On the productivity and efficiency of education : the role of innovations in Dutch secondary education

Citation for published version (APA):

Haelermans, C. M. G. (2012). *On the productivity and efficiency of education : the role of innovations in Dutch secondary education*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20120704ch>

Document status and date:

Published: 04/07/2012

DOI:

[10.26481/dis.20120704ch](https://doi.org/10.26481/dis.20120704ch)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 26 Apr. 2023

On the productivity and efficiency of education

The role of innovations in Dutch secondary education

© Carla Haelermans, Maastricht 2012.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission in writing, from the author.

ISBN 978-94-003-0040-8.

Cover design: Raadhuis voor creatieve communicatie, Alkmaar

Printed by: CPI Wöhrmann print service, the Netherlands

This book is no. I of the TIER Research Series, a PhD thesis series published by TIER.

ON THE PRODUCTIVITY AND EFFICIENCY OF EDUCATION

The role of innovations in Dutch secondary education

Proefschrift

ter verkrijging van de graad van doctor

aan de Universiteit Maastricht,

op gezag van de Rector Magnificus,

Prof mr. G.P.M.F. Mols

volgens het besluit van het College van Decanen,

in het openbaar te verdedigen

op woensdag 4 juli 2012 om 16.00 uur

door

Carla Maria Gertruda Haelermans

Promotores:

Prof. Dr. W.N.J. Groot

Prof. Dr. H. Maassen van den Brink

Copromotor:

Dr. J.L.T. Blank (Technische Universiteit Delft)

Beoordelingscommissie:

Prof. Dr. L. Borghans (voorzitter)

Prof. Dr. C.P. van Beers (Technische Universiteit Delft)

Prof. Dr. J. Dronkers

Prof. Dr. K.J. Hayes (Southern Methodist University, United States of America)

Prof. Dr. L.L.G. Soete

Acknowledgements

Writing a PhD-thesis is like riding a roller coaster. Upwards progress is made slowly and is usually followed by the joyful feeling of acceleration after reaching the top. Unfortunately, this feeling usually doesn't last long. Sometimes you feel as if the world is upside down, and sometimes everything is going by so fast that you can hardly remember it. Now and then you feel like you're in this ride alone, but often, you've got company.

In 2007, Jos invited me to the amusement park in Delft to try the roller coaster. A couple of years later, another amusement park opened in Maastricht and I decided to try the new roller coaster from Wim and Henriëtte over there as well. During both roller coaster rides, I have learned a lot and met many people. Right now, the roller coaster is slowing down, which gives me time to thank a few people, before the ride is finally over.

I am very grateful to Jos, who offered me the opportunity to start my PhD-research and introduced me to the world of productivity and efficiency research. I am also very thankful to Wim and Henriëtte for inviting me to TIER and for offering me the chance to dig deeper into education economics and finish my thesis at this amazing education economics institute. Without the supervision of these three people and the confidence they had in me, I would not have come this far.

I would also like to thank two people that I very pleasantly worked with during the past few years. Kristof, thank you for the great collaboration and the great papers that we produced together. Lex, thank you for everything that you have done for me: the supervision of my master thesis, writing and publishing a paper together, everything that I learned from you about being an academic, and the great advice that you always gave me.

Next, I want to thank my colleagues at IPSE Studies and at TIER, from whom I learned a lot. In Delft I had an amazing peer group that stimulated me to produce the best work possible. These people became good friends, and I would like to thank them personally. Casper, Emiel, Harald, Hester and Maartje, thank you for being the best peer group ever. Nienke and Stephanie, thank you for always having good advice and for the good times that we have together. In Maastricht, I would like to thank the other PhD candidates, Dimona, Erik, Iryna, Johan, Marieke and Sofie, with whom I am having very good times.

Thank you to my old roommates, Olga and Patrick in Delft, and Eline in Maastricht, and to my lovely friends, of whom I would like to mention Esther and Rinske specifically for keeping me company online. The past few years would not have been the same without you!

Last but not least, I want to thank my parents and sister for the faith they have in me and the support I always get from them. Lastly, Arjen, I would not have been able to finish this journey without you. Your unconditional support and love helped me through the bad times and made the good times even better than they already were. Thank you!

Carla Haelermans
Maastricht, May 2012

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Aims and Contributions of the Thesis	9
1.3	Thesis Outline	12
2	A Meta-Analysis on the Efficiency of Educational Institutes	17
2.1	Introduction	17
2.2	Previous Research and Reviews	21
2.3	Meta-Analysis – Descriptive Statistics and Simple Analysis	28
2.4	Meta-Analysis – The Extended Model and Publication Bias	35
2.5	Meta-Analysis – Meta-Regression Analysis	44
2.6	Conclusion	49

3 On the Allocation of Resources for Secondary Schools 53

3.1	Introduction	53
3.2	Dutch Secondary Education	57
3.3	Methodology	58
3.4	Data	62
3.5	Results	68
3.6	Conclusion	77

4. Diffusion and Determinants of Innovations 79

4.1	Introduction	79
4.2	Literature Review	80
4.3	Data	86
4.4	Methodology	93
4.5	Results	94
4.6	Conclusion	102

5 The Relationship between Innovations and School Productivity 103

5.1	Introduction	103
5.2	Literature Review	108
5.3	Data	113
5.4	Methodology	118
5.5	Results	123
5.6	Conclusion	131

6	The role of innovations in secondary school efficiency – A conditional efficiency model	133
6.1	Introduction	133
6.2	Methodology	137
6.3	Institutional Setting and Data	141
6.4	Results	145
6.5	Comparison of Chapter 5 and Chapter 6	150
6.6	Conclusion	152
7	Conclusions and Discussion	155
7.1	Conclusions	155
7.2	Policy Implications	161
7.3	Study Limitations and Further Research	164
	Bibliography	165
	Appendix A	183
	Appendix B	193
	Appendix C	199
	Summary	201
	Samenvatting	207
	Biography	215

List of Tables

Table 2-1 – Sources for the meta-analysis	32
Table 2-2 – Summary statistic, pooled average efficiency score, and test for fixed effects	34
Table 2-3 – Average efficiency scores by characteristics of the study and the population	37
Table 2-4 – Random-effects meta-regression	41
Table 2-5 – Egger’s test and Funnel graph’s Asymmetry Test (difference from 0 presented here)	44
Table 2-6 – Hedges’ simple meta-regression model	47
Table 2-7 – Hedges extended meta-regression model	48
Table 3-1 – Summary statistics (for 2007)	67
Table 3-2 – Technical efficiency scores of Dutch secondary schools 2002-2007 in percentages	69
Table 3-3 – Technical efficiency scores of Dutch secondary schools averaged over years 2007 in percentages	70
Table 3-4 – Misallocation of resources in percentages (data for 2007)	73
Table 3-5 – Median of misallocation of resources 2002-2007 in percentages	73
Table 3-6 – Estimation results for control variables in the model	76
Table 4-1 – Response among Dutch secondary schools	87
Table 4-2 – Representativeness of the sample – Logit analysis	88

Table 4-3 – Weighted sample descriptives of the determinants, 2007 (n=126)	92
Table 4-4 – Marginal effects of relation determinants and innovations, 2002-2007 (n=756)	101
Table 5-1 – Descriptive statistics, Dutch secondary education 2007 (n=126)	117
Table 5-2 – Model specifications (I-IV) for the nature of innovations	122
Table 5-3 –Log likelihoods and degrees of freedom of the various models (n=378)	123
Table 5-4 – Relevant parameter estimates: model with output biased technical change	126
Table 5-5 – Relationship between innovations and student achievement	130
Table 6-1 – Descriptive statistics 2007 (n=119)	144
Table 6-2 – Order-m efficiency scores	146
Table 6-3 – Influence of environment on educational performance	147
Table 7-1 – Overview of efficiency scores in the whole thesis	157

List of Figures

Figure 1-1 – Thesis outline	13
Figure 2-1 – Effect size vs. standard error of the effect size	43
Figure 3-1 – The predicted difference between the inflow and outflow of teachers	54
Figure 3-2 – The production frontier with and without environmental variables	62
Figure 3-3 – Over/Underutilization of teachers (in percentages)	74
Figure 3-4 – Productivity development between 2002 and 2007	77
Figure 4-1 – Diffusion of innovations (Rogers, 2003)	83
Figure 4-2 – Percentage diffusion of innovation clusters, 2002 – 2009 (n=153)	95
Figure 5-1 – Productivity development between 2005 and 2007	129
Figure 6-1 – Partial regression plots of the five innovation clusters in Model 3	149

1 Introduction

“The foundation of every state is the education of its youth”

- Diogenes Laertius -

1.1 Introduction

Productivity of education is a topic of intense debate among politicians, teachers, and trade unions and employers in education. Until recently, there was little political discussion on productivity in education. However, the EU's Lisbon goal for its Member States to become competitive knowledge economies has put it high on the political agenda of most countries. In the face of the current economic crisis and the austerity measures and the budget cuts that come with that, this goal has become even more of a challenge than before. This challenge is to improve educational output with the same amount of money or less. This puts the productivity of education and the interventions to improve this productivity on the agenda.

Productivity is the main source of countries' economic growth and the progress they have made during the past centuries. Over the past few hundred years, there have been numerous changes in production processes and the way people work. This has led to an increase in (labour) productivity, and this is the main reason for our current

prosperity. Productivity drives our economic development; therefore, it is important not only to focus on interventions that improve productivity, but also to measure and monitor the development of productivity and productivity growth.

Changing productivity is often achieved by rigorously changing processes and, in many cases, the key to changes in productivity is innovation. The educational sector has also been focusing on innovation over recent years, but these innovations have rarely focused on productivity. In education, the motivation to change is often pedagogical or practical, and hardly ever involves economic and organizational issues such as productivity. Taking into account the concept of educational productivity is just not part of the mind-set of the educational world.

Productivity is a sensitive issue among teachers, trade unions, and employers in the educational sector. Teachers are not production workers, they say, and every attempt to introduce measures of productivity is seen as questioning the quality of their work. Opponents of measuring educational productivity also argue that increasing productivity in education is equal to increasing work pressure and decreasing educational quality.

However, the health care sector is a perfect example to demonstrate that increasing productivity and maintaining, or even increasing, quality at the same time is perfectly possible. Such an insight is needed in education as well.

In addition to that, private sector productivity is rising, and, in order to avoid or limit Baumol's disease¹, the public sector, and education in particular, has to keep up with this.

In education, it is necessary both to increase productivity and to measure it in a relevant way. As responsibility, accountability, and transparency are more and more becoming the norm in education, it is important to gain insight into educational productivity, first of all, in order to fulfil these requirements, but also because schools are evaluated with the indicators that are used to measure productivity.

¹ For the education sector, Baumol's disease is the tendency of the (salary) costs to rise while labour productivity and student performance lag behind.

Schools need to start acting on the basis of the information they obtain from productivity indicators in order to be prepared for these assessments. Furthermore, as we want the resources invested in education to be well spent, it is important to operate as efficiently as possible, and to generate the highest possible educational output with the given budget. To do this we need to obtain insight into the productivity of education.

Concluding, productivity has become an important issue in education, and insight into the productivity of education provides tools for limiting Baumol's disease, and for realizing accountability and transparency.

1.1.1 Productivity, efficiency and innovations in education

Productivity is defined as the ratio of outputs to inputs. In education, both inputs and outputs of productivity can be measured in several ways. Depending on the choice of inputs and outputs, different types of productivity can be measured. Different measures could, for example, apply to teachers' productivity (Schalock et al., 1993), productivity of schools (Hoxby, 2002, 2003; 2004), or productivity of education in general. To avoid Baumol's disease, the productivity of teachers is the measure that should be used. However, if one is interested in the yields of government educational investments, productivity of education in general is a better measure. In other words, different measures of productivity should be used that relate to the question concerned.

There are several options to measure educational inputs. Educational inputs can, for example, be measured as total expenditure per school, expenditure per student, investments in education by the government, available computers per student, number of teachers, student teacher ratio, teacher quality, et cetera. It is important to choose the educational input that best covers the educational sector. Expenditure per student is a good measure of input, because both the financial aspect and the number of students are covered by it. There is a trade-off, however, because when expenditure per student is used, no information is included about the number of teachers and the number of managers employed. If it is possible to

include multiple inputs, it would be better to use the expenditure per personnel group per student as input. In practice, the choice of educational inputs is data-driven, and financial information is not always available. That is why some studies use, for example, student teacher ratios or the number of teachers (or personnel) per student as input.

Ideally, educational outputs represent educational performance and quality of education. Performance can refer either to student performance per se, such as test results, or to dropout rates, grade repetition, graduation numbers, and even plain student numbers. Educational quality is a difficult output; as it is not exactly clear what quality is. Often, educational quality is measured by student performance. However, usually, quality includes not only student performance in terms of figures but also the learning process, the school environment, the atmosphere at the school, etc. Many of these aspects are very hard, if not impossible, to measure on a comparable scale.

The different inputs and outputs that are used define the focus and, in many cases, also the quality of the productivity research. Therefore, productivity measures cannot easily be compared.

Productivity in education is often measured over time, and productivity figures mainly present the development of productivity over the years. Productivity compared with other schools can also give interesting insights. Productivity relative to others is referred to as efficiency. Efficiency of education is defined as productivity of education compared with the best-performing school.

Efficiency of education can refer to technical efficiency, in which the total inputs of the school are compared with the outputs. Technical inefficiency arises from failure to provide maximum services given the available resources (Fried et al., 2008). Efficiency can also refer to allocative efficiency, which studies the mix of resources and their prices compared with outputs. Allocative inefficiency arises “from resources being applied in the wrong mix, given their prices” (Fried et al., 2008, p. 5). Knowledge about allocative efficiency gives insight into the mix of

resources that a school has, and the share of each type of resource with respect to expenditure. Schools and government often do not have any knowledge about the optimal allocation of these resources within a school and between schools.

In order to improve the productivity and efficiency of education, many schools are implementing innovations. In this thesis, the concept of innovation is defined as: *'an idea, process or product that is new for a particular organization when it is introduced by that organization'* (based on Dosi (2000), Rogers (2003), and Stoneman (2001)). This definition is in accordance with the innovation questionnaire² that was sent out to gather data for the productivity and efficiency research in this thesis. Innovations in this context are, for example, new courses, language innovations, profiling of the school, didactic innovations, the use of specialists (e.g. dyslexia specialist), IT innovations, new school buildings, teacher professionalization innovations, and collaborations with primary education and higher education.

Although many schools invest in innovations, the effect of innovations on the productivity and efficiency of schools remains unclear. Therefore, more research is needed on the effects of innovations on productivity and efficiency. Studying innovations in relation to the productivity and efficiency of education is important for two reasons.

Firstly, large investments are being made in innovations in education, but scientific research has yet to prove the effectiveness of these innovations in. Secondly, innovation is a potential factor to improve the productivity and efficiency of schools. Studying innovations and productivity contributes to the discussion on how to limit or even avoid Baumol's disease, and shows how output can be increased without increasing personnel resources as well.

² For more information, see Chapter 4.

1.1.2. The literature on productivity, efficiency and innovations in education

Productivity

Studies on productivity have shown that productivity in education has decreased over the years. Reports on the development of productivity in the public sector in the Netherlands are published by the Netherlands Institute for Social Research (2007, 2010, 2012). These publications show that, in the Netherlands, educational productivity has decreased over the years. International studies on productivity development in education also show downward trends (e.g. Barber & Mourhsed, 2007; Hoxby, 2004; Leigh & Ryan, 2008).

Studies on productivity in education differ widely with respect to the inputs and outputs they have used for productivity. Nonetheless, all these studies find decreasing educational productivity. According to these publications, the decreasing productivity is mainly due to rising salaries in other sectors (whereas salaries in education were stable or had even decreased), the lack of real competition and market power, and the lack of a pay-for-performance system. Other explanations for decreasing productivity in education are economic developments, IT and innovation, organizational changes, class size, and non-financial terms of employment (Groot & Maassen van den Brink, 2011; Ojen van et al., 2011; SBO, 2011).

Efficiency

Similar to the productivity studies, the studies on efficiency in education differ widely, with respect to both content and methodology. The content differs mostly in the types of inputs and outputs used in the measurement of productivity. With respect to outputs, some studies use only the number of students of certain years (e.g. Ouellette & Vierstraete, 2005), whereas other studies use the number of students and the subsequent student performance (e.g. Blank et al., 2007b). There are also studies that use only the student performance or changes in student performance as an output measure (e.g. Afonso & St. Aubyn, 2006; Conroy & Arguea, 2008; Millimet & Collier, 2008). Studies that use, amongst other outputs, student performance - which can be independently compared and cannot be

influenced by the schools – as an output measure, are preferable to studies that use, for example, only student numbers as output. One reason is that in most countries schools are evaluated based on student performance and not on the number of students. Another reason is that such an independent performance measure is comparable between schools and cannot be influenced by them.

With respect to inputs, studies use costs of personnel or materials (e.g. Grosskopf et al., 1997), costs per student (e.g. Chakraborty et al., 2000; Ruggiero, 2007), or total costs of a school (e.g. Kempkes & Pohl, 2008). Other studies do not use costs as inputs, but use, for example, the number of teachers (e.g. Barbetta & Turati, 2003), teacher characteristics (such as experience: Conroy & Arguea, 2008)), school characteristics, or student characteristics. Studies that use costs combined with the number of students and number of teachers as inputs are preferable to studies that use other inputs.

A reason for the preference for including costs is that, in many countries, schools receive a lump-sum payment and are evaluated on the basis of their performance. Another reason is that a combined input of both costs and number of students or teachers provides more information than using only costs.

Besides differences in content, there are methodological differences between efficiency studies. These methodological differences between efficiency studies do not necessarily influence the quality of the studies, but do influence the outcomes and interpretation of the results. The most important methodological distinction is the difference between studies that use parametric methods, such as stochastic frontier analysis (e.g. Agasisti & Johnes, 2010; De Witte et al., 2010; Grosskopf et al., 2009), and studies that use nonparametric methods, such as Data Envelopment Analysis or Free Disposal Hull (e.g. Agasisti & Johnes, 2009; Tyagi et al., 2009). There is also a methodological difference between studies that use a cost function or input-oriented analysis (e.g. Denaux, 2009; Kempkes & Pohl, 2010), and studies that use a production function or output-oriented analysis (e.g. Agasisti & Perez-Esparrells, 2010; Flegg & Allen, 2007).

Many studies on the efficiency of education use a two-stage model to study the influence of certain factors on the productivity and/or the efficiency of a school. These studies find that: public ownership and charter schools lead to higher efficiency (e.g. Barbetta & Turati, 2003; Grosskopf et al., 2009); competition fosters productivity (e.g. Millimet & Collier, 2008; Ni, 2009); and scale and market structure also play a major role in the efficiency and productivity of schools (e.g. Barnett et al., 2002; Chakraborty et al., 2000).

Despite the importance of knowledge of allocative efficiency, there are only few studies that focus on allocative efficiency. Frequently, information on the prices of educational services is not available. Most studies on allocative efficiency investigate higher education. These studies often focus on allocative efficiency with respect to allocating time and money over teaching outputs and research outputs (Johnes & Johnes, 2009; Tauer et al., 2007). A few other studies on primary and secondary education look at the allocative efficiency of outputs. For instance, Wenger (2000) examines the trade-offs between test scores and graduation rates.

Studies that examine other educational sectors often do not analyse the allocative efficiency of resources (inputs). There are a few studies that look at the allocative efficiency of resources in other sectors than higher education: Banker et al. (2004), Grosskopf et al. (1997, 2001), and Blank et al. (2011). The first three studies focus on allocative efficiency in school districts in the US. School districts in the US operate at a different administrative level than single primary or secondary schools in Europe. Therefore, they differ significantly from studies using European data. The study by Blank et al. (2011) is the only one that looks at secondary education and that uses input prices to calculate misallocation scores.

Innovations

Studies on innovation in education mostly focus on the effect of a single individual innovation on outputs, e.g. student performance. For example, there are some quantitative studies on individual innovations such as the impact of IT use in

education (e.g. Angrist & Lavy, 2002; Leuven et al., 2007) and studies that relate teacher professionalization and teacher quality to school performance and student performance (e.g. Clotfelter et al., 2007; Miller et al., 2008; Rivkin et al., 2005). There are also some studies that use case studies to evaluate individual innovations, such as Reid (2003), who evaluated an absenteeism scheme and Van Grinsven and Tillema (2006) used a case study to examine the relationship between student self-regulation and learning effects. Some studies focus on innovations at an aggregate level. However, these studies are often qualitative analyses (Kirschner et al., 2004; Lubinski, 2003; Orozco-Gómez, 2006), and do not consider the effects of these innovations on educational performance. Rather, they describe the types and numbers of innovations that are being implemented in education.

Previous research on innovations in education has hardly studied the effect of innovation in general on student or school performance. The studies that look at multiple innovations simultaneously or at all the innovations implemented in the school are mainly of a qualitative nature or case studies. Most studies on innovations in education do not search for quantitative evidence on the importance or effects of innovations at the aggregate (school) level, nor do they relate innovations economically to productivity.

1.2 Aims and Contributions of the Thesis

The following research gaps are identified: 1) the empirical literature on allocative efficiency in education is underdeveloped, particularly for secondary education; 2) quantitative data on innovations in (secondary) education is lacking; and 3) the empirical literature on factors explaining differences in productivity and efficiency lacks information on the role of innovations (aggregated at the school level).

In order to fill these research gaps, the aim of this thesis is threefold: to develop a theoretical base; to use innovative empirical analyses; and to use a unique data set. *First*, this thesis develops a ‘theoretical’ base for the analyses by providing a

literature overview of efficiency and productivity studies in education. The aim is not only to give a descriptive literature review, but also to perform a quantitative meta-analysis on the efficiency of educational institutes. *Second*, this thesis adds to the existing literature on productivity and efficiency of education by empirically estimating the technical efficiency and misallocation of Dutch secondary education schools, using innovative methods, both parametric and nonparametric. *Third*, this thesis gathers and analyses unique data on innovations (aggregated) at the school level and includes these data in the models as described in the second aim.

Scientific contribution

The scientific contribution of this thesis lies, first of all, in the inclusion of specific innovation information on innovations in the productivity and efficiency framework of education. As argued in the previous section, there is hardly any quantitative research on the effects of multiple innovations. Most studies look at one particular innovation. Therefore, information is gathered on several individual innovations, which are then aggregated and grouped into five clusters in order to include them in the productivity model and analyse them.

Second, the scientific contribution lies in the rich data and specific methodology used for the application to Dutch secondary education. The data and methodology make it possible to estimate the technical efficiency, the misallocation, and the conditional efficiency of secondary education in the Netherlands.

As there is an ongoing discussion on the choice of parametric versus nonparametric methods in efficiency research, this thesis includes both methods and compares their results. The starting point is a parametric output distance function. The advantage of parametric models is, for example, that they can handle noise and endogeneity in the data. Parametric models converge easily and have a low risk of the dimensionality problem. Lastly, parametric models make it possible to study statistical inference quite easily. The disadvantage of parametric models is that the functional form has to be assumed a priori. In this thesis, two chapters, Chapters 3

and 5, use the parametric production function and one of these chapters extends the base model by including specific information on innovations.

