The role of innovations in secondary school performance - Evidence from a conditional efficiency model

Innovative Applications of O.R.

The role of innovations in secondary school performance – Evidence from a conditional efficiency model

Carla Haelermans [a,b,*], Kristof De Witte [a,c]

[a] Maastricht University, TIER, Faculty of Humanities and Sciences, PO Box 616, 6200MD Maastricht, The Netherlands
[b] Centre for Innovations and Public Sector Efficiency Studies, Delft University of Technology, The Netherlands
[c] Katholieke Universiteit Leuven (KULeuven), Faculty of Business and Economics, Naamsestraat 69, 3000 Leuven, Belgium

ABSTRACT

This paper studies the influence of educational innovations on school performance. We apply a tailored, fully nonparametric conditional efficiency model to study secondary school efficiency in the Netherlands. The application uses official school data and a self-collected questionnaire on recent innovations in schools. In the nonparametric model, it is assumed that schools aim to maximize educational attainments of students under a budget constraint. The results suggest that innovations are positively related to efficiency. We find that profiling, pedagogic, process and education chain innovations are significantly related to school efficiency, whereas innovations in the professionalization of teachers are insignificantly related to school efficiency. Furthermore, the number of locations per school and the number of schools per governing body are negatively and significantly related to school efficiency. School type and region significantly influence school efficiency, whereas share of disadvantaged students, degree of urbanization and student/teacher ratio do not have significant influence.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Under tightening budget constraints, governments and schools are reducing the resources for educational innovations. Innovations correspond to a broad concept, which is defined here as: '*an idea, process or product that is new for an organization at the time it is introduced*' (cfr. Dosi, 2000; Rogers, 2003; Stoneman, 2001). Practical examples in the context of schools are, e.g., the change in teaching style, teaching facilities and teacher professionalization. Despite the attention to educational innovations and despite the attention to evidence-based education research, little is known on the relationship between school efficiency and innovations. Do innovations foster educational attainments if one accounts for budget constraints and characteristics of the school and the students?

This paper contributes to the literature in two ways. First, from an empirical point of view, it defines various categories of innovations and examines which of those innovation categories significantly correlate with school performance. School performance is measured as a ratio of student's educational achievements to school resources. School resources and innovations have never been considered simultaneously. Moreover, to allow for causal interpretations, previous literature paid significant attention to the influence of introducing single innovations, and largely ignored the (existing) mix of innovations. Illustrative are studies on the use of information technology in schools (e.g. Angrist and Lavy, 2002; Leuven et al., 2007); on smoothing the transition between schools (e.g. Valentine et al., 2009); or on teacher professionalization (e.g. Clotfelter et al., 2007; Miller et al., 2008; Rivkin et al., 2005). Studying the mix of innovations, compared to one single innovation at a time represents better the reality of the learning environment in the school, and with this, its effects on student performance.

Estimating the influence of innovations might be intricate to endogeneity issues. The endogeneity can arise from three sources. First, there might be omitted variable bias in that schools with more resources innovate more, but also produce and attract better students (Berg van der, 2008; Hedges et al., 1994; Vignoles et al., 2000). We capture this source of endogeneity by including the school budget in the efficiency estimation. Also parental background may influence both variables (Houtenville and Smith Conway, 2008; Preston et al., 2011). More involved parents may insist on educational innovations, and may simultaneously have higher ability children. By controlling for urbanisation and disadvantageous neighborhood, we account for this.

Endogeneity can also arise from measurement errors. This is unlikely in the current application. On the one hand, the data are reliable. The self-reported data on innovations are representative and have been carefully checked for measurement errors (see

* Corresponding author at: Maastricht University, TIER, Faculty of Humanities and Sciences, PO Box 616, 6200MD Maastricht, The Netherlands. Tel.: +31 433884458.

E-mail addresses: carla.haelermans@maastrichtuniversity.nl (C. Haelermans), k.dewitte@maastrichtuniversity.nl, Kristof.dewitte@econ.kuleuven.be (K. De Witte).

Haelermans (2010) for an extensive discussion). Furthermore, the non-innovation data are nationally collected by the Ministry of Education. On the other hand, the applied methodology mitigates the influence of a typical observations (e.g., arising from small and undetected measurement errors in the data). We therefore ignore this potential source of endogeneity.

Third, endogeneity can be due to reverse causality: innovative schools attract higher ability students. Previous literature indicated that this is not the case (Reynolds et al., 2000). It has been argued that not student ability but competition (Lubienski, 2003), teacher attitudes (Ghaith and Yaghi, 1997) and teacher beliefs (Hermans et al., 2008) can be considered as determinants of innovations.

As a second contribution, we measure school performance in a fully nonparametric framework, which implies that we do not impose any *a priori* specification on the functional form of the production technology. Previous literature indicated that information on the functional form is largely absent (Rothstein, 2010; Yatchew, 1998) leading to a specification bias. In particular, we estimate relative school performance by a model rooted in the popular Data Envelopment Analysis (DEA) literature (Charnes et al., 1978; Deprins et al., 1984). The model follows the robust order-m technique of Cazals et al. (2002) which mitigates the influence of outlying observations. This is convenient in the setting at hand, as some schools might be rather atypical (e.g., due to different student characteristics). To test the influence of innovations on school performance, the model is further adapted to a conditional efficiency framework. This allows us to include background variables (e.g., innovations, school size, region) in the efficiency estimations (Daraio and Simar, 2005, 2007; De Witte and Kortelainen, 2013). As a major advantage, the conditional efficiency model avoids the separability condition, which assumes that the background variables do not influence the input or output mix. Obviously, in the setting at hand, the use of innovations is expected (though not necessarily) to influence the educational attainments (i.e., the output of schools). Previous research often applies a two-stage model in which the background variables are regressed on the efficiency scores.[1] It is clear that in such a two-stage model, the innovations can only explain the efficiency score, but not influence it. As innovations are endogenous and, therefore, do influence efficiency, this is not desirable. The conditional efficiency model avoids this pitfall.

As a disadvantage, the conditional efficiency approach is descriptive and does not allow for causal interpretations. These would require (quasi-)experimental settings, which are barely exploited by efficiency tools. As a practical disadvantage, the computational burden to compute the bandwidths for the nonparametric model is large (see also De Witte and Kortelainen, 2013).

This paper focuses on innovation in Dutch secondary schools, for which we have a rich data set on educational attainments, school resources and student characteristics. From a representative questionnaire we set out to secondary schools, we distinguish five innovation clusters: (1) profiling, (2) pedagogic, (3) process, (4) teacher professionalization and (5) education chain innovations. The impact of each innovation cluster on school efficiency is examined. In doing so, we account for various differences between schools (e.g., size of the school, student teacher ratio, school type and region).

The remainder of this paper is structured as follows: Section 2 provides a literature review on the effect of innovations in education. Section 3 describes the empirical methodology and Section 4

explains the institutional setting and the data. Section 5 presents the results and relates them to the literature. The paper ends with some concluding remarks in Section 6.

