely used in empirical applications, are estimated using a balanced panel dataset of the Spanish secondary schools observed during the 2003-2012 period. The results suggest that Spanish schools are moderately inefficient and that inefficiency is presumably not caused by something unexpected within each year, but rather by persistent factors such as school management decisions or regional/state regulations. Therefore, understanding more deeply the sources of inefficiency in secondary education in Spain is sufficiently important to justify our ongoing research on these topics.

REFERENCES

Abbott, M., & Doucouliagos, C. (2003). The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review*, *22*(1), 89–97.

Agasisti, T., & Dal Bianco, A. (2009). Reforming the university sector: Effects on teaching efficiency—evidence from Italy. *Higher Education*, *57*(4), 477–498.

Agasisti, T., & Munda, G. (2017). *Efficiency of investment in compulsory education: An overview of methodological approaches* (*Joint Research Centre Technical Report* No. 106681).

Agasisti, T., & Wolszczak-Derlacz, J. (2015). Exploring efficiency differentials between Italian and Polish universities, 2001-11. *Science and Public Policy*, *43*(1), 128–142.

Agasisti, T., & Zoido, P. (2018). Comparing the efficiency of schools through international benchmarking: Results from an empirical analysis of OECD PISA 2012 data. *Educational Researcher*, *20*(10), 1–11.

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

Aparicio, J., Cordero, J. M., Gonzalez, M., & Lopez-Espin, J. J. (2018). Using non-radial DEA to assess school efficiency in a cross-country perspective: An empirical analysis of OECD countries. *Omega*, *79*, 9–20.

Avvisati, F. (2016*). REPEST: Stata module to run estimations with weighted replicate samples and plausible values.* Retrieved from https://ideas.repec.org/c/boc/bocode/s457918.html.

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

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*(2), 325–332.

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

Bessent, A. M., & Bessent, E. W. (1980). Determining the comparative efficiency of schools through data envelopment analysis. *Educational Administration Quarterly*, *16*(2), 57–75.

Bessent, A., Bessent, W., Kennington, J., & Reagan, B. (1982). An application of mathematical programming to assess productivity in the Houston independent school district. *Management Science*, *28*(12), 1355–1367.

Blackburn, V., Brennan, S., & Ruggiero, J. (2014). Measuring efficiency in Australian schools: A preliminary analysis. *Socio-Economic Planning Sciences*, *48*(1), 4–9.

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

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An Introduction to Efficiency and Productivity Analysis* (2nd ed.). New York: Springer.

Conroy, S. J., & Arguea, N. M. (2008). An estimation of technical efficiency for Florida public elementary schools. *Economics of Education Review*, *27*(6), 655–663.

Cooper, S. T., & Cohn, E. (1997). Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review*, *16*(3), 313–327.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). Data envelopment analysis: History, models, and interpretations. In W. W. Cooper, L. M. Seiford & J. Zhu (Eds.), *Handbook on Data Envelopment Analysis* (2nd ed.) (pp. 1–39). New York: Springer.

De Witte, K., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society*, *68*(4), 339–363.

Dimmock, C. (2005). *School-Based Management and School Effectiveness*. London and New York: Routledge.

Emrouznejad, A., & Yang, G. L. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, *61*, 4–8.

Färe, R., Grosskopf, S., & Weber, W. L. (1989). Measuring school district performance. *Public Finance Quarterly*, *17*(4), 409–428.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, *120*(3), 253–290.

Figlio, D., & Loeb, S. (2011). School accountability. In E. A. Hanushek, S. Machin & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 3, pp. 383–421). Amsterdam: Elsevier.

Filippini, M., & Greene, W. (2016). Persistent and transient productive inefficiency: A maximum simulated likelihood approach. *Journal of Productivity Analysis*, *45*(2), 187–196.

Gralka, S. (2018). Persistent inefficiency in the higher education sector: Evidence from Germany. *Education Economics*, *26*(4), 373–392.

Greene, W. (2004). Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics*, *13*(10), 959–980.

Greene, W. (2005). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, *23*(1), 7–32.

Grosskopf, S., Hayes, K. J., & Taylor, L. L. (2014). Efficiency in education: Research and implications. *Applied Economic Perspectives and Policy*, *36*(2), 175–210.

Halkos, G., Tzeremes, N. G., & Kourtzidis, S. A. (2012). Measuring public owned university departments' efficiency: A bootstrapped DEA approach. *Journal of Economics and Econometrics*, *55*(2), 1–24.