The nonparametric Free Disposal Hull (FDH) is used to estimate efficiency conditional on innovations and other variables. The FDH is used to compare the results on innovations using both parametric and nonparametric methods. The main scientific contribution of this thesis is the inclusion of innovations at the aggregate level and the use of different innovative (non)parametric methods.

Societal contribution

The societal contribution of this thesis is that the results of the analysis on the effects of innovations in secondary education can be used to determine whether it is desirable to further disseminate these innovations. As argued before, innovation in education is being stimulated by both governments and the schools themselves. However, little empirical research exists on the relationship between innovations in education and the performance of students and schools. In this thesis a unique data set of individual innovations (grouped into clusters based on content) implemented in Dutch secondary schools is collected and aggregated at the school level. This makes it possible to provide new information with respect to the innovativeness of education, specifically for Dutch secondary education at the (aggregate) school level.

Innovations are not considered separately; all innovations present at the school are taken into account. This thesis also provides information on the determinants of innovations, and insight into the relationship between innovations and educational performance in Dutch secondary education.

The societal contribution is also found in estimating the technical efficiency and misallocation of resources, as these give insight into how the secondary education sector is performing as a whole and give information on the diversity between schools with respect to efficiency scores. This is specifically important because government budgets are under pressure, and this research provides information on how to generate the same quality with the same or less money.

Lastly, the societal contribution is the policy advice on innovations and school efficiency that can be given, on the basis of this research, to both the national and the local authorities, and to school management.

For the national and local authorities, the interesting aspects of this thesis are the determinants of innovations and the relationship of innovations to school performance. Unfortunately, this thesis cannot provide information on the causality of the effect of innovations on school performance, but it can provide insight into the direction and magnitude of the relationship between innovations and school performance. For school management, the interesting aspects lie in: the efficiency gains for schools; the potential gains by allocating the money differently over the resources; and the potential efficiency gains if other school and environmental factors change.

1.3 Thesis Outline

The thesis consists of four parts (see Figure 1-1). The current Chapter 1, which constitutes part I of this thesis, introduces the thesis as a whole, and describes its aims and contributions, and the basic outline of the thesis.

Figure 1-1 – Thesis outline

Part II contains two chapters on the efficiency of education. *Chapter 2* is a meta-analysis on the efficiency of education. The first part describes the literature on efficiency in education and indicates the differences between the several studies that have been published. Then, the second (quantitative) part of the meta-analysis is performed and publication bias is addressed. The meta-regression, corrected for publication bias, is used to determine the factors that influence the magnitude of the efficiency score in education. Some of these aspects are analysed further in Parts II and III of this thesis. In this meta-analysis, primary, secondary, and tertiary education are taken into account. In the remainder of the thesis the focus is on secondary education.

Chapter 3 is an applied study comparable to the individual studies used for the meta-analysis, but, in addition, it adds a misallocation analysis. In Chapter 3, technical efficiency scores and the misallocation of Dutch secondary education are estimated, using a parametric indirect output distance function. The efficiency scores are estimated by including in the model resources, prices and budgets, types of

personnel, material, pupil performance, and environmental characteristics. In view of the predicted teacher shortage in the Netherlands, the emphasis in this chapter is placed on the misallocation scores. The insights gained from the misallocation scores may contribute to the discussion on how to decrease the teacher shortage.

Part III contains the chapters that incorporate innovations in the education productivity model. First, in *Chapter 4*, the diffusion and determinants of innovations in education are discussed. Innovations are divided into five clusters, and the diffusion path of the innovation clusters is described. Then the determinants are regressed on the share of applied innovations in a certain cluster. Finally, policy implications are derived from the findings.

In *Chapter 5*, the relationship between innovations and school productivity is studied. The prevalence of innovations is explained by the determinants found in Chapter 4, and the five innovation clusters enter the production function in the next step. The productivity of the school is assessed using a budget-restricted indirect output distance function. Compared with the model in Chapter 3, this model is extended with innovations as explanatory variables, represented by the five innovation clusters. In Chapter 5, the nature of innovations (neutral, input-oriented, output-oriented, or input/output-oriented) is examined by comparing all possible models. Finally, the relationship between the innovation clusters and productivity is studied using the best-performing model.

In *Chapter 6* the focus is on the relationship between the five innovation clusters and educational performance. A nonparametric conditional efficiency approach is used to estimate the relative performance of schools when conditioning on innovations and student, school, and environmental characteristics. The specific setting of the model allows both continuous and discrete variables to be included in the conditional efficiency framework. Therefore, it is possible to include all types of variables in the model at the same time, which makes the model more realistic. Nonparametric bootstrap-based significance tests are used to test for the statistically significant relationship between the innovation clusters and other variables and the

educational performance. Lastly, this chapter discusses and compares the methods and findings on innovations in both Chapters 5 and 6.

Finally, in *Chapter 7 (part IV)*, the conclusions of this thesis are presented. Concluding statements on the combined findings of the individual chapters are presented, as well as an overall conclusion. In the second half of this chapter, the policy implications of the findings are discussed and possible future research directions are indicated.

2 A Meta-Analysis on the Efficiency of Educational Institutes³

“Much education today is monumentally ineffective. All too often we are giving young people cut flowers when we should be teaching them to grow their own plants”

- John W. Gardner -

2.1 Introduction

Previous research has indicated that educational performance is not only important to individuals with respect to their further career and wages (e.g. Ashenfelter et al., 1999; Becker, 1964; Card, 1999; Mincer, 1958) but also relates to the other activities of individuals and to the performance of a country. For example, it is known that educational performance is negatively related to crime (e.g. Groot & Maassen van den Brink, 2010; Lochner & Moretti, 2004), and positively related to health (e.g. Conti et al., 2010; Silles, 2009). Other studies have indicated that, on the national level, educational performance increases economic growth and competitiveness (e.g. Barro, 1991; Hanushek & Woessmann, 2009; Krueger & Lindahl, 2001) and the

³ This chapter is based on C. Haelermans (2011). A meta-analysis on the efficiency of educational institutes. *TIER working paper*.

productivity of individuals and firms (e.g. Büchel, 2002; Moretti, 2004). Although the here mentioned studies represent only a very small part of all economic research on education, they do show the importance of education.

Because of the high importance of educational performance, it is important to find a way to value and measure this performance. In order to value educational performance, many studies compare schools and countries and rank these on a regular basis. Examples of such rankings are the three-yearly international PISA (Program for International Student Assessment) comparisons (OECD) and the four-yearly TIMSS study (Trends in International Mathematics and Science study) (National Centre for Education Statistics) which compare the performance of countries in terms of learning outcomes. Further examples are the MBA program comparison of Hsu, Chao and James (2009), or the yearly rankings of *The Economist* (e.g. 2010), the *Financial Times* (e.g. 2010) and the *Shanghai ranking* (2011). In scientific research such comparisons can be made by using efficiency analyses, among other types of analysis. There are many types of efficiency analyses (e.g. parametric vs. non-parametric methods), which have different units of observations, such as countries, schools or students, and include different types of variables. All these efficiency studies have different characteristics, so one cannot simply compare them.

Efficiency studies have evolved over the years, relying increasingly on more sophisticated econometric methods, a variety of methodologies, more complex models that include more years of data, different types and amounts of inputs and outputs, and more variables to calculate the efficiency scores which are used to compare the educational performance of schools and/or students. Hence, the characteristics of these studies have also changed and have become more diverse.

The diversity of efficiency studies first of all lies in the choice of estimation method: either a parametric. Semi-parametric or a non-parametric method may be chosen. Other differences between studies arise from the choice of the unit of observation and which variables to use as inputs and outputs. With respect to inputs there is a difference between studies that use both quantities and prices of

2. A Meta-Analysis on the Efficiency of Educational Institutes

resources, such as different types of personnel and materials, and studies that use either one of these. For outputs, there is a difference between studies that incorporate some kind of quality control (e.g. class size or student performance) in their output measure and studies that only use student numbers (e.g. Barbetta & Turati, 2003; Johnes, 2006; 2011). Other differences may arise from the type of data (panel data compared with cross-sectional data or only one year of data), the characteristics of the educational system, and the educational sector (e.g. primary, secondary, or higher education) being compared.

These differences between the methods employed by studies lead to differences in the quality of the respective studies, as some only cover part of the output of an educational institute (for example, only the number of students), whereas others include many outputs (e.g. also student performance) which gives a better representation of reality. There are differences in the way in which quality of education is incorporated in the analysis. Quality is a major problem in comparing efficiency studies and their estimates with one another. Another issue arises from the fact that there are not many overview studies. What is needed is a good overview of the studies on the efficiency of education institutes. To the best of our knowledge, no such quantitative review exists. Furthermore, it is necessary to find a way to compare efficiency studies with one another at the same level in order to discuss the 'average' efficiency score. It is of vital importance for the interpretation of efficiency research to know how the different study characteristics and authors' choices influence the magnitude of the efficiency rates. This is the particular contribution of this chapter.

By providing a literature overview on efficiency studies in education, and by carrying out a meta-analysis on the efficiency scores, the intention is to fill the research gap. Furthermore, this chapter provides insights in how to compare educational efficiency studies. Lastly, this chapter contributes in not only analyzing factors that are known to influence the efficiency score, e.g. parametric vs. non-parametric or the returns to scale, but also analyze and explain differences in factors

that are of economic interest, such as the country, data period and the unit of observation.

The literature review discusses studies on efficiency, but also considers the studies that do not meet the criteria for being included in the meta-analysis (e.g. studies that do not present efficiency scores or standard deviations, or studies that only present scale efficiency). Using meta-analysis has several advantages. Meta-analyses take into account the robustness of the individual estimates; make it possible to include meta-information on the efficiency analysis of education; and identify differences between estimates. Meta-analyses can also help to find an overall structure among those papers that are individually hampered by imprecise measurements. However, the danger of this is that – due to the process of editing and refereeing – publication bias occurs. The meta-analysis can deal with this as well, by explicitly taking into account publication bias. A meta-analysis makes it possible to compare the different parametric and non-parametric approaches described above, and define how the choice of a specific approach affects the estimates. Additionally, in the meta-analysis one can also distinguish between, for example, different types of data or different circumstances that are used by individual studies. In sum, meta-analysis is more structured than a literature review and seems a good approach to overcome the comparability problem between studies, as the meta-information and type of analysis differ between these studies.

In Section 2.2, this chapter first explains the concept of efficiency, and then focuses on previous research and reviews related to the efficiency of education (primary, secondary and higher education). Next, the differences in these studies that make efficiency studies hard to compare are discussed. In Section 2.3, the selection of the studies, the database, and the descriptive statistics are discussed. Section 2.4 presents the results of the extended model, including a simple meta-regression analysis, and discusses the issue of publication bias. This is followed in Section 2.5 by a short explanation of Hedges' model and the application of the meta-regression analysis corrected for publication bias. Section 2.6 provides a conclusion and discussion.

2.2 Previous Research and Reviews

Meta-analysis

Meta-analysis was first used in the mid-1970s, mostly in the field of health-related studies (Lipsey & Wilson, 2001), and has been gaining popularity in economics since the beginning of the 21st century. Although the methodology has changed and improved, the basic concept of the analysis: namely, the calculation of the effect size and the weighting based on significance and sample size, has not changed. Currently, its use is not as extensive as in health research, yet several economics studies have used meta-analyses to overview the literature.

Several papers, such as those written by Pereira and Martins (2004), Ashenfelter et al. (1999), Groot and Maassen van den Brink (2000) and Haelermans and Borghans (2012), have used a meta-analysis to investigate the returns to education or (on-the-job) training. As far as we are aware, no meta-analysis exists on the efficiency of education. There are a few meta-analyses on the efficiency of, for example, farmers (Phillips, 1994), urban public transport (Brons et al., 2005), and agriculture in developing countries (Thiam et al., 2001). However, none of these studies corrects for publication bias.

The concept of efficiency measurement

The theoretical foundation of efficiency analysis lies in production theory. In production theory, production, revenue or profit is maximized or costs are minimized by decision-making units (Fried et al., 2008). In production theory inputs and outputs are incorporated in the model, which are, for example, used to determine the optimal production. In efficiency research, decision-making units are benchmarked based on their current production compared with the optimal production or compared with the best performing decision-making unit. Efficiency measurement makes a comparison between the actual performance of a decision-making unit and the optimal performance that is possible for this unit. Optimal performance is defined by a frontier which represents combinations of inputs and

outputs that give optimal results. As, in most cases, the optimal situation is unknown, an approximation of this frontier is needed, which is often called the best-practice frontier. A best-practice frontier represents the maximization of production, revenue or profits, or the minimization of costs (Fried et al., 2008).

In recent years, efficiency studies have become more and more common in education research. As efficiency research is not always used to produce efficiency scores, it is also known under the name of ‘productivity research’. Many studies use the productivity/efficiency framework to analyze productivity (e.g. Simpson, 2009), the influence of certain factors on productivity (e.g. Naper, 2010), or to test different versions of a model (e.g. Baker, 2001). Another group of efficiency studies computes and presents efficiency scores, such as technical efficiency scores, scale efficiency scores, or allocative efficiency scores. Technical efficiency scores represent the efficiency of the institute as a whole, based on inputs and outputs, i.e. an institute is considered technically efficient if output cannot be increased without increasing at least one input or decreasing at least one other output (Fried et al., 2008, ch. 1.). Scale efficiency scores, in turn, represent the deviation from the optimal scale for that institute. Allocative efficiency is the efficiency with respect to the share of resources. Allocative efficiency, for example, answers the question whether a school needs more or less managers or teachers. The majority of the studies on efficiency present only technical efficiency scores and only a small share of these efficiency computing studies use the two-step approach to try to explain the differences in the efficiency of the units of observation by their characteristics⁴.

Most of the studies that analyze efficiency scores estimate and present technical efficiency scores. In these studies a frontier is estimated, hence educational institutes on the frontier are considered best-practice institutes, and get an efficiency score of 1. The other educational institutes lie below the frontier, and get

⁴ Note the critique by Simar and Wilson (2007) on the two-step procedure. They mention the following problems with these procedures: 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.

2. A Meta-Analysis on the Efficiency of Educational Institutes

efficiency scores relative to the best-performing institutes. This is, for example, a percentage between 0 and 1 (where a value of 0.75 means an efficiency score of 75 per cent, and thereby an inefficiency score of 25 per cent). The magnitude of technical efficiency scores depends on several choices made by the author concerned which will be discussed in Section 2.4. The present study only focuses on the technical efficiency scores themselves and the variety in efficiency between institutes. Average efficiency scores vary from quite low, e.g. 0.35, (Barbetta & Turati, 2003) to very high, e.g. 0.99 (see Hsu et al., 2009). Of course, these studies have different characteristics which influence this low or high score.

Efficiency research in education

Of the published efficiency studies, some only present the average efficiency scores (e.g. Abbott & Doucouliagos, 2009; Borge & Naper, 2006; Ouellette & Vierstraete, 2005), while others also present the standard deviation, and often also the minimum and maximum (e.g. Agasisti & Salerno, 2007; Flegg et al., 2004; Grosskopf et al., 2009; McMillan & Chan, 2006), and a third group presents all the efficiency scores for all the individual units of observation (see, for example, Chakraborty et al., 2001; Cordero-Ferrera et al., 2008; Kempkes & Pohl, 2010; Tauer et al., 2007). Some of the authors make a comparison between two groups of institutions and draw a conclusion related to the difference in efficiency scores between these groups (e.g. Barbetta & Turati, 2003; Grosskopf et al., 2009). However, most of the efficiency studies look at different units of observation, and conclude that efficiency scores vary greatly among the units of observation (Agasisti & Salerno, 2007; Cokgezen, 2009).

Some studies also present scale efficiency scores. These scores represent whether schools could operate more efficiently if they would either increase or decrease scale. In most cases, these efficiency scores are presented next to the technical efficiency scores for comparison (e.g. Abbott & Doucouliagos, 2002; Agasisti & Perez-Esparrells, 2010; Hsu et al., 2009; Hu et al., 2009; Johnes, 1995). However, in a few cases the study fully focuses on scale efficiency (Glass & McKillop,

1995). Due to the low number of studies that include scale efficiency, this meta-analysis will focus on technical efficiency scores only. Some other studies analyze not only technical efficiency or scale efficiency but also allocative efficiency. Most of these studies on the allocation of resources conclude that a better allocation of resources will increase productivity (e.g. Conroy & Arguea, 2008; Grosskopf et al., 1997; Johnes & Johnes, 2009; Soares de Mello et al., 2006; Tauer et al., 2007). Some studies show that a different allocation of resources does not necessarily lead to higher productivity (Blank et al., 2007a; Blank et al., 2011). As explained earlier, in this meta-analysis, the focus is only on technical efficiency scores. However, in order to interpret the meta-analysis results, it was necessary to explain the broader context of efficiency analysis.

Differences in efficiency research in education

The main difference between efficiency studies is the choice of method and the choice of the inputs and outputs, with the latter being the most important for the quality of the study. In earlier studies, efficiency models mainly used one output and often only one input (see Coelli et al., 2005; Kumbhakar & Lovell, 2000). However, over the years, as efficiency analysis became more developed and was increasingly applied in education, researchers started to include multiple inputs and outputs.

Efficiency studies in primary education (PE) and secondary education (SE) mainly use the number of students and student performance as outputs in the efficiency model, while some of these studies only use the number of students of certain years (e.g. Barbetta & Turati, 2003; Ouellette & Vierstraete, 2005; Smet & Nonneman, 1998). Other studies use the number of students and the subsequent student performance (Blank et al., 2007a, 2007b). Lastly, there are studies that only use student performance or change in student performance as an output measure (e.g. Afonso & St. Aubyn, 2006; Blank et al., 2007b; Conroy & Arguea, 2008; Dijkgraaf et al., 2008; Grosskopf et al., 2009; Millimet & Collier, 2008). However, the inputs used in efficiency studies on PE and SE are diverse. Some of the studies use costs as input, which can be costs of personnel or materials (see Blank et al., 2007a, 2007b;

2. A Meta-Analysis on the Efficiency of Educational Institutes

Grosskopf et al., 1997; Ouellette & Vierstraete, 2005), costs per student (e.g. Chakraborty et al., 2000; Hoxby, 2004; Ruggiero, 2007), or total costs of a school (Izadi et al., 2002; Kempkes & Pohl, 2008; Smet & Nonneman, 1998). Other studies do not use costs as inputs, but, for example, number of teachers (e.g. Barbetta & Turati, 2003), teacher characteristics, such as experience (Conroy & Arguea, 2008), school characteristics, or student characteristics. Most of the studies described above use data from the US or the UK. However, some efficiency studies use data on, for example, the Netherlands, Italy or Australia and New Zealand.

Efficiency studies in higher education (HE) mainly use two types of outputs: first of all, outputs that are related to the teaching activities of higher education; and, second, outputs that are related to research activities, such as publishing and applying for research grants. Furthermore, some studies use only teaching outputs (e.g. Abbott & Doucouliagos, 2009; Agasisti & Salerno, 2007; Stevens, 2005) or only research outputs (among others, Adams et al., 2005; Cherchye & Vanden Abeele, 2005; Groot & García-Valderrama, 2006). Other studies use a combination of the two, such as Flegg et al. (2004), Johnes, Johnes, Thanassoulis, Lenton and Emrouznejad (2005) and Worthington and Lee (2008). Specifications of these different outputs are, for instance, scientific publications, number of undergraduate, graduate or PhD students, income from research, and income from grants. Studies that combine these two types of outputs often focus on the existence of economies of scope. Some studies deviate from the standard outputs, as described above. These studies use, for example, quality and reputation of the higher education institutes as output (Dawson et al., 2009). With respect to inputs, there is also a large distinction between studies that include costs (e.g. Athanassopoulos & Shale, 1997; Johnes, 2006, 2008; Worthington & Lee, 2008) and studies that do not (e.g. Adams et al., 2005; Ferrari & Laureti, 2005). Besides costs, studies use number of students and personnel as inputs (e.g. NG & Li, 2009; Soares de Mello et al., 2006). Lastly, some studies use inputs such as IT knowledge (Sosin et al., 2004). There are also studies that look at the educational sector as a whole. These studies mostly use student numbers as outputs (see, for example, O'Mahony & Stevens, 2009; Waldo, 2007). In

higher education, inputs and outputs vary as well, but not as much as in primary and secondary education.

To conclude, the inputs and outputs used in studies on both primary/secondary education and higher education are very diverse. Many of the inputs and outputs reflect indicators that are often referred to as ‘quality indicators’, for example, student performance, student teacher ratio, and the number of students (OECD, 2010). Including these indicators may lead to a better representation of reality in the efficiency study, and might lead to a larger spread among the efficiency scores, as more information on the differences between educational institutes is taken into account.

There are several studies that take into account the background characteristics of schools or students. These background characteristics are, for example, family and municipality information (Wenger, 2000), or urbanization (e.g. Blank et al., 2007a, 2007b; Foreman-Peck & Foreman-Peck, 2006). Although other studies may also take into account background characteristics, they often do not explicitly mention them in the description of the analysis and the variables used. The majority of the studies do not include background characteristics in their analysis, which is a major disadvantage of these studies, as the best practice and thereby the efficiency scores may not reflect reality. Concluding, studies differ in the amount of background characteristics they take into account (if at all), which influences the magnitude of the efficiency scores. Taking into account all possible background characteristics may lead to a larger range in efficiency scores as schools become more different from each other if more variables are included in the analysis.

Finally, it is important to choose the method of analysis. One could choose a parametric estimation method (for example, Stochastic Frontier Analysis (SFA)), or a non-parametric estimation method (for example Data Envelopment Analysis (DEA), or Free Disposal Hull (FDH)). There is an ongoing debate on the advantages and disadvantages of both methods (Van den Broek et al., 1980). The main advantage of non-parametric methods is that there is no preset assumption on the functional form of the model (Coelli et al., 2005), whereas the main advantage of parametric

methods is that it is easy to deal with environmental characteristics, statistical tests can be conducted, and outliers can easily be identified (Coelli et al., 2005)⁵.

In general the two methods⁶ should tell the same story, as they both calculate technical efficiency scores by identifying a frontier and comparing the decision-making units to the ones that lie on the frontier, as the benchmark. Furthermore, they require the same data, and the basic method for both SFA and DEA requires data on both inputs and outputs, but not necessarily on prices (Coelli et al., 2005). In general it is true that the better the quality of the data, the more SFA and DEA generate the same results (Fried et al., 2008). However, one major condition is that the specified functional form in the SFA should be the correct one. When opting for non-parametric estimation, the next choice is whether to use constant returns to scale (CRS) or variable returns to scale (VRS), whereas for the parametric analysis it is necessary to choose a specific type of model, e.g. production model, costs model, or profit model. In non-parametric analysis, VRS is a subset of CRS and therefore, by definition, VRS cannot have lower efficiency scores than CRS. There are other characteristics that might lead to differences in efficiency scores, but authors are not always in control of these characteristics. These are, for example, the country and the data period. Another aspect that depends on the data is the unit of observation. Sometimes it is possible to choose a certain unit of observation, but in most cases the researcher is bound to the data that can be gathered or is already available. In their analysis, many authors estimate several efficiency scores, and then compare the scores of the groups. The general recognition of the importance of all these characteristics for the magnitude of the efficiency scores is underlined by comparison analyses, as the authors make a distinction between groups of schools or methods because they believe there is a difference between them. Comparisons are, for example, made between: SFA and DEA (e.g. McMillan & Chan, 2006); CRS

⁵ Stochastic Frontier Models may hamper from the “wrong skewness” problem as described in Simar and Wilson (2010). Unfortunately, it is not possible to take this into account for the meta-analysis. It might, however, downsize the results from parametric analysis, as models with very low or no inefficiency might be rejected.