2. Literature review on innovations in education

Schools innovate in various ways. Some innovations are visible to prospective parents. Illustrative are laptops for all students or a digital white board. Other innovations are invisible to outsiders. One can think of new didactical approaches or another system for teacher pay. There is an extensive body of literature on school innovations, but so far, most studies only cover one particular innovation, without taking into account the presence of the other innovations that the school might be working on. An advantage of studying only one innovation is that causal effects can be studied. However, ignoring other innovations might over- or underestimate this causal effect. This study includes all recent innovations at the school simultaneously. These innovations are clustered along five types: (1) profiling, (2) pedagogic, (3) process, (4) teacher professionalization and (5) education chain innovations. The clustering is similar to recent policy documents with respect to innovations in education (see e.g. Onderwijsraad, 2006b). This section discusses previous literature per cluster.

The profiling innovations cluster contains innovations with respect to curriculum changes (e.g., new courses) and the profiling of the school. Dutch secondary schools have a relatively large freedom in the design and organization of courses such that they can profile themselves towards a specific target group, like culture schools, sports schools or schools for high ability students. Several studies exist with respect to curriculum changes. A few examples of these come from Choudhury et al. (2008) and Niedermier et al. (2010). The latter observe that a curriculum change results in above-average and sustained improved performance by the students, whereas the first finds that a structural curriculum change decreases the length of program completion time. The results of Choudhury et al. (2008), however, prove to depend on gender, the school level and education track of the student. All in all, while the literature mostly argues positive results of curriculum changes with respect to student performance, the impact of school profiling is still unclear. To our best knowledge, this paper is the first to consider the influence of profiling on school performance.

The didactic cluster contains innovations that focus on the didactics of the courses: the way classes are taught and the use of specific pedagogical services (e.g., a dyslexia specialist). Increasing quality requirements force schools to increase this pedagogical and didactical focus. Pedagogical changes are, e.g., using different teaching methods (e.g., peer counseling, only group work or project based learning) or involving specialists like a remedial teacher or a speech specialist. There are only few previous studies regarding the effect of pedagogical changes on educational performance. The first is published by Nii and Chin (1996), who find that problem based learning leads to significantly higher grade point averages than using traditional didactic lecturing. Two more recent studies are by Fuchs and Woessmann (2007) and Queen (2009). The former observed that teachers experimenting with pedagogic approaches obtain higher student achievement with their students, while Queen shows that cooperative learning significantly benefits students more than traditional instruction. Overall, we can conclude that, although the number of studies on pedagogic changes in education is low, there seems to be a positive effect on educational performance.

The latter observation might be explained in two ways. First, one could argue that the implementation of a change suffices for student performance to increase, simply because of the change (so-called Hawthorne effect). If this is true it does not matter what

---

[1] Simar and Wilson (2010) extensively criticize the two-step method and present three arguments for this critique: First, the allocative efficiency estimates are artificially bounded by 1. Second, the efficiency estimates are correlated with inputs and outputs in a complex way. Third, there often is a systematic correlation across the efficiency estimates.

type of innovation a school invests in, as long as changes are implemented on a regular basis. Especially with respect to the pedagogic innovations, this might be the case. However, it does not matter whether the pedagogic change causes the positive change in student performance, or simply the fact that the teaching method and environment changed, as the argument about the positive effect of (pedagogic) innovations on school performance still holds. Second, and more likely than the existence of the Hawthorne effect in this context, the positive effects in the literature might be explained by 'publication bias' in which only studies with positive and significant results are published.

The cluster on process innovations consists of innovations which facilitate the learning process of students. Examples consist of IT changes, organizational changes at the school level and major building changes. Triggered by an increasingly digital world, schools introduce more extensive IT use, digital learning materials, laptops for all students or digital white boards (see for example Beauchamp, 2004). Furthermore, some schools are constructing a new building where they incorporate novel pedagogical strategies (e.g., by replacing all class rooms by group and discussion rooms). Even though schools make significant investments in process innovations, the evidence of the influence of IT use on educational performance is mixed and therefore, inconsistent results are found in the literature. Both positive, negative and zero effects are observed in previous literature. With respect to the latter, Goolsbee and Guryan (2006) studied the impact of home internet connections on student performance whereas Rouse and Krueger (2004) studied the relation between the use of a specific instructional computer program and student performance. Neither of these studies find a significant result. In contrast, a large part of the studies on IT use in education find positive effects. For example, Machin et al. (2007) use an instrumental variable approach to identify the causal impact of IT expenditures on performance of students. They observe a significant positive relationship. Punie et al. (2006) also find a positive relationship and conclude that there is evidence that IT use improves educational performance, although IT use at home counts for a major part of that relation. Lastly, Sosin et al. (2004) conclude that IT has a small but positive effect on the performance of students. A negative relationship has been observed by, for example, Leuven et al. (2007) who look at the effect of computer subsidies on performance. Furthermore, Angrist and Lavy (2002) also find a negative relation between IT and school performance.

Unfortunately, little is known on the effects of creating new school buildings. However, studies on group work, with computers and discussion tasks compared to traditional classroom teaching show positive effects (e.g. Sullivan and Pratt, 1996). Overall, the evidence on the effects of process innovations in education seems mixed.

The fourth innovation cluster contains all innovations that are related to teachers. Due to the teacher shortage in some western countries (European Trade Union Committee for Education, 2011), such as the Netherlands (Ministry of Education, 2011b) and Germany and the UK (e.g., King, 2009; Lipsett, 2008; Santiago, 2002), governments are attempting to increase the attractiveness of the teaching occupation. Schools add to this by introducing specific innovations for teachers. Previous literature considered the relation between an innovative approach for teacher pay, teacher quality or teacher training and student performance. These studies mainly find positive relations. For example, Atkinson et al. (2009), Kingdon and Teal (2007) and Lavy (2009) all conclude that a performance related pay scheme increases test scores and value added, although this does not necessarily hold for all types of schools. Furthermore, some studies use teacher professionalization innovations and relate these to school and student performance (e.g. Clotfelter et al., 2007; Miller et al., 2008; Rivkin et al., 2005). Of these studies, Rivkin et al. (2005) focus on the effect of teacher

quality, experience of teachers and the education of teachers on student performance and conclude that teachers have powerful effects on student performance, whereas Miller et al. (2008) study the effect of teacher absence and find that teacher absence significantly decreases the study achievement of students. Next, Clotfelter et al. (2007) find that teacher experience, teacher test scores and regular licensure of teachers have a positive effect on both reading and math scores of students. Croninger et al. (2007) find a positive effect of teachers degree type and experience on reading scores of students. Based on these studies, it seems that professionalization of the teacher positively contributes to educational performance.