Harrison, J., & Rouse, P. (2014). Competition and public high school performance. *Socio-Economic Planning Sciences*, *48*(1), 10–19.

Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review*, *25*(3), 273–288.

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.

Johnes, G., & Schwarzenberger, A. (2011). Differences in cost structure and the evaluation of efficiency: The case of German universities. *Education Economics*, *19*(5), 497–499.

Johnes, G., & Tone, K. (2016). The efficiency of higher education institutions in England revisited: Comparing alternative measures. *Tertiary Education and Management*, 1–15.

Johnes, J., Portela, M., & Thanassoulis, E. (2017). Efficiency in education. *Journal of the Operational Research Society*, *68*, 331–338.

Johnson, A. L., & Ruggiero, J. (2014). Nonparametric measurement of productivity and efficiency in education. *Annals of Operations Research*, *221*(1), 197–210.

Karagiannis, G., & Tzouvelekas, V. (2009). Parametric measurement of time-varying technical inefficiency: Results from competing models. *Agricultural Economics Review*, *10*(1), 50.

Koopmans, T. C. (1951). Analysis of production as an efficient combination of activities. In T. C. Koopmans (Ed.), *Activity Analysis of Production and Allocation. Cowles Commission for Research in Economics: Monograph No. 13* (pp. 33–97). New York: Wiley.

Kumbhakar, S. C. (1990). Production frontiers, panel data and time-varying technical inefficiency. *Journal of Econometrics*, *46*, 201–211.

Kumbhakar, S. C., & Heshmati, A. (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., & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. New York: Cambridge University Press.

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (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., Wang, H. J., & Horncastle, A. P. (2015): *A Practitioner's Guide to Stochastic Frontier Analysis using Stata*. New York, NY: Cambridge University Press.

Laband, D. N., & Lentz, B. F. (2003). New estimates of economies of scale and scope in higher education. *Southern Economic Journal*, 70(1), 172-183.

Levin, H. M. (1974). Measuring efficiency in educational production. *Public Finance Quarterly, 2*(1), 3–24.

Levin, H. M. (1997). Raising school productivity: An x-efficiency approach. *Economics of Education Review, 16*(3), 303–311.

McCarty, T. A., & Yaisawarng, S. (1993). Technical efficiency in New Jersey school districts. In H. O. Fried, C. A. K. Lovell & S. S. Schmidt (Eds.) (1993), *The Measurement of Productive Efficiency: Techniques and Applications* (pp. 271–287.) New York: Oxford University Press.

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

Mohrman, S. A., & Wohlstetter, P. (1994). *School-based management: Organizing for High Performance*. San Francisco: Jossey-Bass.

OECD (2014). *PISA 2012 Technical Report.* Paris: OECD Publishing.

Simar, L., & Wilson, P. W. (2015). Statistical approaches for non-parametric frontier models: A guided tour. *International Statistical Review, 83*(1), 77–110.

Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics, 13*(1), 57–66.

Thanassoulis E., De Witte, K., Johnes J., Johnes, G., Karagiannis, G., & Portela, C. S. (2016). Applications of data envelopment analysis in education. In J. Zhu (Ed.), *Data Envelopment Analysis: A Handbook of Empirical Studies and Applications* (pp. 367– 438). New York: Springer.

Titus, M. A., & Eagan, K. (2016). Examining production efficiency in higher education: The utility of stochastic frontier analysis. In M. B. Paulsen (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. 31, pp. 441–512). New York: Springer.

Titus, M. A., Vamosiu, A., & McClure, K. R. (2017). Are public master's institutions cost efficient? A stochastic frontier and spatial analysis. *Research in Higher Education*, *58*(5), 469–496.

Tran, T. B., Grafton, R. Q., & Kompas, T. (2008). Firm efficiency in a transitional economy: Evidence from Vietnam. *Asian Economic Journal*, *22*(1), 47–66.

Tyagi, P., Yadav, S. P., & Singh, S. P. (2009). Relative performance of academic departments using DEA with sensitivity analysis. *Evaluation and Program Planning*, *32*(2), 168–177.

Wolszczak-Derlacz, J., & Parteka, A. (2011). Efficiency of European public higher education institutions: A two-stage multicountry approach. *Scientometrics*, *89*(3), 887–917.

Wyckoff, J. H., & Lavinge, J. (1991). *The Relative Inefficiency of Public Elementary Schools in New York*. Work Paper, State University of New York, Albany.