⁶ There are other methods that are being used, such as COLS, MOLS and thick frontier. However, only the methods used most often are discussed here.

and VRS (Agasisti & Johnes, 2009; Tyagi et al., 2009); schools in rural vs. urban areas (e.g. Kantabutra & Tang, 2006); universities with and without medical faculties (e.g. Agasisti & Salerno, 2007); private, public, and charter schools (e.g. Barbetta & Turati, 2003; Grosskopf et al., 2009); and several input and output variable combinations (e.g. Tauer et al., 2007). These comparisons also emphasize the difference in magnitude of the average efficiency when using different methods or data. From this section on efficiency research in education, it can be concluded that efficiency scores vary substantially due to differences in the specification of the study. Some of these differences can be explained by the lack of perfect data of good quality (Lovell, 2000). However, other differences may exist because there are substantial differences between educational systems in various countries, differences between the wealth and educational level of different areas within countries, and differences in the type and quality of students.

2.3 Meta-Analysis – Descriptive Statistics and Simple Analysis

Database

The research field of efficiency is broad and a selection had to be made in the publications that were found for this chapter. All published papers were collected from the Economic Literature Index (Econlit), JSTOR, Web of Science, Interscience and Science Direct, and, additionally, unpublished articles were collected from the National Bureau of Economic Research (NBER), OECD, The World Bank, and from working paper series from several individual universities, found through Google Scholar. The keywords used during this search process are the following: “efficiency” (in the title), “efficiency of education”, “efficiency & education” and “efficiency & school”. In order to select studies, the following inclusion criteria were applied:

1. The study includes an empirical analysis about the efficiency of education.
2. The study computes technical efficiency scores, and presents these scores or the mean of the efficiency scores.

2. A Meta-Analysis on the Efficiency of Educational Institutes

3. The study reports the standard deviation from the mean, or it is possible to calculate this.
4. The total number of observations is given, or it is possible to calculate this number.
5. The study has educational institutes or departments as the unit of observation.
6. The study is written in English.

In general, the effect size is the main variable needed for a meta-analysis, and, in the specific case concerned, the effect size is the average efficiency score. In the remainder of this chapter the average efficiency score is used, where other meta-analyses would refer to the effect size. The formula for the average efficiency score (ES) and its standard error (SE) of a study included is:

$$ES_m = \bar{X} = \frac{\sum x_i}{n}, \quad (2-1)$$

$$SE_m = \frac{s}{\sqrt{n}}, \quad (2-2)$$

where $\bar{X}$ is the mean efficiency score; $\sum x_i$ is the sum of the individual efficiency scores of the units of observation; n is the number of observations; and s is the standard deviation of the mean efficiency score. For more information on the methodology of a meta-analysis see Lipsey and Wilson (2001).

The average efficiency score of the units of observation is that reported in the studies. Some studies report more than one efficiency score: for example, comparing SFA and DEA, estimating the efficiency scores for different years, or comparing different groups of students or schools. All the reported average efficiency scores from each study are included, given that the study was made before the set deadline for entries in January 2011. In the sensitivity analysis, the inclusion of multiple efficiency scores per study is controlled for. Due to the selection criteria, many studies were found to have used the above-mentioned keywords had to be

excluded. The excluded studies were not a loss because they did not report average efficiency scores. In total, 190 estimates from 48 studies were selected for this meta-analysis, based on the inclusion criteria⁷.

In the database, variables are created for the author of the study, the year of publication, the country to which the data set relates, the educational sector, the unit of observation, the estimation method the study has used (parametric or non-parametric), the used inputs and outputs, and the total number of observations. Furthermore, the database includes the average efficiency of education and the corresponding standard. The studies used for the analysis were published between 1990 and 2010, with the years of data used ranging from 1977 to 2005. The total number of observations varies from 11 to 2129. Most of the studies use DEA (35 studies), VRS (25 studies), and analyze the higher education sector (34 studies).

Table 2-1 presents information on the authors, year of publication, survey period, and the number of average efficiency scores per study. The number of average efficiency scores per study differs markedly because some studies report efficiency scores for more years or for different estimation methods.

Pooled average efficiency score

Table 2-2 shows the average efficiency scores for all included studies. As can be seen in Table 2-2, the mean efficiency score for all studies is 0.794 (on a scale between 0 and 1), with an average standard deviation of 0.159. This implies that, on average, schools could improve their production process by approximately 20 per cent. Furthermore, the fixed-effects pooled efficiency score is 0.938, and the random-effects pooled efficiency score is 0.799. There are several reasons to prefer the scores of the random-effects model over those of the other two. First of all, in the fixed-effects model the Q-statistic is very significant, implying heterogeneity in the average efficiency scores. Furthermore, the mean efficiency score size does not take into account the standard deviation from the mean and the number of observations per study, which gives a biased estimate. Therefore, the random-effects model is

⁷ The complete database and the corresponding coding scheme are available upon request.

2. A Meta-Analysis on the Efficiency of Educational Institutes

used. The random-effects model gives a smoother distribution of weights, and a parameter which is as significant as in the fixed-effects model. In the remainder of this chapter the random-effects average efficiency score is estimated and presented in the tables.

Table 2-1 – Sources for the meta-analysis

Author and year	No. of efficiency scores in study	Sector	Country	Method	Survey period
Abbott and Doucouliagos (2002)	2	HE	Australia	non-parametric	1995
Agasisti and Dal Bianco (2009)	6	HE	Italy	non-parametric	1998-2003
Agasisti and Johnes (2009b)	5	HE	USA	parametric	2003-2005
Agasisti and Johnes (2009a)	4	HE	England & Italy	non-parametric	2001-2004
Agasisti and Johnes (2010)	2	HE	Italy	parametric	2001-2003
Agasisti and Perez-Esparrells (2010)	4	HE	Italy & Spain	non-parametric	2000-2004
Agasisti and Salerno (2007)	8	HE	Italy	non-parametric	2002
Athanassopoulos and Shale (1997)	3	HE	UK	non-parametric	1992
Avkiran (2001)	3	HE	Australia	non-parametric	1994
Barbetta and Turati (2003)	12	PE/SE	Italy	both	1998
Bonesrønning and Rattsø (1994)	4	PE/SE	Norway	non-parametric	1990
Chakraborty et al. (2001)	2	PE/SE	USA	both	1992
Cherchye and Vanden Abeele (2005)	9	HE	Netherlands	non-parametric	1996-2000
Cokgezen (2009)	3	HE	Turkey	non-parametric	2003
Cordero-Ferrera et al. (2008)	2	PE/SE	Spain	non-parametric	2001
Denaux (2009)	2	PE/SE	USA	non-parametric	2005
Diamond and Medewitz (1990)	1	PE/SE	USA	non-parametric	1977
Essid et al. (2010)	3	PE/SE	Tunisia	non-parametric	2003
Flegg and Allen (2007)	10	HE	England	non-parametric	1994-2003
Flegg et al. (2004)	13	HE	UK	non-parametric	1980-1992
Fu and Huang (2006)	3	HE	Taiwan	non-parametric	2003
Grosskopf et al. (1997)	1	PE/SE	USA	parametric	1987-1989
Grosskopf et al. (2009)	6	PE/SE & HE	USA	parametric	2001
Hsu et al. (2009)	2	HE	USA	non-parametric	2006

Author and year	No. of efficiency scores in study	Sector	Country	Method	Survey period
Hu et al. (2009)	2	PE/SE	China	non-parametric	2003
Izadi et al. (2002)	1	HE	UK	parametric	1994
Johnes (1995)	3	HE	England	non-parametric	1992
Johnes (2006)	4	HE	England	non-parametric	1993
Johnes and Salas-Velasco (2007)	1	HE	Spain	parametric	1998
Johnes et al. (2005)	2	HE	England	both	2000-2002
Kantabutra and Tang (2006)	2	PE/SE	Thailand	non-parametric	2001-2003
Kantabutra and Tang (2010)	2	HE	Thailand	non-parametric	2005
Kempkes and Pohl (2010)	7	HE	Germany	both	1998-2003
Köksal and Nalcaci (2006)	2	HE	Turkey	non-parametric	1999-2000
Madden et al. (1997)	2	HE	Australia	parametric	1987-1991
Mancébon and Bandrés (1999)	1	PE/SE	Spain	non-parametric	1990
McMillan and Chan (2006)	2	HE	Canada	both	1992
McMillan and Datta (1998)	18	HE	Canada	non-parametric	1992
Ng and Li (2000)	9	HE	China	non-parametric	1993-1995
Oliveira and Santos (2005)	1	PE/SE	Portugal	non-parametric	1999-2000
Primont and Domazlicky (2006)	1	PE/SE	USA	non-parametric	2000, 2002
Ray and Jeon (2008)	1	HE	USA	non-parametric	1998-2004
Saito and McIntosh (2003)	2	PE/SE	USA	non-parametric	1994-1996
Soares de Mello et al. (2006)	1	HE	Brazil	non-parametric	1996-2000
Stevens (2005)	4	HE	England	parametric	1995-1998
Tajnikar and Debevec (2008)	2	HE	Slovenia	non-parametric	2004-2005
Tauer et al. (2007)	8	HE	USA	non-parametric	2004-2005
Tyagi et al. (2009)	2	HE	India	non-parametric	2003
Total	190				

Table 2-2 – Summary statistic, pooled average efficiency score, and test for fixed effects

<i>Summary Statistic</i>		
	Mean	Standard deviation
Mean efficiency score	0.794	0.159
<i>Pooled average efficiency score</i>		
	Efficiency score	Standard Error
Fixed effect	0.938	0.000
Random-effect	0.799	0.009
<i>Test for Fixed Effects</i>		
Q-statistic	51000	
p-value	0.000	
Between-studies variance	0.013	
N	190	

Sensitivity analysis

In order to test for the sensitivity of the meta-analysis results several sensitivity tests are performed⁸. These tests are carried out by excluding some of the studies. First the highest 5 and 10 per cent and the lowest 5 and 10 per cent of the studies with respect to the average efficiency score are excluded, thus removing all outliers. Next, the highest and lowest 5 and 10 per cent with respect to the significance level of the average efficiency score are excluded. Performing these analyses with the adapted data sets has hardly any effect on the random pooled average efficiency score. Then, the top and bottom 5 and 10 per cent of the studies with respect to sample size are excluded. This analysis also has a negligible effect on the random pooled average efficiency score. Lastly, only one average efficiency score per study is kept in the data set, as most authors present more than one efficiency score in order, for example, to make comparisons. Keeping only one efficiency score per study avoids interdependencies between the different efficiency scores presented in the

⁸ Results are available upon request.

individual studies. These interdependencies could make the meta-analysis biased towards certain characteristics which may influence the efficiency score and therefore may lead to higher or lower average efficiency scores. Keeping only one efficiency score per study reduces the meta-analysis to 48 observations. This also has hardly any effect on the magnitude of the pooled average efficiency score. In all the sensitivity analyses the random-effects pooled average efficiency score varies between 0.799 and 0.824. There are slightly larger changes in the case of excluding the top and bottom 10 per cent of the average efficiency score, significance, and number of observations than for the other groups. However, the changes are still very small, and there is no reason to exclude studies from the meta-analysis that have not been excluded already in the study-selection stage.

2.4 Meta-Analysis – The Extended Model and Publication Bias

There are several possibilities to analyze the effect of the study characteristics on the efficiency scores. First of all, one could estimate the extended model in which the total observations are divided over several groups, according to their study characteristics. A second option is to use a meta-regression analysis. However, both these estimates might be hampered by publication bias. Therefore, first the extended model is presented, and then the meta-regression analysis, after which it is tested for publication bias.

The extended model

Table 2-3 shows the estimated pooled average efficiency score and its standard error. It was not possible to calculate the significance of the difference between the groups for two reasons. First, there was considerable heterogeneity in both of the subgroups. Second, when estimating the random-effects efficiency scores it is not possible to calculate the significance of the between-study difference. As it is not possible to calculate significant differences in this paragraph the differences in mean

efficiency are only described, whereas the remainder of this chapter also interprets and explains differences in efficiency scores.

Table 2-3 indicates that studies on primary and secondary education have much lower efficiency rates than studies on higher education. The two studies that use data from North America and Europe have higher efficiency rates than studies from other continents. Studies that use non-parametric methods have higher efficiency scores than studies using parametric methods, and studies using variable returns to scale have higher efficiency rates than studies using constant returns to scale and parametric studies. There is hardly any difference in the efficiency scores of studies from before 2000 and after 2000. Lastly, studies that use educational departments as the unit of observation have lower efficiency rates than studies that use schools, school districts or study programs as the unit of observation. Study programs have the highest efficiency, followed by school districts and schools.

Table 2-3 is very informative with respect to the different efficiency scores for studies that use different inputs and outputs. Studies using publications and/or research grants as an output have lower efficiency scores than studies using other outputs, whereas studies using a combination of teacher and student numbers and publications and/or research grants have higher efficiency scores than the other categories. With respect to inputs, studies using costs as inputs have higher efficiency scores than the other categories. Studies using student performance, the number of students, or number of employed staff (mostly in full-time equivalents) have lower efficiency scores.

Table 2-3 – Average efficiency scores by characteristics of the study and the population

	N	Random average efficiency score	z- value	Standard Error
Total	190	0.799	91.24	0.009
Sector				
Primary/Secondary Education	40	0.746	35.68	0.021
Higher Education	150	0.814	83.36	0.010
Continent				
North-America	51	0.869	61.13	0.014
Europe	103	0.773	41.39	0.019
Other	36	0.763	33.93	0.022
Method				
Non-parametric	155	0.815	93.79	0.009
Parametric	35	0.73	27.63	0.026
Returns to scale				
VRS	93	0.839	81.1	0.010
CRS	38	0.796	28.35	0.028
Not stated	59	0.738	35.27	0.021
Outputs				
Students and research output	88	0.856	105.7	0.008
Student performance	70	0.759	48.42	0.016
Teacher and research output	8	0.798	18.99	0.042
Expenditures	10	0.764	15.67	0.049
Publications/research grants	13	0.679	14.84	0.046
Inputs				
Costs	28	0.877	124.8	0.007
Number of staff	47	0.703	29.27	0.024
Costs and number of staff	62	0.857	75.58	0.011

Costs and number of students	4	0.972	111.3	0.009
Number of students	21	0.763	44.9	0.017
Number of staff and students	17	0.842	29.27	0.029
Student performance	11	0.734	13	0.056
Unit of observation				
School	150	0.812	86.73	0.009
Department	31	0.734	31.85	0.023
School district	3	0.834	10.52	0.079
Study program	6	0.881	49.34	0.018
Data period				
Before 2000	115	0.807	76.48	0.011
From 2000 on	75	0.788	55.09	0.014

Random-effects meta-regression

Table 2-4 gives the results of the random-effects meta-regression analysis. The meta-regression analysis is to be preferred over the models presented above because of the possibility of analyzing the significance of the between-group differences. In a meta-regression all the explanatory variables that are created during the coding stage can be regressed on the main variable, in this case the pooled average efficiency score. A standard meta-regression is a random-effects regression in which the relation between the background characteristics and other variables and the pooled average efficiency score can be tested quantitatively. One of the advantages of a meta-regression is its multivariate context (Doucouliagos & Laroche, 2003): a meta-regression makes it possible to investigate the relation between the average efficiency score and one or more of the key variables, where the other variables, the study characteristics, can be used as controls, as presented in Table 2-4.

The results of the meta-regression analysis are comparable to those presented in the extended model. Table 2-4 presents the results of the meta-regression, and shows that most of the study and sample characteristics are of significant influence on the mean efficiency scores. With respect to the results of the

2. A Meta-Analysis on the Efficiency of Educational Institutes

differences in estimation method we find the result which could be expected based on the theory. Studies using parametric methods have significantly lower efficiency scores compared with studies using non-parametric methods. Many studies on productivity research have already concluded that a different method of estimation leads to different efficiency scores. Furthermore, parametric efficiency scores are very often lower than non-parametric efficiency scores (Fried et al., 2008). Moreover, Table 2-4 shows that the type of returns to scale being used does not significantly influence the magnitude of the efficiency score.

With respect to the variables which might be of interest to education economics, we first of all find that earlier studies have significantly lower efficiency scores than recent studies. This might be because earlier studies very often only use one input and/or output, and as Table 2-4 also shows, an increasing number of inputs and/or outputs leads to higher efficiency scores. Another reason might be that in the early years researchers were open to any results, whereas in recent years researchers may not trust their results if efficiency scores are very low, because the 'common' efficiency score findings are higher. Therefore, it is possible that these researchers reject their result, in a similar way as described in Simar and Wilson (2010).

Studies with data from North America have (significantly) higher efficiency scores compared with studies from other continents and European studies. One reason for this continent difference might be that studies using data from North America are able to research school districts (and as Table 2-3 and 2-4 show, school districts have higher efficiency scores than schools or departments), and therefore often use districts as the unit of observation, whereas in Europe there are no school districts. Another reason might be that, in many cases, North American studies only look at private schools, which seem to have higher efficiency scores (e.g. Barbetta & Turati, 2003). Furthermore, Table 2-4 shows that studies in the HE sector have higher efficiency scores than studies in PE or SE. This might be because higher education institutes may be more similar to each other compared with primary and secondary education schools. The similarity might lie in the student population (e.g. with

respect to brightness and social status), but also in the organizational structure. For the number of inputs, we see that an increasing number of inputs lead to significantly higher efficiency scores. This result is also positive and significant for the number of outputs. These results were to be expected because, for DEA, it holds that the higher the number of inputs and outputs, the higher the average efficiency score (Tauer et al., 2007).

With respect to inputs, all other inputs compared with the number of staff give higher efficiency scores. Of these, the inputs that involve costs and the number of students reveal significantly higher efficiency scores. For the outputs, only publications and research grants, a combination of the number of students and publications and research grants, and a combination of teachers and research output as outputs are significantly different from student performance. Efficiency scores are lower when publications and/or research grants are used as outputs in a study, and higher for number of students combined with research performance and the number of teachers combined with research performance. The other output categories do not seem to be significantly different from using student performance as output. This might indicate that including quality indicators such as student or research performance in the analysis significantly influences efficiency scores. Furthermore, this indicates that the choice of inputs and outputs influences the average efficiency scores, and that this choice should not be made lightly.

Lastly, using departments as the unit of observation leads to significantly lower efficiency scores than using schools, school districts or study programs as unit of observation. One reason for the finding that studies with departments as unit of observation have lower efficiency scores might be the fact that departments in higher education institutes in most cases have two outputs, both research and teaching. Some studies on departments within higher education institutes argue that an efficiency gain would be possible if departments would focus more on one output than on the other or vice versa (Tauer et al., 2007).

Table 2-4 – Random-effects meta-regression⁹

	Coefficient	t-stat
Constant	0.627 ***	7.63
Higher Education	0.005	0.14
Parametric	-0.056 *	-1.69
Number of outputs	0.014 *	1.75
Number of inputs	0.009	1.42
Year ^a	-0.005 **	-2.25
Returns to scale ^b		
CRS	-0.028	-0.97
Not specified	-0.052	-1.73
Inputs ^c		
Number of students	0.096 **	2.13
Costs	0.107 **	2.42
Student performance	0.065	1.15
Costs and number of staff	0.069 **	2.21
Costs and number of students	0.233 ***	3.23
Number of staff and students	-0.019	-0.41
Outputs ^d		
Publications/research grants	-0.205 ***	-3.92
Expenditures	0.062	1.14
Students and research output	0.085 **	2.4
Teacher and research output	0.152 *	1.87
Unit of observation ^e		
School	0.177 ***	4.19
School district	0.291 ***	3.31
Study program	0.162 **	2.06
Continent ^f		
Europe	-0.051 ***	-1.78
Other continents	0.044	1.04
Between studies variance	0.013	
N	190	

^a 1977 is the base year⁹ Metareg in STATA used.

^b VRS is reference

^c Number of staff is reference

^d Student performance is reference

^e Department is reference

^f North America is reference

*** significant at the 1% level, ** significant at the 5% level, and * significant at the 10% level.

Publication bias

Publication bias is a type of selection bias, which can arise because not all research that has been done with respect to the subject concerned is published or reported. The reason for this could be that researchers or editors do not trust the results, or do not value the results properly. This particular type of bias is very likely to influence the results of a meta-analysis, meaning that there are studies that cannot be found in common databases or books, which are ‘missing’ in the data set of studies created for the meta-analysis. Due to these ‘missing’ studies, the pooled average efficiency score might be biased.

Egger’s test is a commonly used method to test for publication bias. This test, which is more powerful than other publication bias tests, standardizes the average efficiency score, and does a regression of the average efficiency score on its inverse variance, and tests whether the intercept is significantly different from zero (Egger et al., 1997). However, in the case of efficiency scores, it is likely that there is also bias if the intercept is significantly different from 1, as it is unlikely that a study that gives an average efficiency score of 1 will be published. Both these possibilities are tested here. Begg’s test standardizes the effect sizes, and does an adjusted rank correlation to check whether the effect size and its variance are correlated. A funnel plot gives a visual representation of the potential bias. This test has only moderate power in the case of an N smaller than 25 (Begg & Mazumdar, 1994). Doucouliagos and Stanley (2009) introduced the Funnel graph’s Asymmetry Test (FAT) in which the t-value is regressed on the inverse variance. The estimate of the constant in this regression

provides a test of publication bias. In absence of publication bias the constant has to equal 0.

Figure 2-1 presents the effect size vs. the standard error of the effect size. This plot also shows the existence of publication bias. Furthermore, Begg's and Egger's tests on publication bias are estimated, as is the Funnel graph's Asymmetry Test (FAT). All three methods point towards the existence of publication bias, which is presented in Table 2-5. Therefore, Hedges' correction method is used to correct for this bias.