Finally, education chain innovations denote, on the one hand, the relationship between the various levels of education (i.e., primary, secondary and tertiary education) and, on the other hand, the relationship with the community. The former consists of, e.g., primary school students who take introductory classes at the secondary school, students in graduation year of secondary education who take some lectures at university, and upper secondary education students that find an internship at the business that the school has arrangements with. The latter is represented by, for example, community schools and extracurricular activities. Literature indicates that extracurricular activities positively contribute to school performance and social aspects (e.g. Mahoney, 2000; Story et al., 2003). Studies on the effect of community schools are scarce and underdeveloped, but many studies exist on the relationship between the various levels of education and most of the above described relationships can be found in literature. With respect to the transition from primary to secondary education, Yadav (2010) studied the effect of mentoring on the transition to secondary school for at-risk children and finds positive changes in the output. For the transition from secondary to tertiary education, Bragg and Ruud (2007) studied the impact of specific career and technical education transition programs on student outcomes and find a positive effect of the transition program on a reading test, but no evidence was found for the math test. Lastly, a study by the US Department of Education (2010 a summary of findings from two literature reviews) concludes that most experiments in this field find positive effects but are carried out poorly. The US Department of Education calls for better evidence based research in this respect. Based on present literature, the relation between education chain innovations and educational performance seems to be positive.

Overall, we find that most studies on single innovations, corresponding with the innovations present in our innovation clusters, find positive effects. The only exception is the cluster of process innovations, where evidence from the literature is mixed.

3. Empirical methodology

Performance of schools is estimated against a frontier consisting of best practice observations. In this sense, performance estimation is a relative concept in line with seminal work of Farrell (1957). We apply an efficiency model which is based on the Free Disposal Hull (FDH) methodology (Deprins et al., 1984). The model is well-suited to the setting at hand because of three reasons. First, it is a fully nonparametric methodology which does not require any information on the production process. This is convenient as information on the relationship between the resources and the produced outputs is often unavailable to researchers (e.g. Yatchew, 1998). As parametric models assume a priori a functional form on this relationship, they might be wrongly specified which leads to biased estimation results (Hjalmarsson et al., 1996). Second, given its linear programming nature, the FDH model does not rely on price information (Deprins et al., 1984). This is convenient for

educational settings as information on output prices is unavailable (e.g., there is no price for educational attainments). Third, we use a recent extension of the traditional FDH model which mitigates the influence of outlying observations (e.g., arising from measurement errors or atypical observations; so-called 'robust FDH'; Cazals et al., 2002) and exogenous variables (heterogeneity across schools; so-called 'robust conditional FDH'; Daraio and Simar, 2005). The robust estimates have been shown to posses attractive properties: i.e., they are consistent (i.e., estimate the 'true' inefficiency) and have a fast rate of convergence (see Jeong et al., 2010). We briefly present the model in three steps: (*step 1*) the basic FDH model, (*step 2*) the robust FDH model and (*step 3*) the robust and conditional FDH model. For an in depth discussion, we refer to Daraio and Simar (2007) and Fried et al. (2008).

3.1. The Free Disposal Hull model

Consider a set χ of n schools, which is characterized at the level of the school by p heterogeneous and non-negative inputs x ($x_1 \ldots x_p$) and q heterogeneous and non-negative outputs y ($y_1 \ldots y_q$). The sample χ is then denoted by $\chi = \{(X_i, Y_i) = 1, \ldots, n\}$. The FDH model assumes that the input–output combinations are certainly feasible and that the inputs and outputs are freely disposable. Free disposability means that it should be possible to produce the output y also with more inputs and to produce less outputs with a given input set x. Formally: $\forall (x, y) \in \Psi$, if $x \geqslant x$ and $y \leqslant y$ then $(x, y) \in \Psi$ [where Ψ denotes the production technology set: $\Psi = \{(x, y) | x \in \Re_+^p, y \in \Re_+^q, (x, y)$ is feasible$\}$]. The best practice production set is defined as a free disposable hull of undominated input–output combinations:

$$\Psi_{FDH}\{(x,y) \in \Re_+^{p+q} | x \leqslant X_i, y \leqslant Y_i \in \chi\} \tag{1}$$

In the study at hand, we evaluate efficiency from an output-oriented perspective: with the given resources, what is the output shortfall for a school if it would produce as efficient as the observations on the best practice frontier? The output-oriented inefficiency estimates, $\lambda(x_o, y_o)$, measure the distance to the best practice frontier (see Fried et al., 2008, for further details):

$$\lambda(x_o, y_o) = \sup\{\lambda | (x_o, \lambda y_o) \in \Psi_{FDH}\} \tag{2}$$

As an efficient observation is located on the best practice frontier, it obtains an efficiency score λ equal to 1. An inefficient observation obtains an efficiency score λ higher than 1. The inefficiency ($1 - \lambda$) indicates the potential percentage increase in output if the observation would produce as efficient as its reference partner.

3.2. The robust FDH model

The FDH model in Eq. (2) is deterministic and may be problematic in presence of outlying observations as these heavily influence the best practice frontier. Outlying observations might arise because of measurement error or atypical observations (although we already argued in the introduction that the data are relatively clean from measurement errors, some small and undetected errors could still be present). Cazals et al. (2002) suggested to mitigate the impact of outlying observations in the FDH model, by drawing with replacement subsamples of size $m < n$ among those observations with fewer inputs than the evaluated observation (i.e., among those Y_i such that $x_o \geqslant X_i$). Cazals et al. (2002) have shown that the convergence rate of this order-m estimator is comparable to parametric estimators. Therefore, this estimator avoids the curse of dimensionality problem. Performance is assessed relative to this smaller sample. Following Daraio and Simar (2005), the partial sample size is determined as the value for which the number of super-efficient observations (i.e., $\lambda > 1$) is relatively constant. In

the setting at hand, m corresponds to 50, although alternative values delivered similar outcomes.

After repeating the sampling and efficiency evaluation B times, where B is sufficiently large (larger than 2000), the robust efficiency scores $\lambda^m(x_o, y_o)$ are obtained by taking the arithmetic average of the B inefficiencies.[2]

Thanks to the smaller sample size, an outlying observation will not constitute the reference sample in every draw. This will mitigate the impact of outlying observations. In case the evaluated observation (x_o, y_o) does not constitute its own reference set in every of the B drawings, the efficiency score λ^m will be larger than 1. This so-called super-efficiency indicates that the evaluated observation is performing better than the average m observations in its reference sample (Daraio and Simar, 2007).

3.3. The robust and conditional FDH model

The robust FDH scores can be easily adapted to include heterogeneity among schools (Daraio and Simar, 2005). Denote the exogenous variables, which can – at least in the short run – not be influenced by the school management, by z ($z_1 \ldots z_r$).[3] As a major benefit, the approach accounts for environmental factors in efficiency estimation without assuming the separability condition.

Daraio and Simar (2005) suggested to draw the subsamples of size m by a given probability, which is determined by a Kernel function around the continuous exogenous variables z. Observations (X, Y) with similar exogenous characteristics are drawn with a higher probability than observations which are less similar in z. Similar to before, the robust conditional FDH model draws B times the reference sample of size m with replacement, but now with a probability

$$\frac{K(z_o - Z_i)}{h} \bigg/ \sum_{j=1}^{n} \frac{K(z_o - Z_i)}{h}$$

among those Y_i such that $X_i \leqslant x_o$; where $K(\cdot)$ denotes a Kernel function and h the appropriate bandwidth (estimated by cross-validation) (for more information on bandwidth selection, see Badin et al., 2010). Finally, the B efficiency evaluations are averaged to obtain the robust conditional efficiency estimates $\lambda^m(x_o, y_o | z_o)$. The interpretation of the efficiency scores is similar to the robust FDH model. The convergence rate of the conditional estimator of Daraio and Simar (2005) depends on the dimension of Z, implying that the curse of dimensionality is not completely avoided but may exist for the continuous variables due to the smoothing in z.