Figure 2-1 – Effect size vs. standard error of the effect size

Table 2-5 – Egger’s test and Funnel graph’s Asymmetry Test (difference from 0 presented here)

Egger's test				
	Coefficient	Standard Error	t-stat	p> t
Slope	0.966	0.008	121.680	0.000
Bias (intercept)	-8.615	1.243	-6.930	0.000
N	190			

Funnel graph’s Asymmetry Test (regressing t-stat on 1/se-es)				
	Coefficient	Standard Error	t-stat	p> t
1/standard error effect size	-2761.792	454.221	-6.080	0.000
Bias (intercept)	141.458	14.763	9.580	0.000
N	190			

2.5 Meta-Analysis – Meta-Regression Analysis

Hedges’ model

The standard meta-regression has one major drawback: namely, it ignores the existence of publication bias. If the meta-analysis is positively tested for publication bias, the meta-regression should also incorporate this bias. Since it is likely that the chance of being published is largely influenced by the significance level of the results of a study, an extra aspect has to be added to the meta-regression to account for this chance of publication. Hedges (1992) developed an approach to estimate a meta-regression taking into account the effect of publication bias. This approach has been used in studies by Ashenfelter et al. (1999), Huang, Maassen van den Brink and Groot (2009) and Haelermans and Borghans (2012). Hedges’ meta-regression assumes that there is a weight function, based on significance levels, which determines the probability that the study is observed. Studies with a high significance level are more likely to be observed. To specify the probability of being

2. A Meta-Analysis on the Efficiency of Educational Institutes

observed, Hedges creates three groups of probabilities, based on the p-value. The weight function consists of all studies, and partly depends on the log-likelihood function for the data. From this likelihood function one can derive the heterogeneity in rates of return, the 'true rate of return', the log-likelihood, and the effect of some study or sample characteristics on the average efficiency score. Hedges' paper is referred to for further details on this topic (Hedges, 1992). With respect to the part that is used in this study, the basic idea is that there is a population of studies, each with a measured average efficiency score that equals the true average efficiency score for the group plus an error term. This is represented in Formula 2-3:

$$\hat{ES}_i = ES_i + \eta_i, \quad (2-3)$$

Here $\hat{ES}_i$ is the true efficiency score, ES_i is the efficiency score and η_i is the error term. As the efficiency scores have a mean of 0.794, all efficiency scores are significantly different from zero. As Hedges' model is built on the assumption that a study is rejected based on the significance level compared with zero, Hedges' model cannot immediately be applied. Therefore, the average efficiency score (ES) is redefined as the difference from the median efficiency score instead of the efficiency score itself. Now, the assumption of being significantly different from zero holds. The difference of the true average efficiency score from the median depends on the characteristics of the efficiency analysis and the estimation technique used:

$$ES_i = P(\text{efficiency_analysis_characteristics, data, estimation method, etc}); \quad (2-4)$$

$$t = \frac{\hat{ES}}{\sigma_{\eta}}. \quad (2-5)$$

Here σ_{η} is the standard error. The relative probability for each study to be rejected equals p_i . This function is defined as a step function on three intervals:

$$\begin{aligned}
 p_i = & \begin{aligned} & g1, \quad 2.054 \leq t < 2.576 \quad \Rightarrow p_i < 0.05 \\ & g2, \quad 1.645 \leq t < 2.054 \quad \Rightarrow 0.05 \leq p_i < 0.1 \\ & g3, \quad 0 \leq t < 1.645 \quad \Rightarrow 0.1 \leq p_i \end{aligned} \quad (2-6)
 \end{aligned}$$

These three groups are used because the 5 and the 10 per cent significance level are often used to indicate insignificant results. Results are robust with respect to changes in these intervals.

Meta-regression analysis with Hedges' model

Table 2-6 presents the results of Hedges' simple meta-regression model. Here σ presents the heterogeneity in average efficiency scores, and the constant presents the average efficiency score. Since Hedges' analysis is calculated based on the deviation from the median efficiency score, this average score should be added to the value of the constant. The constant should be interpreted as 0.797 (0.794 – 0.003) in the restricted model and 0.804 (0.794 + 0.010) in the unrestricted model. As Table 2-6 shows in the restricted model (where $g1=g2=g3=1$), the difference with the average efficiency is -0.003 whereas in the unrestricted model, in which $g2$ and $g3$ are allowed to vary, the difference with the average efficiency is +0.010. Both $g2$ and $g3$ are very significant. The likelihood ratio test and the Chi-statistic and corresponding p-value all point towards the existence of publication bias, which is consistent with the publication bias test. From this, it can be concluded that the pooled average efficiency score is 0.804, which is the average efficiency score corrected for publication bias from Hedges' analysis.

Table 2-6 – Hedges’ simple meta-regression model

<i>Restricted model</i>			
	Coefficient		Z-value
g2	1		
g3	1		
Constant	0.003		0.29
σ	0.153		
Log likelihood	257.230		
<i>Unrestricted model</i>			
	Coefficient		Z-value
g2	0.045	**	2.18
g3	0.007	***	4.1
Constant	-0.010	***	-6.37
σ	0.003		
Log likelihood	1002.382		
Chi statistics	7.93		
p-value	0.0049		

*** significant at the 1% level, ** significant at the 5% level, and * significant at the 10% level.

Table 2-7 shows the extended version of Hedges’ meta-regression model. Table 2-7 again provides evidence for the existence of publication bias because of the likelihood value and the chi-statistic and p-value. The results of Hedges’ extended model show that all variables that were statistically significant in the extended model and the uncorrected meta-regression analysis, as presented in Tables 2-3 and 2-4, are significant in Table 2-7 as well. Furthermore, the magnitude of these variables in Table 2-7 is the same as well. Therefore, we can conclude that there is no difference with the effects shown in Tables 2-3 and 2-4 and the explanations given in the previous section for differences in efficiency scores hold for Table 2-7.

Table 2-7 – Hedges extended meta-regression model

Statistics of unrestricted model			
Likelihood value	336.176		
Chi-statistics for publication bias	177.860		
p-value	<0.001		
Estimate of unrestricted model	Coefficient		Z-value
g2	0.772	**	2.10
g3	0.217	***	3.50
Higher Education	-0.004		-0.14
Parametric method	-0.062	**	-2.27
Number of outputs	0.017	*	1.66
Number of inputs	0.008		1.35
Year ^a	-0.005	***	-2.57
Returns to scale ^b			
CRS	-0.026		-1.07
Not specified	-0.040		-1.55
Inputs ^c			
Number of students	0.079	**	2.04
Costs	0.082	*	1.96
Student performance	0.048		1.04
Costs and number of staff	0.059	**	2.16
Costs and number of students	0.231	***	3.25
Number of staff and students	-0.039		-0.97
Outputs ^d			
Publications/research grants	-0.180	***	-4.06
Expenditures	0.056		1.25
Students and research output	0.090	**	2.88
Teacher and research output	0.115	*	1.66
Unit of observation ^e			
School	0.160	***	4.56
School district	0.234	***	2.93
Study program	0.158	**	2.30

2. A Meta-Analysis on the Efficiency of Educational Institutes

Continent ^f			
Europe	-0.066	***	-2.63
Other continents	0.018		0.48
Constant	-0.165	***	-2.64
σ	0.103	***	18.60

^a 1977 is the base year

^b VRS is reference

^c Number of staff is reference

^d Student performance is reference

^e Department is reference

^f North America is reference

*** significant at the 1% level, ** significant at the 5% level, and * significant at the 10% level.

2.6 Conclusion

In this chapter, efficiency of education and the differences between efficiency in education have been discussed. A meta-analysis is used to estimate the pooled average efficiency score. Next, the extended model is presented and meta-regression is used to analyze the effects on the efficiency score of study and population characteristics. Tests for publication bias showed that publication bias has to be taken into account. This chapter applied Hedges' model to account for publication bias in the meta-regression. Finally, the effects of student, teacher, school, and environmental characteristics on the efficiency score were discussed.

The main finding of this chapter is that the efficiency of education is approximately 0.804, so, on average, there is 20 per cent less educational performance than is potentially possible. Publication bias is found and corrected for. The analysis confirms that efficiency scores are significantly different for several study- and population-characteristics, which implies that there are several possible reasons for differences in efficiency estimates. The expected results with respect to

differences in efficiency scores for different methods are confirmed, as the average efficiency scores are significantly lower for studies that use parametric compared with non-parametric estimation techniques. Furthermore, studies using VRS have higher efficiency scores than studies using CRS. Many other authors have found that using a stochastic estimation method leads to lower efficiency scores than using a non-parametric method (Barbetta & Turati, 2003; Chakraborty et al., 2001; Johnes et al., 2005; McMillan & Chan, 2006). The fact that parametric and non-parametric analyses can differ so much in efficiency scores is in most cases due to limited data availability and poor data quality. Furthermore, it might be the case that authors make the wrong assumption about the functional form of the production function or that they use the wrong inputs, which brings systematic errors in the measurement of the parametric efficiency scores (Fried et al., 2008, ch.2).

With respect to the variables of interest to education economics, the results show that the average efficiency scores are significantly lower for studies from Europe, compared with North American studies. Furthermore, the analysis shows that efficiency scores differ significantly for the different units of observations used in the studies. Studies using departments as the unit of observation get significantly lower efficiency scores than using schools, school districts or study programs as unit of observation. Next, the meta-analysis shows that earlier studies have significantly lower efficiency scores than more recent studies. Finally, the analysis shows that the type and number of inputs and outputs have a major influence on the efficiency scores. More inputs and outputs lead to higher efficiency scores. The type of inputs and outputs also plays a major role in this. The average efficiency score turns out to be lower for primary and secondary education than for higher education, however insignificant.

Efficiency studies using European data present significantly lower average efficiency scores than studies using data from North America. Furthermore, using departments as the unit of observation leads to significantly lower efficiency scores than using schools, school districts or study programs as unit of observation. The finding that earlier studies have significantly lower efficiency scores than more

2. A Meta-Analysis on the Efficiency of Educational Institutes

recent studies might be, for example, because of the fact that in earlier days often a single input, single output framework was used. Lastly, the meta-analysis shows the finding that including more outputs in the model leads to significantly higher efficiency scores. On the one hand, more outputs might lead to a smaller diversity between schools (and thereby a higher average efficiency score), as schools can be very different if compared based on a single output, but the difference might average out as more outputs are included. On the other hand, including more outputs may increase the quality of information on the institutes in the data, which would increase the diversity between the institutes, thereby leading to a larger spread in efficiency scores.

A meta-analysis is a strong tool to estimate overall efficiency scores. Even though efficiency studies differ on account of different choices of, for example, estimation method, and the number and type of inputs and outputs, a meta-analysis makes it possible to compare these studies by incorporating these differences. However, no efficiency scores, or the effect of certain characteristics on efficiency have been estimated in this chapter. Many studies estimate the efficiency scores of educational institutes and some of those studies have estimated the two-step model and explained efficiency scores by factors that are not included in the efficiency analysis itself. Giving an overall overview of efficiency estimates in education by means of a meta-analysis, and giving an overview of the determinants of educational efficiency is new. Therefore, the main contribution of this chapter is that, as far as we are aware, this is the first quantitative meta-analysis on efficiency in education.

3 On the Allocation of Resources for Secondary Schools¹⁰

“Education costs money, but then so does ignorance”

- Sir Claus Moser -

3.1 Introduction

Teacher shortage is a serious concern in both European countries (European Commission, 2006) and the United States (Cooley Nichols et al., 2008; Ingersoll & Perda, 2010). For example, in the Netherlands, the expected shortage of teachers in secondary education for the next 10 years is about 500 full-time equivalents (FTEs) per year (Ministry of Education, 2011c). Figure 3-1 presents the gap between the outflow and inflow of teachers. Similar patterns are detected for other European countries, such as Germany and the UK (e.g., King, 2009; Lipsett, 2008; Santiago, 2002).

¹⁰ This chapter is based on C. Haelermans, De Witte, K., & Blank, J. L. T. (2012). On the allocation of resources for secondary schools. *Economics of Education Review*, 31(5), 575-586.

The increasing teacher shortage is due to four reasons. First, on the supply side of teachers: over time 'teaching' has become less attractive for youngsters as the jobs' social status diminished (Hargreaves et al., 2006; Hargreaves et al., 2007). Second, teaching has become a more 'female' occupation which is considered to combine more easily with a family. Consequently, more teachers work part-time (e.g. 0.5 FTE) (Ministry of Education, 2011b). Third, the ageing of the baby boomers (the cohort born after World War II) is accelerating the outflow from the profession (Ministry of Education, 2011c). Fourth, on the demand side: increasing attention to individual counselling and teaching (particularly for youngsters of low socio-economic status, and for students with lower educational attainments) has increased the demand for teachers (Ministry of Education, 2011c).

Figure 3-1 – The predicted difference between the inflow and outflow of teachers

(Source of data used: CentER data/ECORYS/QQQ Delft, 2006).

This chapter analyses whether alternative school resources, such as supporting personnel, school management, and materials, can be – to a limited extent –

3. On the Allocation of Resources of Secondary Schools

substituted for teachers. In addition, it discusses the scope for performance improvements. Indeed, tightening budget constraints and increasing demands are forcing schools to spend resources in terms of employees, management, and material in the most productive way. Much is at stake. The increasing requirements for schools (e.g. more pupil counselling, additional extra-curriculum activities, use of school buildings during weekends, and summer holidays) all induce pressure on the resources which are already in place. Productivity thus becomes an issue in schools (see Ball & Goldman, 1997; Ministry of Education, 2011b, p. 98). Using the available insights, resources can be reallocated among schools, so that higher educational attainments can be obtained.

Although in policy documents the teacher shortage is referred to as an overall phenomenon (Ministry of Education, 2011c), it is of course clear that some subjects are more liable to a shortage of teachers than others (e.g. Ingersoll & Perda, 2010; Mangrubang, 2005). We abstract from this heterogeneity for two reasons. First, our data include the total number of teachers per school, and do not allow us to make a distinction between the different teachers for the different subjects. Second, even if it were possible to make this distinction in the data, the number of teachers per subject per school is very small, which makes the variation in the number of teachers even smaller. The latter makes it impossible to study the teacher shortage per subject.

In line with the previous literature (Fried et al., 2008), performance is considered in two complementary ways: technical efficiency and misallocation of resources. The former denotes the relative ratio between the available resources and the outcomes. Relative to a production frontier, it measures the output shortfall of schools, given the available resources. The latter enriches the technical efficiency estimations by price information. The measure of misallocation of resources yields insights into the under- or over-utilization of school resources. The misallocation of resources in secondary education has been largely overlooked, and is often presented by allocative efficiency. The available allocative efficiency studies mostly consider higher education (Cherchye & Vanden Abeele, 2005; Johnes & Johnes,

2009; Soares de Mello et al., 2006; Tauer et al., 2007). Some rare exceptions are Banker et al. (2004) and Grosskopf et al. (1997, 2001) who both study school districts in Texas.

The remainder of this thesis examines performance in Dutch schools. Studying Dutch schools is attractive and insightful for three reasons. First, standardized performance measures of Dutch students makes educational attainments very comparable. Second, there is information on student achievement, which compares the educational career of a student (both in terms of school track and retentions) with the education track predicted for a student at the end of primary education. Third, Dutch schools receive a annual lump-sum budget from the government, which is at the discretion of the school, such that, within the existing legal framework, the allocation of this budget among the several resources is the decision of the school (*Stb.* 1963; 40). Therefore, a significant heterogeneity in hired resources, in terms of management, teachers, supporting personnel, and material use is observed.

From a methodological perspective, we assume that schools are maximizing educational attainments with a given budget. From this starting point, one can derive an economic model which explains the performance and the optimal use of resources. We use an indirect budget-restricted indirect-output distance function (IOD) model (Blank, 2009; Grosskopf et al., 1997). In particular, technical efficiency and misallocation of resources are estimated against a frontier consisting of best-practice observations. We examine the potential increase in educational attainments under budget constraints.

The remainder of the chapter is structured as follows: Section 3.2 explains Dutch secondary education and Section 3.3 outlines the methodology and the model. Section 3.4 describes that data and data sources, while Section 3.5 discusses the results. Some concluding remarks are presented in Section 3.6.

3.2 Dutch Secondary Education

In this section the Dutch secondary education sector is explained in order to understand the choices, explanations, and interpretations in the remainder of this chapter and the thesis as a whole.

Dutch secondary education offers five different levels of education, which are concluded with a national examination in the final year. Students enter a level of education, based on a test and a recommendation given in elementary school. The first three years of secondary education are called lower secondary education. The last years of secondary education (the total number of years in secondary education depends on the level of education) are called upper secondary education. The five different levels of secondary education is as follows (Dutch abbreviations in brackets)¹¹:

- practical education (PRO, four years);
- practical prevocational school (VMBO-bk, four years);
- theoretical prevocational school (VMBO-gt, four years);
- higher general school education (HAVO, five years);
- pre-academic education (VWO, six years).

Most schools offer a combination of these levels of education. In VMBO, a distinction is made between a theoretical (VMBO-gt) and a practical (VMBO-bk) course of study. This thesis distinguishes eight types of schools, ranging from schools offering solely PRO, VMBO or VWO to the comprehensive schools that offer all levels of education.

In the Netherlands, the concept of a school can refer to several levels of aggregation. The highest level of aggregation is the governing body. A governing body manages or administrates several schools. There used to be a one-on-one relationship between the school and the governing body. However, in the 1990s the Dutch government stimulated an increase in scale. As a result, it is now quite common that a governing body represents more than one school. The second level

¹¹ Eurydice is used for the terminology (Executive Agency of Education Audiovisual and Culture, 2009a, 2009b).

of aggregation is the school as an organization with one or more locations. This is the unit of observation in this thesis. Most schools have two locations. At the lowest level a location itself might be referred to as a school.

Secondary schools in the Netherlands receive a lump-sum payment from the government every year. Within the existing legal framework, the allocation of this budget among the several resources is the decision of the school. Note that the lump-sum budget excludes (large and discontinuous) payments for housing infrastructure (the latter are provided by the municipal and central government). In most cases, the municipality is responsible for providing the accommodation of the school. However, the school itself is responsible for the maintenance of the building. In practice, this means that the school is granted a certain amount of finance from the municipality to fulfil this responsibility.

Lastly, an important aspect for this research is that most schools teach traditionally, with one teacher teaching one group of pupils in one classroom. In this setting, communication is mostly unidirectional: the pupils listen while the teacher speaks. Some schools are moving towards a more modern teaching method in which thinking for oneself, independent group work, and the pupils' own initiative are the keywords.

3.3 Methodology

Economic model

We start from the presumption that schools, with an exogenous budget, are maximizing the educational attainments of students. In a sense, schools are trying to attain best practice in educating students each year, and, thus, try to reach the frontier of the production function with their budget. The best-practice frontier can be determined in various ways (for a discussion, see Fried et al., 2008). This chapter applies the rather strict Corrected Ordinary Least Squares (COLS) method in which the OLS regression is shifted to the observation(s) with the most favourable ratio of

outputs to inputs. These observations are considered best practice and obtain an efficiency score λ equal to 100. Inefficient observations lie below this best-practice frontier and obtain an efficiency score λ smaller than 100. The score $((100/\lambda)-1)*100$ can be interpreted as the percentage output increase if the observation were to perform as efficiently as its best practice. If in the efficiency estimation the observation is further constrained by the total cost of inputs, what is called an indirect-output distance (IOD) function is estimated (for a detailed discussion, see Färe and Primont, 1995).

We apply an IOD function which is similar to the model of Grosskopf et al. (1997). Schools are producing multiple outputs (see below) under budget constraints, which can be thought of as exogenous to the schools. The deterministic model can be written briefly as follows (note that the full budget-restricted output-distance function including the full share equations, is presented in Appendix A-1):

$$\ln(IOD) = f(\ln y, \ln\left(\frac{w}{C}\right), \ln z), \quad (3-1)$$

where IOD stands for the indirect-output distance function with an a priori determined functional form, denoted by f (specified as a Translog); y is the vector of outputs; w denotes the vector of prices of the used resources; C are the total costs; and z is a vector capturing fixed inputs or school characteristics (reflecting observed and unobserved heterogeneity). Several (weak) properties, required for a meaningful estimation, hold for the model. It is straightforward to verify (see, e.g. Coelli et al., 2005; Färe & Primont, 1995) that the model is non-decreasing and quasi-concave in $\frac{w}{C}$. Conditional independence (cfr. a separability condition) is imposed on the model, in that information on one input does not yield information on other inputs. This implies that the inputs are not related to one another, in the sense that one cannot know anything on, for example, the number or costs of teachers if one knows the number or costs of managers. This condition is necessary in order to identify the effects of the different inputs on productivity, and in turn determine the first derivative with respect to costs and the optimal numbers for each input.

The cost shares are immediately derived from the IOD by Roy's identity, in particular, as the first derivative with respect to costs (see Färe & Primont, 1995, p. 92). They correspond to:

$$S_i = \frac{\delta \ln IOD\left(y, \frac{w}{C}, z\right)}{\delta \ln\left(\frac{w_i}{C}\right)} \bigg/ \sum_n \left[\frac{\delta \ln IOD\left(y, \frac{w}{C}, z\right)}{\delta \ln\left(\frac{w_n}{C}\right)} \right] \quad (3-2)$$

Including cost shares (and thus estimating an IOD function instead of an output distance function) is convenient as it allows us to estimate the optimal number of different types of costs (e.g., for personnel, management, supporting personnel, and material use). The difference between the actual costs and the optimal costs yields the scores of misallocation of resources¹².

The indirect-output distance function and the cost-share equations provide a set of equations from which the parameters can be estimated. For the estimation of the system of equations, we compared different starting values for the coefficient of $\frac{w}{C}$. The results proved insensitive to changes within the range of plausible starting values, which points towards robust results. In practice, the system of equations is estimated by system OLS¹³. The estimation of the set of equations corresponds to a least squares estimation with multivariate regression systems under a joint density¹⁴. Ideally, we would want to use the Full Information Maximum Likelihood estimation method. However, due to the complexity of the model and non-linearities, this is not

¹² From a practical perspective, an additional advantage of including cost shares is that the estimations are more efficient as more information is added to the model (Blank, 2010; Greene, 2008; Grosskopf et al., 1997).

¹³ This leads to similar results compared with using SUR as estimation method.

¹⁴ Note that one share equation in the output distance function model has to be dropped because the shares add up to one causing the variance-covariance matrix of the error terms to be singular.

possible and the minimum distance estimator LSQ is used instead. However, this implies adopting the assumption that there are no simultaneity problems with endogenous variables, and that the disturbances are multivariate normal. We are aware that this is most likely not the case for our estimations, but, given the complexity of the model, this is the best we can do. For the interpretation of our results, this means that our results are most likely biased. However, due to the complex nature of the model it is practically impossible to determine the direction of the bias. Therefore, the results will have to be interpreted with caution, and previous and future research will have to confirm these results.

Observed and unobserved heterogeneity

Observed heterogeneity is captured by including exogenous characteristics in the model (denoted by z). The exogenous characteristics cannot be influenced by policy or school management, and these variables are directly included in the output distance function and share equations as the z -variables. Including or excluding exogenous characteristics influences the position of the frontier, so that they have to be taken into account in determining the optimal composition of resources and student attainment. This is graphically presented in Figure 3-2.

To capture the part of heterogeneity that is unobserved, we exploit the panel data structure of the data by including school type, year, and region-fixed effects in the estimation. We observe 8 different school types (ranging from pre-vocational education to pre-university education), 6 years, and 4 regions¹⁵. Region-fixed effects are included, as the Education Inspectorate argues that there is a large dispersion in school attainments across regions (Education Inspectorate, 2011). The fixed effects are estimated for the output-distance function but not for the share equations (as the share equations are a derivative of the output-distance function with respect to

¹⁵ Including school type and year interacted with each other and including region separately (log likelihood (logl)=20597.5, degrees of freedom (df)=153) as fixed effects is tested against using single-fixed effects for school type, year, and region (logl=20589.7, df=113). The tests are in favour of the latter.

$\left(\frac{w}{C}\right)$, and the fixed effects do not appear in the output-distance function as an interaction term with $\left(\frac{w}{C}\right)$.

Figure 3-2 – The production frontier with and without environmental variables

3.4 Data

To estimate the technical efficiency and misallocation of resources of schools, we apply the parametric IOD model to a representative balanced panel data set of 448¹⁶ Dutch secondary schools covering the school years 2002/03 to 2007/08. For convenience the school year 2002/03 is referred to as 2002, and so on. The data are obtained from several existing administrative data sets from the Ministry of Education (DUO¹⁷), the Education Inspectorate¹⁸, and Statistics Netherlands.

¹⁶ The full data needed for estimation of the production model is only available for 448 out of 593 secondary schools.

¹⁷ An organization owned by the Ministry of Education, which collects data on schools and provides the budget for schools, and gives them information on their used resources, students and student performances.

Production

Educational production is commonly defined as some kind of measure of educational attainments in knowledge and skills (Wenger, 2000). We use two measures for educational production: the average student national examination grades per school, and the average student achievement each year during secondary education¹⁹. The former serves as a proxy for attainments, while the latter serves as a proxy for quality.

With respect to the first, every graduating student undertakes this test at the end of secondary education in the subjects for which the student is registered. The exams are graded in a double-blind way, and the school and the teacher of the students cannot directly influence the outcome of the exams²⁰. The average grade per school is based on the average of all national examination grades for all students in all subjects. Note that the level of analysis is the school and not the subject or the student. Our results therefore do not allow any conclusions to be drawn on teacher shortage in individual subjects.

To proxy the quality of the educational process, we include a composite number of the Education Inspectorate: average aggregated corrected achievement. This number 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. Student achievement values are separately calculated for lower secondary education and for upper secondary education. Lower secondary education consists of the first three years of secondary education, for upper secondary education a distinction is made between the following groups:

¹⁸ Organization that inspects the quality of the schools one to three yearly, and publishes results on school and student performance yearly.

¹⁹ A robustness check with only national examination grades as output, and number of students and student achievement as control variables yields very similar results.

²⁰ A school might be able to indirectly influence exam results on an abstract level as they may, for example, be teaching to the test. However, as the exams are nationally developed and cannot directly be influenced by the school, we assume independence of these results.

Part II – Efficiency of Education

- Pre-vocational-practical (VMBO-bk) grade 4;
- Pre-vocational-theoretical (VMBO-gt) grade 4;
- higher general education (HAVO) grade 4 and 5;
- pre-academic education (VWO) grade 4, 5 and 6.

For lower secondary education, student achievement is based on the level of education the student is attaining in the last year of lower secondary education compared with the recommendation given at the end of primary education, corrected for any repeated years. The calculation made by the Education Inspectorate for the achievement of lower secondary education is given by the following:

$$\text{Student achievement first 3 years} = \frac{(\text{educyr3} - \text{recom} - \text{rep} + 2)}{2}, \quad (3-3)$$

where *educyr3* is the level of education the student attains in year 3; *recom* is the recommended level of education given at the end of primary education; and *rep* is a value for the number of any repeated years (which is usually not higher than 1). Both *educyr3* and *recom* are represented by a value, in which, for example, VMBO-bk=1, VMBO-gt=2, HAVO=3 and VWO=4. This implies that, if a student has never repeated a year and attains the same level in year 3 as was recommended in primary education, the student achievement of the first three years of secondary education is equal to 1.

The student achievement measures for upper secondary education of VMBO-bk, VMBO-gt, HAVO and VWO are represented by the percentage of students who graduate without a year of delay or without dropping out. Therefore, the value of upper secondary education achievement has a maximum of 1, whereas the value of the lower secondary education achievement can be above 1. Next, all the student achievement levels (both for lower and upper secondary education) per school are corrected by the Education Inspectorate for the degree of urbanization and the percentage of students living in a poverty area, in order to include a certain amount of quality information in the output measure.

3. On the Allocation of Resources of Secondary Schools

Both output measures are multiplied by the total number of students, as more expenses should lead either to higher performance of the current number of students or to the same performance for a higher number of students. We chose to multiply by the number of students instead of adding the number of students as an extra output, because we believe that student performance corrected for the number of students is a better output measure than only the number of students. Summary statistics are provided in Table 3-1. Table 3-1 shows that the average national examination grade is 6.4 on a 10-point scale. A grade of 5.5 is sufficient to pass the subject. The grades range from 5.4 to 7.2. Despite the large variation, the standard error is relatively low. Student achievement has an average of 0.832 and the total number of students is on average 1,744 per school.

Resources

School resources are summarized in terms of four relatively homogeneous categories: (1) management personnel; (2) teaching personnel; (3) supporting personnel; and (4) material supplies. 'Capital', in the sense of housing (school building) infrastructure is not accounted for owing to data constraints. For the human resources (the first three groups), data are available on full-time equivalents (FTEs) and costs. In particular, the costs correspond to the costs per FTE per year. The total average number of FTEs per school is 163, of which 80 per cent are teachers, 14 per cent supporting personnel, and the remaining 6 per cent managing personnel. The total costs of a school are about 12 million euros per year. Most of the costs are spent on teaching personnel, followed by materials, management personnel, and, lastly, supporting personnel. Besides teaching, a teacher has some management and administrative duties. The different tasks within one function are not officially reported. We assume that there is a homogeneous distribution of these different tasks within one function, between teachers both within one school and across schools.

Prices of the personnel groups are corrected for regional differences, characteristics of the personnel, such as sex, years of experience, and percentage of

part-time jobs. Furthermore, they are corrected for the number of first- and second-year pupils and the number of VMBO-pupils at the school. This correction is necessary since teachers need a different type of teaching qualification for these types of pupils. The same prices are used for the different types of personnel for all schools in one region in the same year. For price differences in material the consumer price index of Statistics Netherlands is used. Different indices are used for different types of material costs, such as energy and water, and cleaning costs. Since we know the composition of the material per school, we can calculate the prices for each type of material by multiplying the index by the cost component of each type and then adding them up. For the material costs which do not have an index, the average price index for materials is used.

Control and environmental variables (z-variables)

Some influences are exogenous to the school, and create heterogeneity across schools. These variables include the number of schools per governing body, the degree of competition, the share of students from a disadvantaged neighbourhood, the degree of urbanization, management and teacher experience, and the number of locations per school. The latter 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). All these variables influence the efficiency of a school, but cannot be influenced by the school management. 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 negative effects of a centralized bureaucracy. The average number of schools per governing body is almost 6 with a minimum of 1 and a maximum of 31. The latter is one specific governing body which manages many schools.

Previous research has shown that competition fosters productivity (Abbott & Doucouliagos, 2009; Millimet & Collier, 2008) and that urbanization is related to productivity (Naper, 2010). Competition is estimated as the number of schools per

3. On the Allocation of Resources of Secondary Schools

10,000 inhabitants in a municipality and has an average value of 0.23. The share of students from a disadvantaged neighbourhood amounts to, on average, 2.7 per cent. Furthermore, we see that the number of locations per school is 2 on average, but ranges from 1 to 10. We observe managers with an average experience of 17 years. The average experience of teachers is 12 years, while the average number of part-time teachers is 26 per cent. All these variables are listed in Table 3-1 below.

Table 3-1 – Summary statistics (for 2007)

	Mean	Standard deviation	Minimum	Maximum
National examination grade	6.4	0.2	5.4	7.2
Student achievement	0.832	0.061	0.571	1.067
Total number of students per school	1744	991	158	5971
Number of FTE management personnel	9	7	1	44
Number of FTE teaching personnel	131	80	15	468
Number of FTE supporting personnel	23	16	3	119
Total FTE	163	97	19	616
Total costs (in 1000 euros)	12311	7944	1078	51481
Management personnel costs per FTE (in euros)	88806	28465	19375	231925
Teaching personnel costs per FTE (in euros)	67528	16086	38581	155061
Supporting personnel costs per FTE (in euros)	21239	7939	10406	93087
Competition	0.23	0.23	0.02	2.89
Number of schools per governing body	5.57	7.57	1	31
Share of students from a disadvantaged neighbourhood	0.027	0.030	0	0.197
Degree of urbanization	0.504	0.319	0.010	0.969
Average years of experience managers	17.0	5.1	0.2	30.8
Average years of experience teachers	12.0	2.4	1.3	18.9
Percentage part-time teachers	0.26	0.08	0.09	0.54
Number of locations per school	2.1	1.5	1	10

3.5 Results

The full estimation results are presented in Appendix A-2. Almost 70 per cent of the parameters are significant at the 5 per cent level; the R-squares of the cost shares are reasonable (ranging up to 15 per cent); and all cost variables have a positive coefficient. The latter indicates that the monotonicity condition is satisfied. The concavity condition is also satisfied, as all elasticities of substitution are negative and the Hessian is negative definite in all cases²¹. These results can be found in Appendix A-3.

Technical efficiency

Table 3-2 presents the average technical efficiency scores for Dutch secondary schools for 2002-2007. Table 3-3 presents the average efficiency scores averaged over 3 years (per school)²². Because of the similarity of the results, and, because of space limitations, we first discuss the results for 2007. Note that in our setting, by construction, efficiency ('doing things right') is closely related to effectiveness ('doing the right things'). Indeed, our output variable corresponds to average pupil test scores and average student achievement. Therefore, in most cases the most efficient school also has a high average national examination grade²³. With respect to efficiency, a deviation (λ) from 100 denotes that the school can improve its output by $((100/\lambda)-1)*100$ per cent if it were to use its given resources as efficiently as its best practice in 2007. The average technical efficiency score for the 448 schools is 70.1 per cent, implying that with the same resources they could improve educational performance by about 40 per cent. 75 per cent of the schools could improve the educational attainments of the students by more than 40 per cent (note that this is controlled for observed heterogeneity among schools, as well as for

²¹ The Hessian has been evaluated for every observation separately.

²² Note that the maximum is not 1 in this table, since the numbers are averaged per school and each year a different school was the best performing school.

²³ Although this does not necessarily need to hold.

resources)²⁴. We observe a minimum efficiency score of 29.5 per cent, which indicates that the least-performing Dutch school could improve the educational attainments of its students by more than 230 per cent if the school were to use its resources as efficiently as its best practice. The standard deviation of the scores is 9.7 per cent, which indicates that, on average, Dutch secondary schools are very comparable and do not differ much in terms of efficiency. From the technical efficiency scores we conclude that, although, in general, Dutch secondary schools perform rather well, efficiency gains are possible for most schools.

The results for other years are similar to the outcomes for 2007. For 2003, the lowest average efficiency score is observed. However, this score is not very different from the scores from 2002 and 2004. The highest average efficiency is observed in 2005. In that year, schools could improve their performance by about 25 per cent if they were to perform as efficiently as the best-performing school. Table 3-3 shows that the efficiency scores averaged over 6 and 3 years do not differ much from the separate efficiency scores per year, as presented in Table 3-2. Rank-order correlations are presented in Appendix A-4, and show that there are high correlations between ranks over the years and also between the averaged efficiency scores over years²⁵.

Table 3-2 – Technical efficiency scores of Dutch secondary schools 2002-2007 in percentages

	2002	2003	2004	2005	2006	2007
Average	67.03	65.23	68.59	74.33	72.25	70.10
Standard deviation	7.66	7.00	7.09	8.11	9.48	9.74
Minimum	37.01	38.39	39.50	45.20	30.57	29.52
2nd quartile	62.30	61.24	64.56	70.27	68.09	65.74
Median	67.10	65.51	68.33	74.73	73.30	70.75
4th quartile	71.49	69.79	73.01	79.57	77.49	76.76
Maximum	100.00	100.00	100.00	100.00	100.00	100.00

²⁴ Running the analysis without fixed effects shows that the relatively high efficiency scores are not related to the inclusion of the fixed effects.

²⁵ Note that all these correlations are significant at the 1 per cent level.

Table 3-3 – Technical efficiency scores of Dutch secondary schools averaged over years 2007 in percentages

	Average over 6 years	2002- 2004	2003- 2005	2004- 2006	2005- 2007
Average	69.59	66.95	69.38	71.72	72.23
Standard deviation	6.62	6.78	6.56	7.19	8.67
Minimum	49.07	41.14	46.62	48.44	36.65
2nd quartile	65.98	62.90	65.54	67.92	68.56
Median	69.47	67.05	69.27	71.97	72.90
4th quartile	73.92	71.20	73.78	76.39	77.38
Maximum	89.20	99.67	91.98	94.93	95.51

Misallocation of resources

While technical efficiency does not account for prices and costs, allocative efficiency scores do, and so do misallocation scores. It can be expected that not all schools use the optimal input mix. In this case, higher output can be obtained by re-allocating the inputs. As schools differ on in terms of resources, student population, location etc., the optimal input mix will be different for every school.

The results are presented in Tables 3-4 and 3-5. A positive sign denotes an overutilization of the input (i.e. misallocation could be improved by diminishing this variable), while a negative sign indicates underutilization of the input. On average, the share of management and supporting personnel is lower than the optimum. On the other hand, schools use more teachers and material than would be optimal. The deviation from the optimum for material and supporting personnel is significantly different from zero at the 1 per cent level. The deviation from the optimum for teaching personnel is significant at the 10 per cent level. The deviation from the optimum for management personnel is not significant. In the following description and interpretation of the results we focus on the results of the median in 2007²⁶.

From Table 3-4, we observe an average overutilization of teachers of 0.21 per cent. This average hides a significant dispersion across the distribution. The latter is also visible in Figure 3-3. As the average school hires 131 FTE teachers, an

²⁶ The results for the other years are available from the author upon request.

3. On the Allocation of Resources of Secondary Schools

overutilization of 0.13 per cent indicates that the same or higher test scores could be reached with 130 FTE teachers. It is interesting to note that the 0.21 per cent overutilization in FTE in 2007 amounts to about 130 FTEs out of on a total of 64,000 FTEs (Ministry of Education, 2011b), which is a lot smaller than the teacher gap in 2007 that was predicted in the report of CentER data, ECORYS and QQQ Delft (2006). In the latter report, the teacher shortage in 2007 is estimated to be about 1,000 out of a total of almost 62,000 FTEs (1.6 per cent)²⁷. This might indicate that, for some schools, part of the predicted teacher shortage could be tackled by better allocation of resources. However, the average school operates close to the optimum with respect to the number of teachers, and our results show that Dutch schools are doing rather well with respect to misallocation of resources.

Furthermore, there is an underutilization of 0.55 per cent of managers. This denotes that the median school could increase its management structure from 9 FTE managers to 9.1 FTEs. This small average underutilization hides a significant dispersion across the distribution. While the median school has an underutilization of management, the top 25 per cent of schools have a (large) overutilization. These correspond to the larger schools of the total sample. Although previous research indicates a sufficient number of managers at Dutch secondary schools (Blank et al., 2007b), our results show that many schools have either an underutilization (the bottom 50 per cent) or overutilization (the top 25 per cent) of managers. This mismatch can be explained by the constant warning of an upcoming shortage, not only for teachers but also for managers (CentER data/ECORYS/QQQ Delft, 2006). This has been confirmed by sector organizations (Sectorbestuur Onderwijsarbeidsmarkt, 2009; Van Velden, 2002). Another reason might be the increasing number of the middle managers and organizational units (Odenthal et al., 2007; Simkins, 2000).

Interestingly, schools have 0.32 per cent too few supporting personnel than would be optimal, and 0.24 per cent too much of materials than would be optimal. The underutilization of supporting personnel is rather surprising given its low costs, and given its effectiveness (e.g. Groom & Rose, 2005).

²⁷ This situation was predicted for an economic recession, which was the case in 2007.

Despite the critical teacher shortage, it is interesting to observe that some schools use too many teachers and too few of the other resources. At one end, there is a school which should hire 14.6 per cent more teachers (i.e. increasing them from 70 FTEs to 80 FTEs) to improve the educational attainments of its students. At the other end, a school hires 12.2 per cent too many teachers. The median school is very close to the optimum regarding the number of teachers. It seems that the teacher shortage in the Netherlands could possibly be decreased by reallocating teachers across schools and/or by relying on alternative resources. Not all tasks in a school have to be fulfilled by teachers. For example, supervision of a group of students in self-study, the support and coaching of student projects, or supervision during lunch time can be undertaken by supporting personnel (see e.g. Blatchford & Sumpner, 1998).

Table 3-4 – Misallocation of resources in percentages (data for 2007)

	Average	Standard deviation	Minimum	First Quartile	Median	Third Quartile	Maximum
Deviation from optimal share - management	-0.18	2.89	-6.06	-2.09	-0.55	1.46	20.61
Deviation from optimal share - teaching personnel	0.31*	4.54	-14.62	-3.24	0.21	3.09	12.25
Deviation from optimal share - supporting personnel	-0.26***	1.38	-3.22	-1.24	-0.32	0.57	7.01
Deviation from optimal share - material	0.76***	4.03	-9.96	-1.60	0.24	2.68	16.49

*** Significantly different from 0 at the 1% level, ** significantly different from 0 at the 5% level, and * significantly different from 0 at the 10% level.

Table 3-5 – Median of misallocation of resources 2002-2007 in percentages²⁸

	2002	2003	2004	2005	2006	2007
Deviation from optimal share - management	-0.16	-0.04	-0.15	-0.41	-0.59	-0.55
Deviation from optimal share - teaching personnel	0.91	0.86	0.81	-0.10	-0.59	0.21
Deviation from optimal share - supporting personnel	-0.05	-0.04	-0.01	-0.03	-0.07	-0.32
Deviation from optimal share - material	-1.17	-1.38	-1.11	-0.07	0.60	0.24

²⁸ Other output for all years as presented in Table 3-5 is available from the author upon request.

Figure 3-3 – Over/Underutilization of teachers (in percentages)

Trends in misallocation of resources

The inter-temporal trends in resource utilization allow us to explore the under- and overutilization over time. Teacher shortage has already been an issue for some time (e.g. CentER data/ECORYS/QQQ Delft, 2006; Municipality of Amsterdam, 2003; Nelis, 2004; Stijnen, 2003). Nevertheless, we observe in Figure 3-3 that the median school in the sample has an overutilization of teachers between 2002 and 2004. In those years, a better allocation of teachers over schools could have decreased the teacher shortage for the bottom quartile of schools. Also in this inter-temporal perspective, the numbers are on a similar level as the predictions made by CentER data, ECORYS and QQQ Delft (2006). Furthermore, the resulting numbers from our study are on the same level as the numbers from the report of the Ministry of Education (2011c), in which an annual shortage of about 500 FTEs was predicted. Only between 2005 and 2006 did a median underutilization emerge, which changed to a median overutilization in 2007. A possible reason for this sudden change in trend could be

that schools and municipalities anticipated the teacher shortage, as this had already been an issue for a long time (Municipality of Amsterdam, 2003).

Control variables in the model

The estimation results of the exogenous variables are presented in Table 3-6. They indicate the correlation between the control variables and educational performance or cost shares. In the IOD model, a positive (negative) sign indicates an unfavourable (favourable) influence on performance. We realize that the model includes several interaction terms of these control variables (with each other, with outputs, and with costs), which makes it hard to compare only the single terms that are presented in Table 3-6. However, a calculation of the average values of our variables multiplied by the coefficients from Table A2 in the Appendix shows that the sign of these values calculated per control variable are not very different from the signs of the coefficients that are presented in Table 3-6. The only difference can be found in the coefficient of urbanization, which has a negative sign (so a positive interpretation) in our recalculation.

An increase in competition increases student performance. This is in line with the previous literature (see e.g. Bradley et al., 1999; Millimet & Collier, 2008). An increase in the share of part-time teachers decreases student performance. This is also in line with the previous literature (e.g. Jacoby, 2006). Interestingly, we find that teachers with more experience decrease student performance. It might be possible that their motivation is different from younger teachers, leading to lower student performance. The finding on urbanization, although insignificant, is similar to that in the previous literature. Naper (2010) shows that schools with many students from rural areas tend to be less efficient, indicating the importance of including both the degree of urbanization and a region variable. The negative finding on urbanization in Table 3-6 is different from this finding in the literature. However, the finding from our calculation underlines the literature described by Naper (2010). The share of students from a disadvantaged neighbourhood decrease student performance. In

the literature we also find that disadvantaged students perform less well in education (Becker & Luthar, 2002; Gaziel, 1997).

The number of schools per governing body and the number of locations decrease student performance. An extra location implies a larger number of students and thereby an increase in scale. The positive relationship between scale and school performance has also been stressed in literature (e.g. Chakraborty et al., 2000; Johnes et al., 2005).

Table 3-6 – Estimation results for control variables in the model

Variable	Estimate	Standard Error	T-statistic	P-value
Competition	-0.038	0.008	-4.748	[.000]
Number of schools per governing body	0.012	0.004	2.835	[.005]
Percentage students from disadvantaged neighbourhood	0.032	0.003	9.704	[.000]
Percentage urbanization	0.008	0.009	0.927	[.354]
Number of locations per school	-0.077	0.011	-7.111	[.000]
Years of experience teachers	0.078	0.021	3.700	[.000]
Share part-time teachers	0.068	0.016	4.218	[.000]

Productivity development

Figure 3-4 shows the productivity development between 2002 and 2007, using the parametric model with cost shares as inputs, and student performance, student achievement and student numbers as outputs. Figure 3-4 shows that productivity has decreased over these 6 years. This decrease in productivity is due to the increased investments in education relative to a decreased quality of education (Kuyper & van der Werf, 2007) and decreased value of diplomas (Lange de & Dronkers, 2007). In other words, performance in education has decreased, causing productivity to decrease.

In this chapter we find that 75 per cent of schools can improve the efficiency by more than 30 per cent. We also find that schools are doing rather well with

respect to the allocation of resources. This implies that the possible efficiency gain for most schools cannot be attained by changing allocative behaviour, but may be attained by improving the productivity of the sector.

Figure 3-4 – Productivity development between 2002 and 2007

3.6 Conclusion

This chapter has studied the optimal allocation of resources in secondary education. We use a parametric efficiency model with budget constraints, in particular the budget-restricted indirect-output distance function, to estimate both technical efficiency scores and misallocation of resources for 448 Dutch secondary schools between 2002 and 2007. Four types of resources are distinguished: management, teachers, supporting personnel, and material use. With the given resources, schools are supposed to maximize student performance, as proxied by national examination grades and student achievement. We control for heterogeneity among schools by including control variables, e.g. information on the share of students from disadvantaged neighbourhoods.