De Witte and Kortelainen (2013) studied the possibility to also include discrete variables in the model, compared to only continuous environmental variables. They propose a standard multivariate product kernel for continuous, ordered discrete and unordered discrete variables, in order to smooth these mixed variables. Although the convergence rate of the conditional efficiency estimator depends on the number of environmental variables, nonparametric statistics and econometric theory tells us that the convergence rate of nonparametric estimators for conditional density and distribution functions involving mixed variables do not depend on the number of discrete variables but only on the number of continuous variables (Cazals et al., 2002). The conditional efficiency estimates

---

[2] Cazals et al. (2002) also suggested a perfectly equal integral formulation of this bootstrap. Given its computational efficiency, the R code underlying our analysis uses this integral formulation.

[3] Various alternative techniques to capture heterogeneity exist. In contrast to alternatives, the robust conditional efficiency model assumes that the exogenous variables Z directly influence the shape of the best practice frontier (i.e., the conditional FDH model does not assume a separability condition). Efficiency estimates are thus determined by both the inputs, outputs and exogenous variables (see Fried et al., 2008, for an extensive discussion).

allow us to examine the direction of the influence of the exogenous variation on school performance. In particular, the ratio of the conditional [i.e., accounting for heterogeneity; $\lambda^m (x_o, y_o | z_o)$] to the unconditional [i.e., ignoring heterogeneity; $\lambda^m (x_o, y_o)$] estimates can be (nonparametrically) regressed on the exogenous factor Z (Daraio and Simar, 2005, 2007). Daraio and Simar (2005) use a smooth nonparametric kernel regression to estimate the regression model. This approach allows one to detect positive, negative and neutral effects of the environmental factors on the production process. However, the marginal coefficient on the median is less meaningful, since we regress on a ratio. When Z is continuous and univariate, the visualization is straightforward, as one can use scatter plots of the ratio of conditional to unconditional efficiency scores against Z, and as a smoothed nonparametric regression curve can illustrate the effect of Z on the production process. In an output-oriented efficiency, a horizontal line implies no effect and an increasing (decreasing) smoothed regression curve shows that Z is favorable (unfavorable) to the production process. If Z is multivariate, one can use *partial regression plots* for the visualization of the effect. This means that only one environmental variable is allowed to change and other variables are kept at a fixed value.

Based on the work of Li and Racine (2007) and De Witte and Kortelainen (2013) present a nonparametric bootstrap procedure to obtain statistical inference on the direction of the influence. They propose a local linear regression estimation and use recently developed nonparametric tests and a nonparametric naïve bootstrap procedure to estimate the finite-sample distribution and a critical value of the nonparametric test statistics. Standard errors and p-values of the significance of the influence of Z on λ^m can be obtained. This model does not suffer from similar inference problems as two-stage models with the traditional and deterministic FDH and DEA models. We assess the influence of innovation on school performance by using this procedure.

4. Institutional setting and data

The empirical application exploits a dataset of 119 Dutch secondary education schools in 2007. These schools teach about 204,000 students, which is around 22% of the total amount of students in secondary education in 2007. The 119 schools comprise around 20% of the Dutch secondary schools. The data arise from the Ministry of Education which yearly collects data on expenses (on personnel and material use), size (number of students and personnel) and school characteristics (number of locations, school type and share of students from disadvantaged neighborhoods).[4]

Schools are facing a budget constraint that is based on the number of students, the education track and the student characteristics. The summary statistics, presented in Table 1, reveal that the expenses per student significantly vary among schools. Average expenses amount to 7150 euro per student in 2007, however variation among schools exists because schools have different types of students and receive different amounts of money for these types. Furthermore, school may decide to spend more or less in a certain year, compared to other schools. However, this figure ranges from about 5250 to about 16,400 euro per student. The expenses per student depict the financial possibilities of a school. It serves in the analysis as an input (denoted before by x).

Studying performance in Dutch secondary schools is attractive as there are two official performance measures for schools. Moreover, the measures are perfectly comparable across schools (they are measured by the education inspectorate). A first outcome is student achievement, which compares the educational track of a student in a given year with the education track predicted for a student at the end of primary education (i.e., the outcome of the ability tracking). If a student has never repeated a year and attends the same level in year 3 as was recommended in primary education, the student achievement of the first 3 years of secondary education is equal to 1.[5] Student achievement in upper secondary education is the percentage of students that has finished without repeating a year. We average the student achievement of the first 3 years and the student achievement of upper secondary education to one number. In a sense, this output variable corresponds to previously mentioned work of Choudhury et al. (2008).

Second, there is information on attainments (so-called 'national examination grade'). This is based on a nationwide test for all the subjects students undertake in the graduation year, for example, Dutch, English, Math, et cetera. Students write exams on about 8–10 subjects. The exams are graded in a double blind way such that neither the school nor the personal teacher can directly influence the outcome of the exams. The average national examination grade is measured on a 10 point scale, where a 5.5 suffices to pass the subject.

In the analysis, we use both the average national examination grades per school and the average student achievement per school as output variables. Table 1 indicates that the average national examination grade in 2007 was 6.4 and the average student achievement was 0.86. The student achievement is considerably below 1 because grade repetition is quite common in upper secondary education (Education Inspectorate, 2011). There is significant variation among schools.

Data on innovations are not centrally collected and have been brought together by a questionnaire. To collect information on innovations, an electronic questionnaire was sent to all secondary education schools in the Netherlands. The questionnaire consists of 132 listed innovations and some general questions, and is based on literature on educational innovations and interviews with school management. School directors were asked to fill out whether and when the school introduced a specific innovation. In this study, we use the innovation data of 2007. About one school in five filled out the questionnaire, which left us with a representative sample (see Haelermans, 2010, for an extensive discussion).

For a proper interpretation of the results, this study categorizes the individual innovations. The 132 individual innovations are divided over five clusters of innovations,[6] based on the content of the innovations. We explicitly did not choose to use, for example, a factor analysis because the manual division over the clusters makes it possible to completely take into account the content and nature of the individual innovations and interpret the value of the clusters in a straightforward way. We did also carry out a factor analysis, which gave similar results. These clusters are related to educational policy subjects that have been important over recent years. Examples of these are the focus of publications of the Education Council on teachers and teacher development (e.g. Onderwijsraad, 2006b; Onderwijsraad, 2007, 2009, 2011), on IT (Onderwijsraad, 2008); on better educational content (Onderwijsraad, 2006a); and on the transition between, for example, secondary and tertiary education (Onderwijsraad, 2005). Furthermore, there is the focus of the

---

[4] After removing some obvious measurement errors arising from questionnaire data, we think that it is empirically more sound to keep the sample of observations instead of reducing it by removing 'outlying' observations. Outlying detection procedures are to a large extent unstable as different outcomes are obtained for different procedures (see De Witte and Marques (2010) for an example and solution). To avoid existing differences among the observations, the robust order-m procedure is used. This mitigates the influence of the existing differences.