The results indicate that the average technical efficiency score for the 448 schools amounts to 70 per cent, implying that with the same resources schools could improve their educational performance by about 30 per cent. Moreover, we observe

an overutilization of teachers by 0.21 per cent (which is about 100 FTEs (Ministry of Education, 2011b)) and an underutilization of other personnel resources in 2007. The absolute numbers of our results are a lot lower than the predictions made by CentER data, ECORYS, and QQQ Delft (2006), who predicted a teacher shortage of about 1,000 FTEs in 2007 out of a total of about 62,000 teachers. Our findings imply that a different allocation of resources is not necessary for most schools, and that the national teacher shortage cannot be solved or deduced by the reallocation of resources. More than a 25 per cent of the schools have between 3 and 15 per cent too few teachers, while the top 25 per cent quarter of the schools have between 3 and 12 per cent too many teachers. Half of all these schools belong to the 25 per cent that have an efficiency score below 70 per cent. In line with studies on allocative efficiency, this research shows that these schools could improve their efficiency by using a different mix of their resources, given their prices.

From an intertemporal perspective and in contrast to previous findings, the results indicate that between 2002 and 2004 there was a slight overutilization of teachers. In those years, a better allocation of teachers between schools could have decreased the teacher shortage somewhat for the bottom quartile of schools. Despite the significant heterogeneity across schools, this chapter argues that gains with respect to allocation of resources are possible in some secondary education schools by reallocating resources. On average the Dutch secondary education sector is rather homogeneous, and it will be hard to realize large efficiency gains. Possible efficiency gains can be found in increasing the productivity of the educational sector, which has decreased over these six years.

The conclusion is that, on average, Dutch schools are performing rather well, but that some schools could improve their efficiency. Schools in the bottom 25 per cent could shift some of the tasks from teachers to, for example, computer use, or other schools could hire more teachers instead of support or management personnel.

4. Diffusion and Determinants of Innovations²⁹

“Creativity is thinking up new things.

Innovation is doing new things”

- Theodore Levitt -

4.1 Introduction

Educational innovation is a hot topic in international political discussions. It is a major component of the Lisbon Agenda (Council of the European Union, 2007; European Union, 2009) and has been emphasized ever since. The Lisbon strategy has set the goal for Europe *“to become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion”* (High Level Group chaired by Wim Kok, 2004, p. 8). Investment in human capital development is one of the main topics mentioned to achieve this goal. Another way to achieve the goal is to make use of

²⁹ This chapter is an extended version of C. Haelermans (2010). Innovative power of Dutch secondary education. *Innovation: Management, Policy and Practice*, 12(2), 154-165.

innovation. Educational innovations are therefore very relevant in achieving that goal.

It is argued that innovation in education improves the international competitiveness of a country (Cantwell, 2006; Ch. 20). Moreover, it is said to help to decrease school dropout and improve school quality, performance, and skill development (Hoyle & Collier, 2006). Consequently, policies to stimulate educational innovations have received a great deal of attention and attract vast subsidies. These policies also mean that schools are increasingly inclined to innovate (Giles & Hargreaves, 2006). While little is known about the effectiveness of innovations, it is known that there is great variation between schools in their efforts to innovate. However, it is *not* known what factors determine this variation. Therefore, this chapter sets out a descriptive analysis to find the determinants that explain the differences in innovative power between organizations.

The aim of this chapter is threefold: first, to show the diffusion path of some specific clusters of innovation in Dutch secondary education; second, to research which determinants of innovations matter most; and, third, to relate this to the literature. This chapter contributes to the existing literature by using rather unique Dutch innovation data and by determining what factors have played a role in determining the share of innovations in secondary education.

The remainder of this chapter is structured as follows: Section 4.2 reviews the related literature. A description of the data and data sources follows in Section 4.3, and then the methodology is outlined in Section 4.4, after which the results are presented and related to the literature in Section 4.5, i.e. the results with respect both to the diffusion of innovations and to the determinants of innovations. The chapter ends with concluding remarks.

4.2 Literature Review

This chapter builds on the literature on innovations, the literature on the determinants of innovations and the literature on educational innovations. This

chapter also focuses on the problems of connecting innovations to educational outcomes. This section is primarily based on the existing literature on innovations and the diffusion of innovations. As defined in the introduction of this thesis: *'innovation is an idea, process or product that is new for a particular organization when it is introduced by that organization'*.

The first part of this section is related to the literature on the theory of innovations. In this section, we by no means pretend to give a complete overview. Rather, this section mentions some of the important relevant innovation theories for this thesis and focuses on specific literature with respect to the diffusion and determinants of innovations. Schumpeter is the first to explicitly state a specific theory on innovation. According to Schumpeter, economics is a dynamic process in an environment of social and institutional policy. His work can be classified into two 'Marks'. In Mark I, which is his early work, the entrepreneur is seen as the driver of innovation. The entrepreneur is held responsible for changes in the economic environment and development, technical change and technical knowledge are seen as completely exogenous here. In Mark II, his later work, the company that performs and uses technical change is seen as the main source of innovation. In Mark II, he revises his vision on the role of technical change, making it an action on itself in the economic process, influencing the competitive advantage of the competitor and the position in the market. Schumpeter specifically indicates competition and organizational size as indicators for innovative activity. (Grupp, 1998; Schumpeter, 1942). Arrow (1962) argues that learning by doing lies at the heart of innovation. He directly incorporates innovation in his competition theory developed in 1962 (Arrow, 1962; Grupp, 1998). The results show that there is a higher incentive to innovate in a competitive environment than in a monopoly. In his institutional theory, Freeman emphasizes in his innovation model the industrial, and professionally managed innovative activity and the role of scientific knowledge in this. He stresses the importance of institutionalized research teams working on structured R&D projects, as well as changes like increasing complexity and high developed system and network techniques (Freeman, 1982; Grupp, 1998). From here, innovation theory

has developed and discusses, among others, successes and failures of innovations, the relation between innovations, technology and economic growth, industrial organization and trade issues.

So far, innovation research has mainly focused on the private sector, not much is known about innovations in the private sector yet. In the private sector, the distinction is made between product, process and organizational innovations. The term product innovation is used for new products or substantial improvements in the already existing product. Product innovation is targeted at the external customer and mostly very market-driven. For firms the revenues of product innovation are very explicit and easily observable (Grupp, 1998; Wong et al., 2008). Process innovation concerns a new or substantially improved method or production process. Process innovations are often very implicit and do not immediately show the reduction in costs or a larger production. Process innovation has an internal focus, where the 'customers' are users within the organization. The knowledge affiliated to process innovations is most of the times very complex, difficult to comprehend and also less observable. The innovation is more organization specific and therefore hard to copy. It is almost impossible to exactly copy a process innovation, there are always some organization specific adjustments needed (Grupp, 1998; Stoneman, 2001; Wong et al., 2008). Product and process innovation can be very closely linked, when for example a process innovation is aiming at eventually producing a new product. It is furthermore possible that a process innovation in the one sector is a product innovation in another sector (Grupp, 1998). Lastly, organizational innovations are related to the organizational issues of the company.

The second part of this section is related to the literature on the diffusion of innovations. One of the literature on diffusion of innovations is by Rogers (2003). Rogers's book describes the complete process of the diffusion of innovations: from the elements of diffusion to the innovation-decision process, the rate of adoption, and adopter categories. Rogers also describes the diffusion process in a figure (based on the work of Tarde, 1903) which shows the innovation path (Figure 4-1). Figure 4-1 shows that innovations spread quite slowly at the beginning. The information on

innovations is scarce and many uncertainties and risks exist. At some point in time the innovation has become well known and the diffusion accelerates. By then, most of the potential users have applied the innovation, and the diffusion rate slows down again. Stoneman (2001) also describes the S-shaped diffusion curve in his book on technological diffusion. The S-shaped diffusion curve is often referred to in the literature (see, e.g., Jarne et al., 2007; Sonis, 2009). We also use the S-shaped diffusion curve for comparison. Other theories on diffusion of innovation are the network theory, which states that networks are the locus of innovation in high-tech companies (see for example Powell et al., 1996) and human capital formation theory, in which HRM activities are directly linked to the innovativeness of a company (e.g. Smith et al., 2011).

Figure 4-1 – Diffusion of innovations (Rogers, 2003)

The literature on the determinants of innovation mostly covers the private sectors. Many authors have studied the influence of organizational size on innovativeness,

and mostly found positive relationships (see, e.g., Bettencourt et al., 2007; Camisón-Zornoza et al., 2004; Damanpour, 1990; Morone & Testa, 2008; Shefer & Frenkel, 2005). Urbanization is another factor that is found to influence innovativeness. Johansson and Forslund (2008, pp. 58-59), for example, address the effect of urbanization and diversity on innovativeness. Furthermore, Boyne et al. (2005) find that urban areas are more diverse and have easier access to innovations. However, the hypothesis of Damanpour and Schneider (2006) that there is a positive relation between urbanization and innovativeness was not confirmed by their analysis.

Aghion et al. (2005) find that more competition increases innovations (as was already argued by Schumpeter and Arrows), up to a certain point, as do Carlin et al. (2004). Tang (2006) shows that the relation between competition and innovativeness depends on the type of innovation and the perception of competition. Many other studies in the private sector look at the factors which influence innovations but these factors are not so common for the educational sector; for example, see the chapters in Van Beers et al. (2008). In these chapters, the authors study different aspects of the determinants of innovation in the private sector. HRM, knowledge of employees and employee satisfaction are found to be the main determinants of private sector innovation. Furthermore, they show that R&D activities, the financial status of a company and internationalization are also important factors. Except for the financial status, these factors do not generally apply in education, and are therefore not further discussed.

For the educational sector, one of the few studies on the determinants of innovations, the study by Könings et al. (2007), shows that teachers' characteristics and the way of teaching influence the innovativeness of a school. Furthermore, Lubienski (2003) finds competition to be a determinant of innovations, and Ghaith and Yaghi (1997) find that teachers' attitudes are determinants too. Lastly, Hermans et al. (2008) point at teachers' beliefs as determinants of innovations. Based on this literature in the private sector and in education, teachers' characteristics, competition, urbanization, and school size are expected to be the main determinants in educational innovation.

Few papers address the effect of aggregated innovations in education and their impacts. Technological change or innovation does not necessarily mean positive change. As Grupp (1998) argues, technological change is not always synonymous with improvement; it would be necessary to add to its definition a word such as 'better' or 'new' so as only to include positive innovation. Nonetheless, innovation is assumed to have a positive effect on educational performance (Centre for Educational Research and Innovation, 2008). Many studies investigate the effect of single innovations, but these studies mostly take place at the micro-level at a school, often in a qualitative way (e.g. Dobbins (2009) and Lubienski (2003)). A few studies on single innovations are the following. For example, there is Machin et al. (2007), on the impact of ICT investment and changes on primary school performance. Using two-stage regression analysis, they found a positive relationship between ICT investment and some performance indicators. Kirschner et al. (2004) examined success factors of educational innovations and arrived at a list of instruments which could determine the success of an innovative education project. Other studies on single innovations are by Atkinson et al. (2009) and Kingdon and Teal (2007). Relevant Dutch research on educational innovations appears in Hofman et al. (2007), who give an overview of innovations in the Dutch educational sector, study the effects of innovations, and describe studies on innovations in education. They develop a research set-up to measure the influence of educational innovations. Their results do not provide empirical evidence on the effect of innovations, but do provide a good overview of studies in educational innovation and identify the necessary conditions for examining the effects of innovation.

However, quantitative proof of the effect of innovations aggregated at the school level is still lacking. The latter issue will be discussed in Chapters 5 and 6.

4.3 Data

This chapter also uses the panel data on Dutch secondary schools covering the school years 2002/03 to 2007/08, as described in Chapter 3. This chapter distinguishes three types of data: school characteristics, environmental characteristics, and innovation data. The school characteristics and environmental characteristics were explained and discussed in Chapter 3. The innovation data were gathered by means of a questionnaire.

The innovation questionnaire was sent to the management of 593 Dutch secondary schools. The questionnaire consists of 132 listed innovations and some general questions. The list of innovations was composed very carefully based on innovations in Dutch education, interviews with school directors from 15 Dutch secondary schools and previous research on innovations in education in the Netherlands. Examples of these studies are the Innovation Monitor (Busman et al., 2007; Busman et al., 2006; Klein et al., 2008) and the reports of the project *durven delen doen* (Waslander, 2007; Waslander & van der Weide, 2009). On the questionnaire, school directors were asked to indicate whether the school had introduced a specific innovation, and if so in what year. There was also room to mention other innovations that were not incorporated in the list. In the end, 155 school directors filled out the questionnaire, of which 153 remained after a data check and data manipulation.

Table 4-1 – Response among Dutch secondary schools

School size and type	Total no. schools	Response (n)	Response (per cent)
<i>School size</i>			
Less than or equal to 1500 pupils	332	90	27
More than 1,500 pupils	261	63	24
<i>Type of school</i>			
PRO	71	18	25
VMBO	58	16	28
PRO/VMBO/HAVO/VWO	53	21	40
VMBO/HAVO/VWO	319	75	24
HAVO/VWO	38	8	21
VWO	37	11	30
Other	17	4	24
Total	593	153	26

Table 4-1 shows the response analysis of the sample compared with the entire population. First, a distinction is made between schools with less than 1,500 pupils and schools that exceed 1,500 pupils. Second, the table shows a comparison based on school type. Table 4-1 shows that there is quite an even distribution with respect to school size. With regard to school type there is a slight overrepresentation of PRO/VMBO/HAVO/VWO-schools, and a slight underrepresentation of HAVO/VWO-schools in the sample.

The next step is to test for the representativeness of the sample using logit analysis. A logit model is a binary choice model which predicts whether the chance of having a specific outcome is dependent on certain characteristics. The logit analysis suits the data very well, since this analysis makes it possible to test whether the chance of being present in the sample is dependent on school characteristics and environmental variables. The explanatory variables used in the analysis are the number of locations, the size of the school represented by the number of pupils, the number of schools in the governing body, the province in the Netherlands where the

school is located, whether it is solely a VMBO/PRO school or solely a HAVO/VWO school, the school type, and the degree of urbanization. Table 4-2 shows the odds ratios of the logit analysis. Table 4-2 shows that most of the variables are not significant at the 5 per cent level. This indicates that the chance of being present in the sample is not dependent on those variables. However, there is a significant constant and a significant degree of urbanization. The selection bias probability of the degree of urbanization is 0.99, which makes the sample a little overrepresented by schools in a less urban area. The chi-square of the individual variables is significant at the 1 per cent level. It may therefore be concluded that the sample is almost, but not entirely, representative for the entire population. In the remainder of this research, the inverse probability weighting (IPW) is used to control for this. IPW uses the inverse of the estimated selection bias probability to correct for the bias. For detailed information on IPW, see Sullivan Pepe (2004; Ch. 7) and Wooldridge (Wooldridge, 2002; Ch. 17) .

Table 4-2 – Representativeness of the sample – Logit analysis

Variable	Odds Ratio	P-value
Constant	0.451	0.002
Number of locations	1.015	0.732
School size	1.000	0.782
Number of schools per governing body	1.004	0.419
Province in the Netherlands	0.982	0.213
Solely VMBO or PRO	1.142	0.476
Solely HAVO or VWO	1.003	0.983
School type	1.019	0.511
Degree of urbanization	0.992	0.000

Innovations

For a proper interpretation of the results, this study categorizes the individual innovations. The 132 individual innovations are divided over five clusters of innovations³⁰, 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 Ministry of Education on main topics such as innovation and IT in education (Ministry of Education, 2011a). This thesis 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

³⁰ A complete overview of all individual innovations can be found Appendix C-1.

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. Therefore, the value of the innovation function equals the unweighted sum of the corresponding innovations of T_L :

$$a_h(T_L) = \sum_{i \in L} t_{hi} , \quad (4-1)$$

where $a_h(T_L)$ is the number of innovations present in cluster L for school h ; and t_{hi} is innovation i present in school h . Innovations are estimated as a separate function, which will be explained in the methodology section.

Determinants

Several determinants may be related to a school's choice to introduce a certain innovation. The determinants evaluated in this research are very carefully chosen from the literature but are bounded by the availability of the data: first, factors which the literature shows to influence innovations are used as determinants; next, factors that are suspected to have an influence on the innovativeness of the school are added; lastly, factors that are suspected not to have an influence are also added too, in order to check that they definitely do not have an influence. The degree of urbanization is corrected for the bias by inverse probability weighting.

The following possible determinants are analysed:

- School size (number of pupils);
- Number of locations;
- Number of schools per governing body;
- Competition³¹;
- School type;
- Traditional way of teaching;
- Average experience of the teachers;
- Share of male teachers;
- Degree of urbanization;
- School performance;
- School budget.

Because the determinants include school performance and total expenditures and these data are not for all schools available, the number of observations reduces to 127 schools per year for the determinants analysis. As explained in Chapter 3, school performance can be measured using the national examination grades or student achievement. Furthermore, the percentage of students who graduate in the last grade is often used as a measure for school performance. In this chapter, national examination grades are used. However, the other two school performance measures (student achievement and graduation percentage) are also analysed, as a robustness analysis. School budget is the total yearly budget per school. School budget might influence the rate of innovation because it sets the limitations to the possible spending of schools, and it may be the case that schools with a higher budget spend more on innovation. Schools with a higher performance might invest more in innovations because they have high standards to live up to, and might believe that keeping up with developments and investing in innovations might contribute to that.

³¹ Competition is estimated as the number of schools per 10,000 inhabitants in a municipality.

Table 4-3 – Weighted sample descriptives of the determinants, 2007 (n=126)

	Mean	Standard Deviation	Min	Max
School size	1,710.57	1,029.68	158	5,971
Degree of urbanization	0.45	0.33	0	0.97
Number of locations	2.03	1.54	1	8
Number of schools per governing body	5.99	8.37	1	31
Competition	0.29	0.32	0.02	2.89
Share of male teachers (per cent)	57.06	8.38	28.02	81.05
Average experience of teachers (in years)	19.91	3.04	0.50	26.54
Traditional way of teaching (1=yes, 0=no)	0.78	0.42	0	1
Practical/pre-vocational (1=yes, 0=no)	0.12	0.33	0	1
Higher general/pre-academic (1=yes, 0=no)	0.14	0.35	0	1
Average national examination grade	6.40	0.22	5.89	7.13
School budget (*100,000)	122.79	81.22	14.84	569.75
Number of innovations present	36.94	6.88	5.77	55.62

Table 4-3 provides insight into the data set of 126 schools, and shows the weighted sample descriptives for 2007. The school size varies from 158–5,971 pupils. This is the difference between, for example, a school that only started recently and an old school with several sites in an intensely urban area (e.g. Amsterdam). The sample includes some schools from very urban areas, and also some schools from very rural areas of the Netherlands. Furthermore, the number of locations per school varies from 1 to 8, with an average of 2 and the number of schools per governing body varies from 1 to 31, with an average of 6. The share of male teachers varies from 28–81 per cent. The average experience of the teacher is 20 years. In the sample, 78 per cent of the schools still have a traditional way of teaching. Table 4-3 shows that 12 per cent of the schools offer solely practical training education or pre-vocational education, and that 14 per cent of the schools only offer higher general education or pre-academic education. The remaining 74 per cent are comprehensive schools. Finally, in 2007, an average of 37 innovations were applied by schools.

4.4 Methodology

Descriptive analysis is used to assess which determinants play a role in the innovativeness of a school, using a logistic transformation. For every cluster of innovations a separate regression is used. The simple panel regression takes into account the time variable and includes the determinants, an information variable and an error term. The information variable assembles the information available to the management of the school on new and upcoming innovations. The value of the information variable is determined by the average number of new innovations introduced in a certain year.

This method assumes a linear relationship between the variables, normally distributed variables and homoscedasticity. The method fits the data set rather well, since a simple test showed that there is a normal distribution and homoscedasticity, and there are linear relationships. Furthermore, the regression makes it possible to determine the factors that influence the share of innovations in a certain cluster at a school. In other words, it shows the reason for the differences in innovative power between schools. In this methodology the assumption is made that innovations have a positive effect on education, however, there is no evidence for this.

The share of innovations in cluster c is written as:

$$p_c = \frac{1}{1 + \exp - (\alpha_c + \sum_x \beta_{cx} * Det_x + \gamma_{ct} * info + u_x)} \quad (4-2)$$

where:

- p_c = the share of innovations in cluster c ;
- c = cluster of innovations;
- Det_x = vector of determinants x ;
- $Info$ = information variable;
- u_c = error term of estimation for cluster c ;
- $\alpha_c, \beta_{cx}, \gamma_{ct}$ = estimated parameters.

For estimation convenience, Equation 4-2 is rewritten as:

$$\text{Log}\left(\frac{p_c}{1-p_c}\right) = \alpha_c + \sum_x \beta_{cx} \cdot \text{Det}_x + \gamma_{ct} * \text{info} + u_c (c=1, \dots, C). \quad (4-3)$$

Equation 4-3 shows the regression of the determinants on the share of innovations in a certain cluster. Equation 4-3 is estimated with Least Squares. Since the parameters of this regression are hard to interpret, the marginal probabilities are calculated as well. The marginal probability is shown in Equation 4-4.

$$MC_{cx} = \beta_{cx} \cdot p_c \cdot (1 - p_c). \quad (4-4)$$

4.5 Results

First the results are discussed with respect to the diffusion of innovations. Next, the results of the determinants of innovations are presented.

Diffusion of innovations

Figure 4-2 depicts the diffusion rate of the innovation clusters for the original sample of 153 schools. Figure 4-2 shows how the innovation clusters have developed over the years. The diffusion rate is calculated by comparing the average number of innovations per cluster per year for all the schools with the maximum number of innovations per cluster. For example, if a cluster has a maximum of 15 innovations and the average number of introduced innovations is 6, then the diffusion rate is 40 per cent.

The maximum number of innovations per cluster is determined by the sum of the innovations in the questionnaire and the innovations added to the questionnaire by the schools for a specific cluster. As indicated above, the list of innovations was very carefully composed. Furthermore, the fact that schools could add missing

innovations to the list should lead to a representative maximum number of innovations in the clusters. However, we realize that not all schools participated in the study, and that therefore some very specific innovations, which could have increased the maximum number of innovations for a cluster, could have been missed. It is probable that other innovations do exist in schools that did not take part in the questionnaire. Therefore, in this chapter, these particular innovations could not be taken into account. Also, the lack of specific data on the innovations that schools have introduced means that it is not possible to distinguish between many small innovations (which may be less effective) and a few large innovations which lead to major changes. Inevitably, Figure 4-2 on the diffusion of innovations is solely based on the number of innovations.

Figure 4-2 – Percentage diffusion of innovation clusters, 2002 – 2009 (n=153)

Figure 4-2 shows that not all innovation clusters have had the same diffusion rate during the period 2002-2009. However, we see that the shape of the curves is similar to the one described in literature (Rogers, 2003). The curve for innovations with respect to new courses/profiling is the flattest curve. Furthermore, this cluster also had the lowest growth between 2002 and 2009. Over these years innovations increased by a little over 16 percentage points. There was an increase of 29 percentage points in process innovations; 20 percentage points in pedagogical/didactical innovations; and 28 percentage points in education chain innovations. The innovations with respect to the professionalization of the teachers experienced the highest increase, of about 40 percentage points

Figure 4-2 also shows that innovation with respect to the professionalization of the teachers is the most diffused over schools, with more than 55 per cent of the total innovations in that cluster in 2009. Innovations with respect to new courses/profiling are the least diffused, compared with the other four clusters. One has to keep in mind, however, that not all innovations are applicable to all types of schools. Some innovations are, for example, specifically meant for specialized schools and therefore cannot be introduced by comprehensive schools.

Lastly, Figure 4-2 shows a large increase in the year 2006 for process innovations, educational chain innovations, and innovations with respect to the professionalization of teachers. The results from the questionnaire with the single innovations shows that the process innovations increase is mainly due to IT-innovations, such as digital white boards, digital exams, and an electronic learning environment. Furthermore, the increase in innovations with respect to the professionalization of the teacher seems to be due to increased working in teams, the role of the teacher as a coach, and the electronic competency file. The increase in education chain innovations is mainly due to projects with universities and elementary schools.

Determinants of innovations

To determine the influence of the determinants on the share of innovations within one cluster, first, the regression is estimated (Equation 4-3), after which the marginal effects of the variables in this regression are calculated. Table 4-4 shows the results of the estimation of the marginal effects of all the determinants on the share of applied innovations in a certain cluster and the significance level. For all five regressions it holds that the regression explains more than 95 per cent of the variation, indicating that the list of determinants is rather informative. It can be seen from Table 4-4 that the innovativeness of a school is related to size. This relationship is threefold: influence of school size, an extra location, or belonging to a larger governing body.

With respect to the school size, Table 4-4 illustrates that there is a positive significant effect on all innovation clusters. The effect per 1,000 pupils varies from a 3.6 per cent higher share of new courses/profiling innovations to a 5.5 per cent higher share for professionalization of the teachers' innovations. Furthermore, Table 4-4 shows that there is a negative significant relationship between the number of locations and most of the innovation clusters. This effect varies from a 0.8 to a 2.9 per cent lower share of innovations in a certain cluster. The last size effect is the number of schools in a governing body. The number of schools per governing body is positively related to process innovations (a 0.04 per cent higher share) and education chain innovations (a 0.2 per cent higher share). There is a negative relationship between the number of schools per governing body and the share of innovations in the professionalization of teachers cluster, the new courses cluster, and the pedagogic cluster. Schools belonging to smaller governing bodies, which may have less joint organizational agreements, apparently choose to work more on the professionalization of the teacher, and focus on new courses and the content of these courses.

The results show a positive relation between innovations and size, where size is a combination of the large positive effect of school size and the smaller negative effect of the number of locations. The positive effect of size is expected, as larger

schools may have more possibilities and personnel to invest in innovations and implement them. Furthermore, scale is often related to making an organization more professional. This in turn might be related to having more information available on innovations and the expected costs of innovations. Moreover, larger organizations have more resources, which might enable them to work on more innovations. However, the analysis shows that the school budget is not related to the investment in innovations. One reason for the negative effect of the number of locations could be that the number of pupils per location is smaller when there is more than one location. The negative sign then decreases the positive effect of the school size. Another explanation is that having more locations implies having more decision-making units. This might lead to a more complicated process regarding the decision to innovate, which could slow down the innovation pace. In addition, the results show that a larger governing body has a positive marginal effect on process innovations, and on education chain innovations. A reason for the latter could be that there is extensive collaboration between the governing body and other schools (elementary schools or higher education) and businesses in the area. The governing bodies may be stimulating their schools to make use of this collaboration. It is also likely that a governing body may intensively interfere in the processes at work in the schools and stimulate the schools to have similar processes and learn from each other. The findings are comparable to the findings in the literature, as discussed in section 4.2 (e.g. Bettencourt et al., 2007; Camisón-Zornoza et al., 2004; Shefer & Frenkel, 2005).

Table 4-4 suggests that competition has a negative effect on new courses/profiling innovations, and a positive effect on the other innovation types. For example, the share of innovations of a school in a municipality with 10,000 inhabitants and two competitors in pedagogical/didactical innovations is 5.5 per cent higher than that of a school in the same municipality with only one competitor. This share difference is 3.9 per cent for process innovations, -0.9 per cent for new courses/profiling, 1 per cent for teacher professionalization, and 2.2 per cent for education chain innovations. It was expected that more competition would lead to

more innovation, as schools would have to make an extra effort to attract the students. The positive effect of competition is also confirmed by the literature (e.g. Aghion et al., 2005; Tang, 2006). The degree of urbanization is positively related to the new courses, pedagogic and process innovations clusters, and negatively related to the teacher professionalization and education chain clusters. Schools in a rural area seem to focus more on the professionalization of the teacher, possibly in order to attract more teachers because the predicted teacher shortage is larger in rural than in urban areas (Ministry of Education, 2011c). The positive results of new courses and pedagogic innovations is as expected, as schools in a more urban area have to compete more to attract good students, and may want to distinguish themselves from other schools by investing in different courses or a specific profile, or by teaching in a different way. The finding of urbanization in the literature is mixed (Boyne et al., 2005; Damanpour & Schneider, 2006), which is in line with our mixed findings for the different innovation clusters.

In the regression, three school types are distinguished: the VMBO/PRO schools; the HAVO/VWO schools; and the comprehensive schools. The schools that solely offer VMBO and/or PRO have a lower share of innovations with respect to new courses/profiling and education chain innovations, and a higher share of innovations with respect to the other innovation types. Schools offering solely HAVO and/or VWO are positively related to all innovation clusters. This indicates the diversity between the different school types. HAVO/VWO schools have a higher innovativeness than VMBO/PRO and comprehensive schools for all clusters of innovations. VMBO/PRO schools seem to be working less with other schools in the vicinity, and are working less on new courses and profiling than HAVO/VWO and comprehensive schools. Different school types clearly have a different innovational focus. The diversity of (educational) institutions and the importance of diversity is also stressed by Ostrom (2005).

From Table 4-4 one can see that the way of teaching and the type of teachers also have an effect on innovations. Having a traditional way of teaching is only significantly related to the new courses/profiling and teacher professionalization

innovation clusters. Schools with a traditional way of teaching seem to be less open to innovation and change, and apparently stick more to ‘the way it has always been’. This is an expected result. Furthermore, the results show that the teachers’ experience is positively related to all innovation clusters. The share of male teachers is negatively related to all innovation clusters. The latter seems to indicate that female teachers are more progressive and more open to innovation and change with respect to the way of teaching and the processes used. It might also be the case that females are more willing to adopt new ideas and suggestions from their superior compared with males. Furthermore, it seems that some level of experience is positively related to the number of innovations at a school. A possible reason for this is that more experienced teachers may be better embedded in the system, and have a more stable working situation, whereas teachers with less experience may need some years to fully adapt. More stable teachers may be more amenable to change and more willing to put effort into the introduction of an innovation and making it a success. It seems that some experience is needed to be ready for innovations and major changes. The finding that teacher characteristics matter for innovativeness is confirmed by the study of Könings (2007).

Lastly, as expected, Table 4-4 shows no significant relationship between school performance and investing in innovations and school expenditure and investing in four of the five innovation clusters. The latter is only significant for the new courses/profiling cluster, and the coefficient is very close to zero, indicating that there is no relationship. In Table 4-4, school performance is measured in terms of the national examination grade. However, the same result is found (for all innovation clusters) when graduation percentage or student achievement is used. This result is in line with the finding in the literature that student performance is not associated with innovativeness (Reynolds et al., 2000). Furthermore, most studies do not actually indicate higher-ability students as a determinant of innovations in education, but use the factors described above, for example, competition (Lubienski, 2003), teacher attitudes (Ghaith & Yaghi, 1997), and teacher beliefs (Hermans et al., 2008) as determinants.

Table 4-4 – Marginal effects of relation determinants and innovations, 2002-2007 (n=756)

	New courses / profiling	Pedagogical / didactical	Process	Professionalization of teachers	Education chain
Traditional way of teaching	0.006 ***	-0.001	0.000	0.001 **	0.000
Number of locations	-0.029 ***	-0.008 ***	-0.024 ***	-0.019 ***	-0.014 ***
School size (per 1000 pupils)	0.055 ***	0.036 ***	0.050 ***	0.052 ***	0.042 ***
Number of schools per governing body	-0.001 ***	-0.001 ***	0.000 ***	-0.001 ***	0.002 ***
Competition	-0.009 ***	0.055 ***	0.039 ***	0.010 ***	0.022 ***
Average experience of teachers	0.013 ***	0.014 ***	0.013 ***	0.011 ***	0.011 ***
Degree of urbanization (per 10%)	0.043 ***	0.018 ***	0.001 ***	-0.004 ***	-0.050 ***
Share of male teachers	-0.001 ***	-0.002 ***	-0.001 ***	-0.001 ***	-0.001 ***
Solely VMBO or PRO	-0.028 ***	0.031 ***	0.002 ***	0.060 ***	-0.044 ***
Solely HAVO or VWO	0.014 ***	0.018 ***	0.037 ***	0.053 ***	0.012 ***
School performance (per 0.1)	0.000	0.001	0.003	0.001	0.001
Total school expenditures (per 1000 Euro)	0.000 **	0.000	0.000	0.000	0.000
Information	0.002	0.002 ***	0.004 ***	0.001 **	0.002
Log-likelihood	1108.89	102.04	554.173	1673.54	1327.21
R-squared	96.8	96.0	89.0	99.2	98.7

Please note that there is no marginal effect of the constant.

*** = significant at the 1% level; ** = significant at the 5% level.

4.6 Conclusion

This chapter has investigated the diffusion of innovations and the determinants of innovations in Dutch secondary education. A questionnaire was sent out to gather information on innovations at secondary schools and relate these to existing school and environmental information. In this study, five clusters of innovations were distinguished, namely: new courses/profiling innovations; pedagogical/didactical innovations; process innovations; professionalization of the teachers innovations; and education chain innovations. This chapter showed the diffusion path and investigates the determinants of all five clusters.

This is a rather unique data set which gives information on innovations present in secondary schools. Furthermore, information is included on the characteristics of the school and the environment. All this together made it possible to calculate the odds ratio of the share of innovations per cluster. Then a logistic transformation was used to regress the determinants on the share of applied innovations in a certain cluster. The results show that scale, competition, school type, teacher experience, and the way of teaching are the main determinants of innovations. Larger schools, in a more competitive environment with a progressive way of teaching seem to innovate more. The influence of school type and degree of urbanization is mixed. Moreover, school performance and school budget are not related to the investment decision in innovations.

One weakness of this chapter is the list of determinants. Other determinants might also influence the decision to innovate (e.g. governing body characteristics). However, this information could not be included because there are no data available on these determinants. Another weakness might be that we do not know whether some determinants could be influenced by the amount of innovations (e.g. traditional way of teaching) instead of the other way around.

5 The Relationship between Innovations and School Productivity³²

“Our future growth relies on competitiveness and innovation, skills and productivity... and these in turn rely on the education of our people”

- Julia Gillard -

5.1 Introduction

Education is considered one of the most critical investments for the future of an individual and a country. Investment in knowledge, skills and capacities is relevant for individuals, as well as for societies. Inevitably, the incentives for a country to invest in education are high (OECD, 2009). However, countries have a budget for public sector expenditures, and it is impossible to invest unlimited resources in education. As expectations on the quality and productivity of educational organizations are high, the question how to meet these expectations has been a topic of political debate. The discussion has continued on how to generate the

³² This chapter is based on 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.

highest school productivity with the resources available (e.g. Grosskopf et al., 1997), as education is one of the most important aspects for the success of a country and for reducing inequality (Barro, 1991; Council of the European Union, 2007; Hanushek & Woessmann, 2009). School productivity is mostly measured by the school expenditures and student performance (e.g. Grosskopf et al., 1997).

Recently, the importance and possible role of (educational) innovations for school productivity has been stressed (Centre for Educational Research and Innovation, 2008) and investments in educational innovations have increased in many OECD countries. Furthermore, the Centre for Educational Research and Innovation of the OECD reports that investments in educational innovations have increased in many OECD countries. Indeed, over recent years, innovation in education has been stimulated tremendously by local authorities, which are acting on the innovative and educational goals of, for example, the Lisbon strategy of the European Union (European Union, 2009; Rodrigues, 2009). The Dutch government, for example, has several subsidy programmes to stimulate innovation not only in general but also in a specific field, such as language and maths. These subsidies have been introduced to stimulate innovation in education (Ministry of Education, 2009, 2010; VO-Raad, 2007). In sum, innovation is seen as a potential factor which influences and relates to school productivity. Innovations might influence the composition of the inputs and outputs because schools might need fewer teachers if they work more with computers, but might need more teachers if they focus more on coaching individual students. With respect to outputs, schools might attract more (or less) students if they innovate more in a certain direction, or might influence the performance of their students by focusing on different innovations.

Many authors have investigated the research framework of innovations in education. These studies mostly analyse the effects or relationships of a single innovation at the micro level at a school, or analyse innovations in a qualitative way (see, e.g., Dobbins, 2009; Lubienski, 2003). For instance, with respect to studies on single innovations at the micro-level, there are some quantitative studies on performance pay for a teacher, which is considered one of the innovations that

5. The Relationship Between Innovations and School Productivity

Dutch schools could apply. These studies, for example, by Atkinson et al. (2009), Kingdon and Teal (2007), Dee and Keys (2004) and Lavy (2009), in general draw the conclusion that a performance-related pay scheme increases test scores and value added, although this does not hold for all types of schools. There are several studies on individual innovations such as the impact of IT use in education, but there is no consensus in these studies on the effects of IT use (e.g. Angrist & Lavy, 2002; Leuven et al., 2007) and other individual innovations such as interventions at the tune of transition between primary and secondary or secondary and tertiary education, which seem to have a positive effect on student performance (e.g. Valentine et al., 2009). Lastly, many studies use teacher professionalization innovations and positively relate these to school and student performance (e.g. Clotfelter et al., 2007; Miller et al., 2008; Rivkin et al., 2005). However, most of the studies on innovations in education only look at the micro-level, and do not use a productivity framework for the analysis, whereas our study looks at the meso-level (all innovations at a school) in a productivity framework. Furthermore, many studies on innovations in education do not look for quantitative evidence on the importance or effects of innovations at a macro-level. These studies analyse innovations in education in a qualitative way (see, e.g., Centre for Educational Research and Innovation, 2008; Orozco-Gómez, 2006).

Numerous other studies have investigated the efficiency and productivity of education. Most of these studies analyse school productivity in a quantitative way, using governance variables, school characteristics and environmental characteristics as explanatory variables (see, e.g., Grosskopf et al., 1997; Mayston, 2003). Many of these studies focus on the allocation of resources and economies of scale and scope. Studies focusing on the allocation of resources in primary and secondary education are, for example, those of Ruggiero (2007), Cordero-Ferrera et al. (2008) and Conroy and Arugea (2008). Ruggiero (2007) examines the efficiency of public schools in the US, and concludes that reallocation of resources can improve their efficiency score, although it is not mentioned how these resources should be reallocated. Conroy and

Arugea (2008) also conclude, for the US, that primary schools do not maximize their efficiency, which is why reallocation of resources is necessary. Cordero-Ferrera et al. (2008) conclude that efficiency scores vary tremendously amongst schools in Spain. Other studies focus on the allocation of resources in higher education, and mostly conclude that departments and research groups can be more efficient in reallocating their resources (see, e.g., Abbott & Doucouliagos, 2009; Cherchye & Vanden Abeele, 2005; Gimenez & Martinez, 2005; Groot & García-Valderrama, 2006; Stevens, 2005; Tauer et al., 2007). Studies focusing on economies of scale and scope mostly conclude that schools can perform better by improving their scale (e.g. Agasisti & Salerno, 2007; Barnett et al., 2002; Chakraborty et al., 2000; Johnes et al., 2005; NG & Li, 2009).

Other studies focus on specific factors that influence school productivity, such as competition, IT, privatization, teacher quality and teacher characteristics, and urbanization and political composition. Studies focusing on competition mostly conclude that more competition leads to higher productivity and efficiency (Abbott & Doucouliagos, 2009; Millimet & Collier, 2008; Noailly & Koning, 2009), whereas studies focusing on privatization find that public schools are in most cases performing better than private schools, possibly due to the different role that the type of schools play in the market (Barbetta & Turati, 2003; Cokgezen, 2009). External factors such as the political composition of a municipality influence school productivity as well, as Borge and Naper (2006) and Waldo (2007) demonstrate in their research, in which they show that municipalities where socialists are in charge of education are associated with a lower efficiency. Naper (2010) finds that schools in urban environments are more efficient than schools in rural environments; an explanation for this finding is not given. The studies that come closest to including innovations in the school productivity model are the studies that focus on the factor IT in relation to productivity. Sosin et al. (2004), for example, conclude that IT has a small but positive effect on the performance of students, whereas Dawson et al. (2009) find no effect of technology on the quality and reputation of universities.

However, in the studies on technology and productivity there is no general agreement on the effect of IT/technology on productivity in education.

As the literature described above illustrates, the research framework of productivity in education has not yet included explicitly-defined innovations on a large scale in the analysis. Furthermore, the research framework on innovations in education hardly gives any (quantitative) evidence at the aggregate level. Almost all studies use a proxy for innovations and changes in the school, whereas hardly any study includes data on explicitly-defined innovations.

This chapter uses quantitative analysis and a budget-constrained indirect output distance function model, including specific data on innovations, to investigate the relationship between innovations and school productivity. This chapter follows the same approach as Blank and van Hult (2009) for analysing the relationship between innovations and the productivity of Dutch hospitals. They showed that the use of innovation variables in a productivity analysis is by far to be preferred to the commonly used time trend as a proxy for technical change. Blank and van Hult (2009) used log likelihood ratios to compare models that have specific innovation information with models which have only a time trend, and show significant differences.

The main aim of this chapter is to research the nature of technical change and also to explore the relation between innovations and school productivity. The nature of technical change refers to the possible different relationship of innovations to school productivity, to cost shares of resources, or to student performance.

The contribution of this chapter to the already-existing literature is that it uses explicitly-defined innovation data, and confirms the earlier findings of Blank and van Hult (2009) that models which include specific innovation variables data provide far better estimates than models which use a time trend. Another contribution of this chapter is that we show empirically that a relationship exists between the clusters of innovation, school productivity and student performance. Most of these findings can be added to the literature. However, it is important to

note that this study does not attempt to establish a causal relationship between innovations and productivity.

The remainder of this chapter is structured as follows: Section 5.2 contains the literature on innovations in education. Section 5.3 describes the data and data sources, and Section 5.4 explains the methodology and productivity model. Section 5.5 first presents the results with respect to the comparison of the different models including innovations, and then presents the results of the final model. The chapter ends with some concluding remarks in Section 5.6.

5.2 Literature Review

New courses/profiling innovations

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 quite considerable freedom in the design and organization of courses, such that they can profile themselves towards a specific target group, such as culture schools, sports schools, or schools for high-ability students. Several studies exist with respect to curriculum changes, and they mainly describe positive results. For example, Choudhury et al. (2008) study the effect of a structural curriculum change in the programme of a professional course on property/casualty insurance in the US on the time students use for the course. This output variable is quite comparable to our output variable, as the measure of student performance that we use also represents the duration of education, among other things (see data section). Choudhury et al. use a pre- and post-change time framework for their analysis, and find an effect of the curriculum change, but they also find that it is contingent on the level of education and gender. A second example of this relationship is from Niedermier et al. (2010), who study the effect of curriculum change in a psychiatry course on the exam performance of students. In this study the performance of students that followed the new curriculum are compared with the national data on examination

5. The Relationship Between Innovations and School Productivity

grades in psychiatry. They find that the curriculum change resulted in above-average and sustained improved performance by the students. Lastly, a meta-analysis by Shymansky et al. (1983) assesses the effect of a new science curriculum on student performance and reveals very positive performance of students in the new curriculum and even finds that students in the new curriculum have a performance which could be up to 60 per cent higher than that of students who followed the traditional curriculum.

Pedagogic/didactic innovations

The pedagogic/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 are forcing schools to alter their pedagogical and didactical focus. Innovations that schools are working on, are, for example, using different teaching methods (e.g. peer counselling, only group work, or project-based learning), or involving specialists such as a remedial teacher or a speech specialist. There are only few previous studies regarding the effect of pedagogical changes on educational performance. The earliest of these is 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 et al. (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 does seem to be a positive effect on educational performance.

Process innovations

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 are introducing more extensive IT use, digital learning materials, laptops for all students or digital white boards (see, e.g., Beauchamp, 2004). Furthermore, some schools are constructing a new building where they can incorporate novel pedagogical strategies (e.g. by replacing all classrooms by group and discussion rooms). Even though schools are making significant investments in process innovations, the evidence of the influence of IT use on educational performance is mixed, and therefore, inconsistent results can be found in the literature (see, e.g., Kirkpatrick & Cuban, 1998). In order to cope with this inconsistency, we discuss all three views (positive relationship, negative relationship, no relationship at all) that are expressed in the literature. First, there is the group of studies that find no significant effects of IT on educational performance. Among these are the studies by Dawson et al. (2009), who find no effect of IT on the quality and reputation of universities, and Goolsbee and Guryan (2006), who find no impact of Internet connections on student performance. Lastly, there is the study of Rouse and Krueger (2004), who also do not find any relationship between the use of a specific instructional computer program (FFW) and student performance. An argument in favour of these findings is that schools need to invest in IT to keep up with worldwide developments, in order to keep students at the same achievement level as before, because they might perform worse if schools do not keep up with educational developments, as this might demotivate students.

Second, there is the group of studies that find positive relations between IT and educational performance. Among these studies is that of Machin et al. (2007), who use an instrumental variable approach to identify the causal impact of IT expenditures on the performance of students, and the study from Punie et al. (2006), who conclude that there is evidence that IT use improves educational performance, although IT use at home accounts for a major part of that relationship. Sosin et al. (2004) conclude that IT has a small, but positive, effect on the performance of students. An overview which includes both studies that do find an effect and studies that do not find an effect is presented in Youssef and Dahmani

(2008). The main explanations for these findings are that IT has an impact on student and teacher characteristics, and in that way influences student performance.

Third, there is the group of studies that find a negative relationship. Studies in this group come, for example, from Leuven et al. (2007), who look at the effect of computer subsidies on performance. Their results indicate that there is even a negative effect. Angrist and Lavy (2002) also find a negative relation between IT and school performance. The main explanation for these findings is that the money available for computers might be being spent on other purposes.

Professionalization of the teacher innovations

The fourth innovation cluster contains all innovations that are related to teachers. Due to the teacher shortage in Western countries (European Trade Union Committee for Education, 2011), governments are attempting to increase the attractiveness of the teaching profession. Schools add to this by introducing specific innovations for teachers. The previous literature considered the relationship between, on the one hand, an innovative approach to teachers' pay, teacher quality or teacher development and training and, on the other hand, student performance. These studies mainly find positive relations. For example, Atkinson et al. (2009), Kingdon and Teal (2007), Dee and Keys (2004) and Lavy (2009) conclude that a performance-related pay scheme increases test scores and value added, although this does not hold for all types of schools. Furthermore, Clotfelter et al. (2007) find that teacher experience, teacher test scores, and regular licensing of teachers all have a positive effect on both reading and maths scores of students, with a larger impact for the latter. Croninger et al. (2007) find that there is a positive effect of teachers' degree type and experience on the reading scores of students.