[5] Note that this value can exceed 1 if a student performs better than primary school recommendation. This might be the case if the student is enrolled in a higher education track than predicted.

[6] A full overview of the 132 single innovations and the clusters to which they belong is available upon request.

Table 1
Descriptive statistics 2007 (n = 119 schools).

	Average	St. dev	Min	Max
Inputs				
Expenses per student (in €)	7156.87	1610.07	5245.84	16415.15
Outputs				
Average national examination grades	6.41	0.21	5.94	7.13
Average student achievement	0.86	0.05	0.73	0.99
Operational environment				
Profiling innovations	5.23	2.81	0	12
Pedagogic innovations	14.6	4.94	2	26
Process innovations	12.87	4.08	3	23
Professionalization of the teacher innovations	7.2	2.87	0	14
Education chain innovations	4.39	2.15	0	10
Total number of students	1729	1032	265	5971
Locations	2.06	1.55	1	8
Schools per governing body	5.63	7.9	1	31
Degree of urbanization (in percent)	45	33	0	97
Student/teacher ratio	13.62	2.77	7.9	29.48
Percentage students from disadvantaged neighborhoods (in percent)	2.7	3.7	0	19.6

Ministry of Education on main topics such as innovation and IT in education (Ministry of Education, 2011a). This paper distinguishes the following five clusters of innovations. Note that the maximum number of innovations belonging to a cluster is mentioned in parentheses.

– New courses/profiling (29).
– Pedagogical/didactical (40).
– Process (35).
– Professionalization of teachers (16).
– Education chain (12).

Innovations in the new courses/profiling cluster are related to the introduction of new courses, such as new mathematics courses; to language innovations, such as bilingual education; and to the profile of school, for example a sports school or a culture school. Pedagogical/didactical innovations are innovations which are mainly associated with the content of the courses and the way classes are taught. The pedagogical/didactical innovations also include the services a school makes use of (e.g. psychologist, dyslexia specialist), projects and pupil innovations, such as pupils as coaches or development plans for pupils. Process innovations are, for example, IT, infrastructural (e.g. a totally new building) and organizational innovations. Innovations with respect to the professionalization of the teacher include all innovations that are related to the teaching staff. Lastly, education chain innovations mainly refer to innovations arising from collaborations between the secondary school and, for example, elementary schools, universities, or business companies.

The value of a cluster is determined by the total number of innovations present at a school belonging to that cluster in a certain year. The value in the next year is the number of newly introduced innovations during that year added to the value of the previous year. Table 1 shows that, on average, most innovations are pedagogic, followed by process innovations.

Schools mutually differ on various other aspects. In a sense, the innovations are even dependent on these background characteristics. In order to study the relationship between innovations and school performance, it is important to take into account the heterogeneity between schools. Including exogenous variables in the analysis also reduces the omitted variable bias, which could create a source of endogeneity.

Various sources of heterogeneity have been taken into account. First, school size (see for example Chakraborty et al., 2000; Flegg et al., 2004; Johnes, 2006), which is represented by the number of students, but also by the number of locations per school and the number of schools per governing body. Schools with several locations often have more students or offer more types of education, which they separate by location. The number of locations also reflects the geographical spread of a school over different sites. This might lead, on the one hand, to additional costs due to extra travel time (for teachers), or on the other hand, to scale economies (e.g., in management). A governing body is a coordinating organ, consisting of a group of managers that determines the 'school of thought' of its represented schools. This determination of the course can be very strict, up to the way of teaching, or can be very loose, in which the individual schools basically make their own decisions. The number of schools per governing body may, on the one hand, reflect scale economies thanks to participation in a schools' network and shared services; on the other hand, it captures the harms of a centralized bureaucracy. Literature on bureaucracy in education has shown that schools with more administrators and teachers have more bureaucracy (Marlow, 2001) and that bureaucracy has a negative influence (Bohte, 2001), but also that the relation between bureaucracy and school performance depends on the way performance is measured (e.g., test scores give a negative relation whereas attendance gives a positive relation) (Smith and Larimer, 2004). Table 1 shows that in 2007 Dutch schools had an average of two locations and belonged to a governing body with on average five schools.

Second, we control for region where the school is located (Ainsworth, 2002) and the degree of urbanization (Naper, 2010; Oliveira and Santos, 2005), which have been indicated to play a role in educational performance. The degree of urbanization is used as proxy for population density. It is measured by Statistics Netherlands as the percentage of large and very large urbanized squared kilometers within one area; it ranges from 0 till 97%. Following common practice in the Netherlands (Education Inspectorate, 2011), school location is proxied by the four large Dutch regions: the North, West ("Randstad"), South and East region of the Netherlands. Regions in the Netherlands are very different with respect to the amount of rural and urban areas, the population density and also the characteristics of the population, the labor market and the competition among schools. As region is a categorical variable, it is not presented in Table 1. The combination of including degree of urbanization and region corrects for different regional characteristics, competition among schools and population density. This allows us to correct for characteristics that the school cannot influence, but will influence the performance of the school.

Third, class size can influence student performance (Krueger, 2003; Lubienski et al., 2008; Nye et al., 2000). The student teacher ratio often serves a proxy for class size. The student teacher ratio

represents the number of students per teacher and is on average 13.6.

Finally, we account for the observation that disadvantaged students obtain lower educational attainments (Becker and Luthar, 2002; Gaziel, 1997) and for the existence of different school types (Lubienski et al., 2008; Silva Portela and Thanassoulis, 2001). Table 1 shows that the share of disadvantaged students in 2007 was on average 2.7%, but ranges from 0 to about 20%. The share of disadvantaged students represents students from a disadvantaged neighborhood, with high unemployment and low incomes. Unfortunately, we do not know the education or income level of the parents, which would have given us more specific information on the background of the students.

We observe eight different school types, which range from categorical schools for either the highest or the lowest level of education, to comprehensive schools that offer (almost) all types of education. In particular, there are four main school types (i.e., practical prevocational, theoretical prevocational, higher general and pre-academic education) and some combinations of them (e.g., practical prevocational and theoretical prevocational education together; both prevocational types and general higher education together). Schools with different school types (and thereby different levels of education) attract different types of students. Therefore, it is important to take the school type into account. As school type is a categorical variable, it is not presented in Table 1.

5. Results

The results are summarized in Table 2. The first column presents the efficiency results when we do not account for heterogeneity among schools (the so-called 'unconditional' efficiency) and only take into account the expenses per student and the two output variables. We observe an average efficiency score of 71.4%. In other words, with a given budget, Dutch secondary schools can increase their student performance by 28.6% if they would perform equally efficient as the best practices. However, there is large variation in the performance of secondary education schools, as can be seen from the sizeable standard deviation of 0.226 around this average efficiency. Furthermore, some schools have an efficiency score significantly larger than 1 (i.e. $\theta^m(x,y) < 1$) such that some observations can be viewed as super-efficient: they perform better than the average m observations in their reference sample.

Three alternative conditional efficiency models have been developed. These models include stepwise additional information on the innovation clusters and the heterogeneity variables. In Model 1, we examine the influence of the amount of innovations in all the five innovation clusters, by analyzing their correlation with the efficiency score via a nonparametric bootstrap procedure (see above). Model 2 adds school type, region, the percentage of disadvantaged students and the degree of urbanization to this baseline model, in order to check whether a different school type indeed leads to difference in school performance and to check the influence of region. The percentage disadvantaged students is a proxy for student quality and the degree of urbanization is a further control for regional difference, apart from the region variable. In model 3 we also include the student teacher ratio and the number of schools per governing body and number of locations.

Table 3 presents the influence of the innovations and heterogeneity variables on the efficiency scores. We present the median influence of these variables, rather than the mean, as the former is less influenced by extreme values. The median value is obtained from the nonparametric regression coefficient. The p-values are obtained from 500 bootstrap samples (see De Witte and Kortelainen, 2013). Note that, due to the structure of the nonparametric bootstrap, we only present whether the exogenous variable is significantly (un)favorable correlated with efficiency. The marginal coefficient on the median is less meaningful.[7]

Model 1 includes the five innovation clusters. The average efficiency score in model 1 does not dramatically change compared to the unconditional model. Average efficiency amounts to 77% with a large dispersion among schools. This indicates that only including information on innovations does not account for all the heterogeneity among schools. In model 1 all five innovation clusters are favorable and significant. The finding on profiling, pedagogic and education chain innovations is similar to findings from previous research using the same data set (Haelermans and Blank, 2012). However, our study is very different from the study of Haelermans and Blank (2012), because they use a parametric model and panel data, whereas we use a nonparametric model and 1 year of data. The results indicate a positive and significant relation between process innovations (IT-innovations) and school efficiency. As argued in the literature there is little consensus on the effect of IT-innovations in education. Our results confirm the positive effect. The positive findings on the influence of the pedagogic innovations and teacher professionalization innovations are in line with previous literature.

From model 2 onwards, the exogenous characteristics seem to capture all differences in initial inefficiency. In contrast to the previous model, which did not include disadvantaged students, urbanization, school type, and region, the average efficiency score no longer significantly deviates from 1. The standard deviation around the mean efficiency of 0.988 reduces significantly to 0.05. Both the efficiency score and the standard deviation indicate that a large part of the variation in inefficiency observed in the unconditional efficiency estimates is related to the innovations and to heterogeneity between schools. The school type is insignificantly related to efficiency. Furthermore, model 2 includes the share of disadvantaged students, the degree of urbanization and the region. The first two variables are unfavorable and insignificant, which corresponds to previous findings (Becker and Luthar, 2002; Naper, 2010). Urbanization has an insignificant sign, indicating that, controlled for the other sources of heterogeneity, schools located in urban or rural areas have about the same performance.

Finally, in addition to previous variables, model 3 includes the number of locations, the schools per governing body and the student teacher ratio. Similar to model 2, the variables absorb the inefficiency. In other words, the high efficiency score (0.99) and low standard deviation (0.01) imply that most heterogeneity between schools has been accounted for. The number of schools per governing body is significant and negatively related to school efficiency. As argued in the literature, the latter variable can capture scale economies and bureaucracy. Our finding is in line with the latter idea (e.g. Marlow, 2001), but not with the literature on scale (e.g. Chakraborty et al., 2000). In the model, the number of locations is significant but negatively related to school efficiency. Multiple locations mainly represent the scale of the school, as schools are only allowed to open an extra location if this is necessary considering the amount of

Table 2
Order-m efficiency scores.

	Unconditional (robust FDH)	Model 1	Model 2	Model 3
Average	0.714	0.770	0.988	0.997
St. dev.	0.227	0.225	0.050	0.012
Min	0.192	0.200	0.642	0.915
Max	1.088	1.000	1.000	1.000

---

[7] Whether a variable has a favorable or unfavorable influence relies on underlying nonparametric visualizations (see De Witte and Kortelainen, 2013), which are available upon request.

Table 3
Influence of innovation clusters and control variables on educational performance.

	Model 1	Model 2	Model 3
Innovation clusters			
Profiling innovations	Favorable (0.042)[**]	Favorable (0.106)	Favorable (<2e−16)[***]
Pedagogic innovations	Favorable (0.038)[**]	Favorable (<2e−16)[***]	Favorable (<2e−16)[***]
Process innovations	Favorable (0.032)[**]	Favorable (<2e−16)[***]	Favorable (0.002)[***]
Professionalization of the teacher innovations	Favorable (0.032)[**]	Favorable (<2e−16)[***]	Favorable (0.240)
Education chain innovations	Favorable (0.048)[**]	Favorable (<2e−16)[***]	Favorable (0.012)[**]
Control variables			
Number of locations per school			Unfavorable (0.008)[***]
Number of schools per governing body			Unfavorable (<2e−16)[***]
School type (eight categories)		(0.240)	(0.032)[**]
Region (four categories)		(0.004)[***]	(<2e−16)[***]
Percentage students from disadvantaged neighborhood		Unfavorable (0.432)	Unfavorable (1.000)
Degree of urbanization		Unfavorable (0.996)	Unfavorable (1.000)
Student/teacher ratio			Favorable (1.000)

Note: $n = 119$; bootstrapped p-values between brackets (500 bootstraps).
[*] Significance at 10% level.
[**] Significance at 5% level.
[***] Significance at 1% level.

students at the school. This contrasts previous literature on scale economies (Chakraborty et al., 2000; Johnes, 2006) as commonly positive scale effects are observed. Nevertheless, it is intuitive as an additional location may require teachers to travel from one location to the other on 1 day, leaving less time to actually teach, which might decrease the efficiency of the school (i.e., for a given budget, lower educational attainments are obtained). Next, we see that the student teacher ratio is negative but insignificant. The negative relationship implies that if the student teacher ratio is smaller (i.e., less students per teacher) the efficiency is higher. Although insignificant, this negative finding is in line with the literature (e.g. Nye et al., 2000). Finally, we observe that, by including these extra heterogeneity variables, the school type variable has become significant. This last result is also observed in the literature (e.g. Lubienski et al., 2008).

6. Conclusion and further research

This paper developed a tailored, fully nonparametric conditional efficiency model to study secondary school efficiency in the Netherlands. The specific setting of the model allows for studying the complex production process of education, without *a priori* information on the production function. The model is based on the work of Cazals et al. (2002), Daraio and Simar (2005, 2007) and De Witte and Kortelainen (2013). Nonparametric bootstrap-based significance tests have been applied to examine the statistical significance of the control variables.

In the nonparametric model, it is assumed that schools aim to maximize educational attainments of students under a given budget. We use data on the Netherlands, as educational attainments can be rigorously compared across schools and as school budget can be relatively freely allocated within the school. The efficiency scores are linked to a questionnaire on the innovation mix in schools. We distinguish five innovation clusters. The results indicate that innovation is positively related to efficiency. Profiling, pedagogic, process and education chain innovations are all significantly related to school efficiency. Furthermore, we find that teacher professionalization innovations are positively but not significantly related to school efficiency. The latter contrasts previous literature. Next, we find that school type and region are significantly related to efficiency. Our results show that efficiency scores are significantly different for the four different regions and for the different school types. Furthermore, the number of locations per school and the number of schools per governing body are significantly and negatively related to school efficiency.

The findings suggest that the effort of governments to stimulate innovations in education, and the effort of schools to implement these innovations, is not in vain. Innovations seem to positively influence school performance, taking into factors influencing performance. This positive relationship confirms the assumptions made on innovations in education and can be the basis for further development and stimulation of innovations in schools and research on this. Note that this paper does not allow for a causal interpretation of the results. Causal effects still need to be studied, but this paper provides some indications for future studies on the relationship between innovations and school performance, and the direction of this relationship.

Acknowledgements

References

Ainsworth, J.W., 2002. Why does it take a village? The mediation of neighborhood effects on educational achievement. Social Forces 81 (1), 117–152.
Angrist, J., Lavy, V., 2002. New evidence on classroom computers and pupil learning. Economic Journal 112, 735–765.
Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., et al., 2009. Evaluating the impact of performance-related pay for teachers in England. Labour Economics 16, 251–261.
Badin, L., Daraio, C., Simar, L., 2010. Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. European Journal of Operational Research 201 (1), 640–663.
Beauchamp, G., 2004. Teacher use of the interactive whiteboard in primary schools: towards an effective transition framework. Technology, Pedagogy and Education 13 (3), 327–348.
Becker, B.E., Luthar, S.S., 2002. Social-emotional factors affecting achievement outcomes among disadvantaged students: closing the achievement gap. Educational Psychologist 37 (4), 197–214.
Berg van der, S., 2008. How effective are poor schools? Poverty and educational outcomes in South Africa. Studies in Educational Evaluation 34, 145–154.
Bohte, J., 2001. School bureaucracy and student performance at the local level. Public Administration Review 61 (1), 92–99.
Bragg, D.D., Ruud, C.M., 2007. Career Pathways, Academic Performance, and Transition to College and Careers: The Impact of Two Select Career and Technical Education (CTE) Transition. University of Illenois, Champaign, IL.
Cazals, C., Florens, J.P., Simar, L., 2002. Nonparametric frontier estimation: a robust approach. Journal of Econometrics 106 (1), 1–25.
Chakraborty, K., Biswas, B., Lewis, W.C., 2000. Economies of scale in public education: an econometric analysis. Contemporary Economic Policy 18 (2), 238–247.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2 (6), 429–444.
Choudhury, A., Jones, J.R., Gamage, J., Ostaszewski, K., 2008. Structural change in the CPCU curriculum and its effect on the completion time. Academy of Educational Leadership Journal 12 (2).

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., 2007. Teacher credentials and student achievement: longitudinal analysis with student fixed effects. Economics of Education Review 26, 673–682.

Croninger, R.G., Rice, J.K., Rathburn, A., Nishio, M., 2007. Teacher qualification and early learning: effects of certification, degree and experience on first-grade student achievement. Economics of Education Review 26, 312–423.

Daraio, C., Simar, L., 2005. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. Journal of Productivity Analysis 24 (1), 93–121.

Daraio, C., Simar, L., 2007. Advanced robust and nonparametric methods in efficiency analysis. Methodology and applications. Springer, Heidelberg.

De Witte, K., Kortelainen, M., 2013. What explains performance of students in a heterogeneous environment? conditional efficiency estimation with continuous and discrete environmental variables. Applied Economics 45 (17), 2401–2412.

De Witte, K., Marques, R., 2010. Influential observations in frontier models, a robust non-oriented approach to the water sector. Annals of Operations Research 181 (1), 377–392.

Deprins, D., Simar, L., Tulkens, H., 1984. Measuring labor efficiency in post offices. In: Marchand, M., Pestieau, P., Tulkens, H. (Eds.), The Performance of Public Enterprises: Concept and Measurements. North-Holland, Amsterdam, pp. 243–267.

Dosi, G., 2000. The research on innovation diffusion: an assessment. In: Dosi, G. (Ed.), Innovation, Organization and Economic Dynamics. Edward Elgar Publishing Limited, Cheltenham, Glos, USA.

Education Inspectorate, 2011. De staat van het onderwijs – Onderwijsverslag 09/10. Inspectie van het Onderwijs, Utrecht.

European Trade Union Committee for Education, 2011. Teacher Shortage in Europe. European Trade Union Committee for Education, Brussels.

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

Flegg, A.T., Allen, D.O., Field, K., Thurlow, T.W., 2004. Measuring the efficiency of British Universities: a multi-period data envelopment analysis. Education Economics 12 (3), 231–249.

Fried, H.O., Lovell, C.A.K., Schmidt, S.S., 2008. The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press, New York.

Fuchs, T., Woessmann, L., 2007. What accounts for international differences in student performance? A re-examination using PISA data. Empirical Economics 32 (433–464).

Gaziel, H.H., 1997. Impact of school culture on effectiveness of secondary schools with disadvantaged students. The Journal of Educational Research 90 (5), 310–318.

Ghaith, G., Yaghi, H., 1997. Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. Teaching and Teacher Education 13 (4), 451–458.

Goolsbee, A., Guryan, J., 2006. The impact of internet subsidies in public schools. Review of Economics and Statistics 88 (2), 336–347.

Haelermans, C., 2010. Innovative power of Dutch secondary education. Innovation: Management, Policy and Practice 12 (2), 154–165.

Haelermans, C., Blank, J.L.T., 2012. Is a schools' performance related to educational change? – A study on the relation between innovations and secondary school productivity. Computers and Education 59 (3), 884–892.

Hedges, L.V., Laine, R.D., Greenwald, R., 1994. An exchange: Part I: Does money matter? A Meta-analysis of studies of the effects of differential school inputs on student outcomes. Educational Researcher 23 (3), 5–14.

Hermans, R., Tondeur, J., van Braak, J., Valcke, M., 2008. The impact of primary school teachers' educational beliefs and the classroom use of computers. Computers & Education 51, 1499–1509.

Hjalmarsson, L., Kumbhakar, S.C., Heshmati, A., 1996. DEA, DFA and SFA: a comparison. Journal of Productivity Analysis 7, 303–327.

Houtenville, A.J., Smith Conway, K., 2008. Parental effort, school resources and student achievement. The Journal of Human Resources XLIII (2), 437–453.

Jeong, S., Park, B., Simar, L., 2010. Nonparametric conditional efficiency measures: asymptotic properties. Annals of Operations Research 173 (1), 105–122.

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

King, N., 2009. Drastic Shortage of Teachers Feared in Germany. Deutsche Welle.

Kingdon, G.G., Teal, F., 2007. Does performance related pay for teachers improve student performance? Some evidence from India. Economics of Education Review 26, 473–486.

Krueger, A., 2003. Economic considerations and class size. The Economic Journal 113 (485), F34–F63.

Lavy, V., 2009. Performance pay and teachers' effort, productivity and grading ethics. American Economic Review 99 (5), 1979–2011.

Leuven, E., Lindahl, M., Oosterbeek, H., Webbink, D., 2007. The effect of extra funding for disadvantaged pupils on achievement. The Review of Economic and Statistics 89 (4), 721–736.

Li, Q., Racine, J.S., 2007. Nonparametric Economics: Theory and Practice. Princeton University Press.

Lipsett, A., 2008. Teacher Shortage in Key Subjects Feared. The Guardian.

Lubienski, C., 2003. Innovation in education markets: theory and evidence on the impact of competition and choice in charter schools. American Educational Research Journal 40 (2), 395–443.

Lubienski, S.T., Lubienski, C., Crane, C.C., 2008. Achievement differences and school type: the role of school climate, teacher certification and instruction. American Journal of Education 115, 97–138.

Machin, S., McNally, S., Silva, O., 2007. New technologies in schools: is there a payoff? The Economic Journal 117, 1145–1167.

Mahoney, J.L., 2000. School extracurricular activity participation as a moderator in the development of antisocial patterns. Child Development 71 (2), 502–516.

Marlow, M.L., 2001. Bureaucracy and student performance in US public schools. Applied Economics 33, 1341–1350.

Miller, R.T., Murnane, R.J., Willett, J.B., 2008. Do worker absences affect productivity? The case of teachers. International Labour Review 147 (1), 71–89.

Ministry of Education, 2011a. Innovatie en ICT in het onderwijs. <http://www.rijksoverheid.nl/onderwerpen/innovatie-en-ict-in-het-onderwijs> (retrieved 21.06.11).

Ministry of Education, 2011b. Nota werken in het onderwijs 2011. Ministry of Education, Science and Culture, Den Haag.

Naper, L.R., 2010. Teaching hiring practices and educational efficiency. Economics of Education Review 29, 658–668.

Niedermier, J., Way, D., Kasick, D., Kuperschmidt, R., 2010. Effect of curriculum change on exam performance in a 4-week psychiatry clerkship. Academic Psychiatry 34 (3), 216–219.

Nii, L.J., Chin, A., 1996. Comparative trial of problem-based learning versus didactic lectures on clerkship performance. American Journal of Pharmaceutical Education 60, 162–164.

Nye, B., Hedges, L.V., Konstantopoulos, S., 2000. The effects of small classes on academic achievement: the results of the Tennessee class size experiment. American Educational Research Journal 37 (1), 123–151.

Oliveira, M.A., Santos, C., 2005. Assessing school efficiency in Portugal using FDH and bootstrapping. Applied Economics 37, 957–968.

Onderwijsraad, 2005. Betere overgangen in het onderwijs. Onderwijsraad, Den Haag.

Onderwijsraad, 2006a. Versteviging van kennis in het onderwijs. Onderwijsraad, Den Haag.

Onderwijsraad, 2006b. Waardering voor het leraarschap. Onderwijsraad, Den Haag.

Onderwijsraad, 2007. Leraarschap is eigenaarschap. Onderwijsraad, Den Haag.

Onderwijsraad, 2008. Onderwijs en open leermiddelen. Onderwijsraad, Den Haag.

Onderwijsraad, 2009. Tien jaar adviezen over leraren en lerarenopleidingen: overzicht en balans. Onderwijsraad, Den Haag.

Onderwijsraad, 2011. Excellente leraren als voorbeeld. Onderwijsraad, Den Haag.

Preston, C., Goldring, E., Berends, M., Cannata, M., 2011. School innovation in district context: comparing traditional public schools and charter schools. Economics of Education Review. http://dx.doi.org/10.1016/j.econedurev.2011.07.016.

Punie, Y., Zinnbauer, D., Cabrera, M., 2006. A Review of the Impact of ICT on Learning. European Commission, Brussels.

Queen, S., 2009. Effect of Cooperative Learning and Traditional Strategies on Academic Performance in Middle School Language Arts. Walden University, Minneapolis, Minnesota.

Reynolds, D., Teddlie, C., Hopkins, D., Stringfiled, S., 2000. Linking school effectiveness and school improvement. In: Reynolds, D., Teddlie, C. (Eds.), The International Handbook of School Effectiveness Research. RefineCatch Limited, Suffolk.

Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools and academic achievement. Econometrica 72 (2), 417–458.

Rogers, E.M., 2003. Diffusion of Innovations, fifth ed. Free Press, New York.

Rothstein, J., 2010. Teacher quality in educational production: tracking, decay, and student achievement. The Quarterly Journal of Economics 125 (1), 175–214.

Rouse, C.E., Krueger, A.B., 2004. Putting computerized instruction to the test: a randomized evaluation of a 'scientifically-based' reading program. Economics of Education Review 23 (4), 323–338.

Santiago, P., 2002. Teacher Demand and Supply: Improving Teacher Quality and Addressing Teacher Shortages. OECD.

Silva Portela, M.C.A., Thanassoulis, E., 2001. Decomposing school and school-type efficiency. European Journal of Operational Research 132, 357–373.

Simar, L., Wilson, P.W., 2010. Inferences from cross-sectional, stochastic frontier models. Econometric Reviews 29 (1), 62–98.

Smith, K.B., Larimer, C.W., 2004. A mixed relationship: bureaucracy and school performance. Public Administration Review 64 (6), 27–36.

Sosin, K., Blecha, B.J., Agarwal, R., Bartlett, R.L., Daniel, J.I., 2004. Efficiency in the use of technology in economic education: some preliminary results. Recent Research in Economic Education 94 (2), 253–258.

Stoneman, P., 2001. The Economics of Technological Diffusion. Oxford.

Story, M., Sherwood, N.E., Himes, J.H., Davis, M., Jacobs, D.R.J., Cartwrigth, Y., et al., 2003. An after-school obesity prevention program for African–American girls: the Minnesota GEMS pilot study. Ethnicity and Disease 13 (1), 54–64.

Sullivan, N., Pratt, E., 1996. A comparative study of two ESL writing environment: a computer-assisted classroom and a traditional oral classroom. System 29 (4), 491–501.

US Department of Education, 2010. Postsecondary Education Transition: A Summary of the Findings from Two Literature Reviews. Office of Vocational and Adult Education, Jessup, MD.

Valentine, J.C., Hirschy, A.S., Bremer, C.D., Novillo, W., Castellano, M., Banister, A., 2009. Systematic Reviews of Research: Postsecondary Transitions – Identifying Effective Models and Practices. Louisville, KY.

Vignoles, A., Levacic, R., Walker, J., Machin, S., Reynolds, D., 2000. The Relationship between Resources Allocation and Pupil Attainment: A Review, London.

Yadav, V., 2010. Secondary school transition: does mentoring help 'at-risk' children? Community Practitioner 83 (4), 24–28.

Yatchew, A., 1998. Nonparametric regression techniques in economics. Journal of Economic Literature 36, 669–721.