Education chain innovations

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, with the community. The former consists of, for example, primary

school students who take introductory classes at the secondary school, students in the graduation year of secondary education who take some lectures at university, and upper secondary education students who find an internship at a business with which the school has arrangements. The latter is represented by, for example, community schools and extracurricular activities. The 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 relationships described above can be found in literature. Most studies in the literature find a positive relationship between education chain (transition) innovations and student performance. Concerning the transition from primary to secondary education, there is the work of Yadav (2010), who studies the effect of mentoring on the transition to secondary school for at-risk children. The author uses a mentoring intervention and finds positive changes in the output measures, and concludes that there is a positive effect of the mentoring intervention for these students. Although this study does not use specific student performance as an output value, it nevertheless shows the relevance of specific measures to improve the transition.

Regarding the transition from secondary to higher education, for example, Bragg and Ruud (2007) study the impact of specific career and technical education transition programmes on student outcomes. They compare individuals who participated in the programme with people who did not participate, and they find a positive effect of the transition programme on a reading test, but no evidence was found for the maths test. Furthermore, Valentine et al. (2009) carried out a systematic literature review, and found 19 studies of interventions in order to improve the transition and thereby the performance of students. The review concludes, based on the 19 studies, that there is small but potentially important effect on the short-term grades of college students. However, they also mention that not all of the included studies use good methodologies, and that some other studies have very weak interventions. Nonetheless, they conclude that the studies that did

use comprehensive interventions and correct methodologies find more effective results than the other studies. A third study by the US Department of Education (2010), which contains a summary of the findings from two literature reviews, mainly concludes that most experiments in this field are carried out inadequately, which makes it impossible to speak about the effectiveness of these experiments. The study calls for better evidence-based research in this respect.

5.3 Data

This chapter uses a balanced panel data set consisting of 126 Dutch secondary schools covering the school years 2002/03 to 2007/08, including a total of 756 observations. This chapter distinguishes four types of data: student achievement (production indicators); resources data; environmental and school variables; and innovation data. The descriptive statistics of the student achievement, resources and innovation data are presented in Table 5-1. The descriptive statistics of the environmental and school variables are presented in Table 4-3 of the previous chapter.

Student achievement

As described in the data section of Chapter 3, there are two official performance measures of students in Dutch secondary education which are comparable across schools: the national examination grades and students' achievement. In this study, we are interested in the relationship between innovations and student achievement. We have only six years of data, three years of which we lose due to the use of lagged variables, so we only have innovations data left for three years. Furthermore, all students from all educational types need more than three years in secondary education in order to take their national examinations, and these grades are only available for graduating students, for obvious reasons. Student achievement, on the other hand, is calculated every year, and includes all the students at the school. Student achievement seems a better measure to determine the possible influence of

innovations on the students' performance in the three years of data that we have available.

Therefore, in this study we use a corrected value of student achievement for lower secondary education and for all types of upper secondary education, multiplied by the number of students, as a production measure, which is calculated by the Dutch education inspectorate every year (see Chapter 3).

Table 5-1 shows student achievement in 2007 for all groups. The achievement for lower secondary education is on average 1.02, implying that in lower secondary education most students attend secondary education at the level that was advised in primary education. Some students graduate from a higher level of education compared with the level recommended in primary school, and some students graduate from a lower level of education, but in the end, at the school level this seems to cancel out. Furthermore, repeating the year does not happen very often in lower secondary education. However, in the upper part of secondary education it is very common (see Education Inspectorate, 2011, Figure 8.5a). For VMBO-bk and VMBO-gt the achievement scores are .62 and .74, respectively, for HAVO the return is .53, and for VWO the return is around .62.

Resources

In this chapter we use the same homogeneous groups of resources as in Chapter 3: management personnel; teaching personnel; support personnel; and material supplies. For the personnel groups, data on the number of full time equivalents (FTEs), costs, and prices are available. In particular, the costs correspond to the costs per FTE per year. The prices of the personnel groups and materials are the same as in Chapter 3.

Table 5-1 presents the descriptive statistics associated with the resources data in 2007. As Table 5-1 shows, the total average number of FTEs per school equals 165, consisting for 80 per cent of teachers, 15 per cent of support personnel, and the remaining 5 per cent of managing personnel. The total costs of a school are about

12.3 million euros per year. Most of the costs are spent on teaching personnel, followed by materials, managing personnel, and, finally, support personnel.

Control and school variables

The control variables relate to (fixed) characteristics of the school environment. These variables include the number of locations; the number of schools per governing body; school size (number of pupils); competition; school type (solely VMBO, solely HAVO/VWO or comprehensive schools); whether the school has a traditional way of teaching; teacher experience; share of male teachers; and degree of urbanization. Chapter 4 showed that the variables described above are related to a school's decision to innovate (Haelermans, 2010). The number of locations, for example, reflects the geographical spread of a school over different sites which may lead to additional costs due to extra travel time (for teachers) and scale effects. The number of schools per governing body may reflect the balance of benefits of participation in a school's network, the benefits of shared services and the negative effects of a centralized bureaucracy. Furthermore, previous research has shown that competition is also related to productivity (Abbott & Doucouliagos, 2009; Millimet & Collier, 2008). Although one may argue that teacher experience and the share of male teachers are part of the managerial decision making of a school, these are still fixed variables in the short and medium run, since employees are highly protected by law and schools cannot just hire someone else.

Innovations

The innovation data are exactly the same as the data presented in Chapter 4. Table 5-1 shows that, in 2007, the average number of innovations per school was 36, with an average of 4 for new courses/profiling, an average of 13 for pedagogical/didactical innovations, 13 for process innovations, 6 for innovations with respect to the professionalization of the teacher, and 4 for education chain innovations.

School type, year and region dummy fixed effects

We started from the simple model which does not take into account the panel structure of our data. In order to incorporate this structure, we could use either the fixed effects approach or the first differences approach. A drawback of the first differences approach is that it removes much of the useful information that can be derived from the data. Furthermore, a test for serial correlation in the idiosyncratic errors for the simple model shows that the hypothesis for no serial correlation cannot be rejected (Greene, 2008). The first-order LM test ($p=0.28$) for serial correlation cannot be rejected, implying that using first differences is not necessarily advisable. Therefore, there is no reason to choose the first differences method over the fixed effects method. With fixed effects, it is possible to incorporate the heterogeneity that exists between schools, school types and years. However, because of the number of observations it is not possible to include standard fixed effects per school. Therefore, we include school type, region and year dummies in the model, as a substitute for regular fixed effects per school. As the simple model (log likelihood (logl)=2962.48, degrees of freedom (df)=77) is nested in the school type, region and year dummy fixed effects model (logl=3086.98 df=91), we can also test these models against each other using the log likelihood ratios. These ratios show that the school type, region and year dummy fixed effects models are significantly better than the simple models. We tested the interactions between regions and years and between school type and years as fixed effects in the model. We tested the use of interaction terms compared with single terms, similar to what was done in Chapter 3, and found that the model using single terms outperforms the model using interaction terms.

5. The Relationship Between Innovations and School Productivity

Table 5-1 – Descriptive statistics, Dutch secondary education 2007 (n=126)

	<i>Mean</i>	<i>Std Dev</i>	<i>Minimum</i>	<i>Maximum</i>
Achievement lower secondary education	1.02	0.06	0.79	1.18
Achievement VMBO-bk	0.62	0.41	0.58	1
Achievement VMBO-gt	0.74	0.32	0.62	0.98
Achievement HAVO	0.53	0.28	0.47	0.88
Achievement VWO	0.62	0.24	0.43	0.91
Number of students lower secondary education	708	440	79	2535
Number of students practical education	18	49	0	273
Number of students VMBO-bk	183	191	0	814
Number of students VMBO-gt	192	155	0	838
Number of students HAVO	298	212	0	1084
Number of students VWO	310	195	0	866
Number of FTE management personnel	9	7	1	44
Number of FTE teaching personnel	132	86	15	468
Number of FTE support personnel	24	17	3	119
Total FTE	165	106	19	616
Total costs (in 1000 Euro)	12280	8123	1484	53973
Management personnel costs (in 1000 euros)	755	644	77	3239
Teaching personnel costs (in 1000 euros)	8455	5502	962	36123
Support personnel costs (in 1000 euros)	480	362	58	2491
Material costs (in 1000 euros)	2110	1550	313	9628
Number of new courses/profiling innovations	4	1	0	7
Number of pedagogical/didactical innovations	13	2	1	20
Number of process innovations	10	2	0	17
Number of professionalization of the teachers' innovations	6	1	2	9
Number of education chain innovations	4	1	1	6
Total number of innovations in 2007	36	7	6	56

5.4 Methodology

Productivity model

For this chapter we use a similar model to that presented in Chapter 3. However, we include innovations in the model that we use in this chapter. The model used in this chapter is as follows:

$$\ln(IOD) = f(\ln y, \ln\left(\frac{w}{C}\right), \ln z), \quad (5-1)$$

where IOD is the indirect output distance function (specified as a Translog); f is the functional form; y is the vector of outputs (student achievement); w is the vector of prices of used resources; C are the total costs; and z is a vector of control variables (e.g. environmental characteristics and innovations).

According to the described production model, the indirect output distance function is estimated. The model constitutes a system with an indirect output distance function and a number of cost share equations. This model and the cost shares are described in Appendix B1.

The models are estimated as multivariate regression systems using various equations with a joint density, which are assumed to be normally distributed. One share equation in the production function model is eliminated because the shares add up to 1, causing the variance-covariance matrix of the error terms to be singular. In order to estimate the system of equations, consisting of nonlinear multivariate regressions with cross-equation constraints, we can use Seemingly Unrelated Regressions (SUR), Full Information Maximum Likelihood (FIML), and the minimum distance estimator LSQ. SUR is an estimator in which the values of the parameters are determined first, before the variance-covariance matrix is estimated. In both LSQ and FIML the parameters of the model and the variance-covariance matrix are simultaneously iterated. LSQ and FIML both assume that the disturbances of the equations are correlated. However, multivariate regression with LSQ assumes that there are no simultaneity problems with endogenous variables on the right-hand side of the equations, whereas the default of FIML assumes simultaneity. Furthermore, if a discrete Hessian is used for iteration in FIML (which provides

convergence close to the optimum, and can be better than the default), the estimator is the same as the multivariate regression estimator when the disturbances are multivariate normal. In our case, there might be simultaneity problems with endogenous variables (because the error terms of the share equations might be correlated with the indirect output distance function, and vice versa), which is why FIML is preferred. Unfortunately, in spite of many attempts, FIML would not converge. It is very likely that this is because it is a very computationally-demanding technique, and we are estimating a very complex model with many non-linearities. Both the minimum distance estimator LSQ and the Seemingly Unrelated Regressions did converge and give plausible results. More importantly, both estimation methods yield similar results. Because LSQ is theoretically similar to FIML with a discrete Hessian, we choose to use the minimum distance estimator LSQ in the remainder of this chapter. However, this implies adopting the assumption that there are no simultaneity problems with endogenous variables, and that the disturbances are multivariate normal. We are aware that this is most likely not the case for our estimations, but given the complexity of the model, this is the best we can do. For the interpretation of our results, this means that it is very likely that our results are biased. However, due to the complex nature of the model it is practically impossible to determine the direction of the bias. Therefore, the results must be interpreted with caution and previous and future research will have to confirm these results.

Explaining innovations

The decision to adopt, or not adopt, an innovation is likely the result of a number of factors. Chapter 4 analysed the determinants of innovations using the same data set and the same types of innovations (Haelermans, 2010). From Chapter 4, we know that the following determinants influence the number of individual innovations within the clusters of innovations: the number of locations, the number of schools per governing body, school size (number of pupils), competition, school type (solely VMBO, solely HAVO/VWO or comprehensive schools), a traditional way of teaching,

teacher experience, share of male teachers, and degree of urbanization. These are the control and school variables that were described in the data section. In order to control for some endogeneity in the decision to innovate, we correct the starting value of the innovation clusters (the total number of innovations per cluster in a specific year) for these determinants. However, as it is unlikely that the value of the determinants in year t will influence the amount of innovations in year t , we take a 1-year lag of the determinants ($t-1$). By doing so, we assume that the situation in a school in year $t-1$ influences the decision to invest in certain innovation clusters in year t , which seems a plausible assumption. Although it is clear that we by no means intend to establish a causal relationship, and although it is also clear that it is not possible to establish a causal relationship with the data concerned, correcting the value of the innovations clusters for the lagged value of these variables at least creates a more independent and endogenous value of the innovation clusters.

Modelling innovations

As mentioned earlier, the modelling of innovations is of specific interest. The current study measures innovations by using the five clusters of innovations described above which are added to the indirect output distance function.

In this chapter, we assume that innovations only relate to the way the students are educated, the way the educational and administrative process is organised, and the way the school is being managed. It does not assume innovations to influence the choice of school by students and their parents (in the short run), the amount of lump-sum payment distributed by the government, or the teacher population of a school. Furthermore, following the literature by, for example, Gunday et al. (May 2011), Lachenmeier and Rottmann (2011) and Ken et al. (2008), we assume a time lag between the year the innovations were introduced and the time from when they could start to possibly influence school performance³³. To incorporate a sufficient time lag we use the innovation variables lagged by two years. As we have also used a lag of one year for the variables in the innovation function,

³³ This does not imply a causal relationship.

only three years of data remain in this chapter. This hardly changes the estimation results of the model, and also does not change the estimation possibilities. Now 378 observations remain for the productivity analysis.

Comparable to the approach of Blank and van Hulst (2009), in this chapter the nature of innovations is tested. However, Blank and van Hulst compare models which have innovations and a time trend with a model which has only a time trend and a model without any technical change indicator. As we assume that including a time trend in addition to the innovations leads to some of the effects of innovations being captured in the time trend, we take this into account and include only innovations instead of also a combination of the time trend and the innovations. We first distinguish three types of models (Models 1, 2 and 3), each with a different way of including changes over time. Each of these models includes student achievement, resources and their prices, and environmental variables. Furthermore, these models include some time indication. First of all there is the most extended type of model (Model 1) which includes technical change (TC) by including specific information on the five clusters of innovations. Next, we assume the parameters of the innovations to be zero but include a time trend as a proxy for technical change, leading to Model 2. Last, we assume the parameters of both the time trend and the innovations to be zero, leading to Model 3, which does not include a time indication at all. Model type 3 is nested in Model types 1 and 2 and Model type 2 is nested in Model 1. Model types 1 and 2 can have different specifications and different ways of including the innovations and/or time trend in the model. In these model types, the nature of innovations can be specified in four different ways (I-IV) as is shown in Table 5-2.

Table 5-2 – Model specifications (I-IV) for the nature of innovations

	<u><i>Model spec I</i></u>	<u><i>Model spec II</i></u>	<u><i>Model spec III</i></u>	<u><i>Model spec IV</i></u>
	Disembodied Technical Change	Input-biased Technical Change	Output-biased Technical Change	Input/Output - biased Technical Change
Technical change	X	X	X	X
Technical change * resources cost shares		X		X
Technical change * student achievement			X	X

Model specification I represents disembodied technical change. It is also the most parsimonious model in which the innovation index or the time trend are directly included in the model as parameters. Model specification II assumes input-biased innovation: aside from the direct terms, the model also includes cross-terms of innovations with cost shares of resources. Model specification III is an output-biased innovation model and includes aside from the direct terms, cross-terms of innovations with student achievement. Model specification IV is an input- and output-biased model and contains aside from the direct terms, cross-terms of innovations with student achievement, as well as cross-terms of innovations with cost shares of resources. Model specification I is nested in Model specifications II, III and IV, and Model specifications II and III are nested in Model specification IV. We use the likelihood ratio test to compare the different Model types (1-3) and the different model specifications (I-IV) for Model types 1 and 2, and test whether the inclusion of extra variables (in this case specifically on innovations) improves the model. The likelihood ratio is calculated as follows:

$$D = -2 * (\ln(\text{likelihood_null_model}) - \ln(\text{likelihood_alternative_model})) . \quad (5-2)$$

The null model is always the most basic model, whereas the alternative model includes more variables. The measure of D is compared with the 5 per cent critical value of the Chi-Square Distribution, using the difference in degrees of freedom between the two compared models. If D exceeds the 5 per cent critical value, the alternative model fits significantly better. The likelihood ratio test can only be used for the comparison of nested models. Therefore, we first compare Model type 3 with Models 1 and 2, to find out whether including variation over time is a good choice. If this is the case, we compare Model specifications I, II, III and IV per model type (1 and 2) and compare all the model types with each other.

5.5 Results

This section presents the results of the likelihood ratio analysis and the results of the model with the best fit. To determine which model this is, we use Table 5-3. Table 5-3 gives the results of the likelihoods. Using a critical, two-sided, value of 0.025 and Formula 5-2, the likelihood ratios can be calculated from Table 5-3.

Table 5-3 –Log likelihoods and degrees of freedom of the various models (n=378)

	No time indication	Disembodied TC Model I	Input-biased TC Model II	Output-biased TC Model III	Input/output-biased TC Model IV
Model type 3	3021.35 (df=48)				
Model type 2		3021.89 (df=49)	3022.32 (df=53)	3023.76 (df=54)	3024.81 (df=58)
Model type 1		3058.24 (df=53)	3065.23 (df=57)	3086.98 (df=58)	3090.42 (df=62)

If we put the numbers from Table 5-3 in Equation 5-2, it can be concluded that in all cases Model type 1 is considerably and significantly better than Model types 2 and 3. This implies that the models including specific information on innovations are performing considerably better than models that only include a time trend to incorporate technical change, or the model that does not incorporate technical change at all. This result has been found in other studies as well, for example by Blank and van Hulst (2009) in the health care sector and by Rahman (1992) for the some of the Canadian manufacturing industries. Thus, comparable to these studies, this chapter shows that for education specific information about technical change improves the fit of the model compared with, for example, just a time trend.

We now turn to Model type 1 to compare the performance of the four model specifications for both types and find the best performing model. First of all, the log likelihoods of Table 5-3 show that the model including disembodied technical change, Model specification I, does not perform better than any of the other models. Second, we see that the most extended model, Model specification IV, does perform significantly better than the input-biased (model specification II), but does not outperform the output-biased (model specification III) model. Although the input- and the output-biased models are not nested, because the output-biased model outperforms the input/output-biased model, which in turn outperforms the input-biased model, we conclude that the output-biased model is to be preferred.

Therefore, Model 1-III, which is the model with non-neutral output biased technical change with innovations and cross-terms with student achievement suits the data best. The likelihoods of the models show that the fit of this model is better than that of the others. However, the results and coefficients of all the above-mentioned models are comparable and robust.

From a statistical point of view the non-neutral output biased model fits the data rather well. Almost two-thirds of the parameters are significant at the 5 per cent level, and all cost coefficients are positive. Therefore, input demand is always positive which means that the monotonicity condition is accounted for. The

concavity condition is also accounted for, since almost all ‘own’ elasticity’s of substitution are negative. Furthermore, we did a robustness check. We ran the school type, region and year dummy fixed effects model and the simple model with both normal innovations and lagged innovations. The results of these four options were comparable for all nine model types (as presented in Table 5-3).

The estimates of the relevant parameters of the model with non-neutral output-biased innovations are presented in Table 5-4^{34 35}. Similar to Chapter 3, in the IOD model, a positive (negative) sign indicates an unfavourable (favourable) influence on productivity. Table 5-4 shows that almost all innovations are significantly related to productivity (a1-a5), but only two are significantly related to the student achievement variables (i1-i6). Process innovations, professionalization of teachers and education chain innovations are positively related to productivity, whereas new courses/profiling innovations and pedagogical/didactical innovations are negatively related to productivity. The T-statistic and p-value show that almost all significant estimates are significant at the 5 per cent level. Innovations are significantly related to the achievement of lower school and practical education students only. We now first discuss the interaction terms, i_1-i_5 , before turning to the interpretation of all the results in total.

The i_1-i_6 parameters measure the relationship between the innovation index and achievement of the five student types. Interpreting the innovation index is not always clear, since it is an unweighted sum of the underlying innovations. Therefore, the separate relationship of the innovations clusters to the production indicators is presented by the following formula:

$$i_x a_t = i_x * a_t, \quad (5-3)$$

³⁴ The full estimation results of the model can be found in Appendix B-1.

³⁵ A technical efficiency analysis for this model, comparable to the analysis in Chapter 3, shows that the average efficiency score in 2007 was 75.8 per cent with a standard deviation of 0.073.

where $i_y a_t$ is the relationship of innovation cluster t to the student achievement y_m .

The results of Formula 5-3 are presented in Table 5-5. Note that in Table 5 we can interpret the sign as it is presented.

Table 5-4 – Relevant parameter estimates: model with output biased technical change

	<i>Parameter</i>	<i>Estimate</i>	<i>Standard error</i>	<i>T-statistic</i>	<i>P-value</i>
a1	New courses/profiling innovations	0.473	0.067	7.054	[.000]
a2	Pedagogical/didactical innovations	0.263	0.106	2.472	[.013]
a3	Process innovations	-0.092	0.152	-0.608	[.543]
a4	Professionalization of the teachers innovations	-0.942	0.177	-5.314	[.000]
a5	Education chain innovations	-0.118	0.062	-1.893	[.058]
i1	Innovation * achievement lower school	-0.603	0.107	-5.655	[.000]
i2	Innovation * achievement practical education	0.530	0.055	9.586	[.000]
i3	Innovation * achievement VMBO-bk	0.034	0.037	0.932	[.351]
i4	Innovation * achievement VMBO-gt	-0.061	0.076	-0.805	[.421]
i5	Innovation * achievement HAVO	0.137	0.088	1.549	[.121]
i6	Innovation * achievement VWO	-0.036	0.052	-0.696	[.486]

Table 5-5 shows that there are only significant relations for the achievement of lower secondary education and practical education students. The achievement of the two types of VMBO, HAVO and VWO students is not significantly related to innovations. The achievement of lower secondary education students is positively related to process innovations, professionalization of teachers innovations and education chain innovations, and negatively to new courses/profiling, pedagogical/didactical and education chain innovations. These results are comparable to the overall effect of innovations shown in Table 5-4. For practical education and, this is exactly the other way around. This difference could be explained by the different types of students that attend these types of education. .

5. The Relationship Between Innovations and School Productivity

First of all, students in lower secondary education represent all types of students in secondary schools, no matter their level and the type of education they will end up in later in secondary school. Therefore, the group of lower secondary education best presents all students in secondary education. The other five groups are specifications of types of students, of which practical education even is a special case, because it is a rather small group of student and only a small share of schools actually offer this type of education. For an explanation of the differences with respect to the content of the innovations clusters, we think that students in lower secondary education might be more influenced by their teachers and the processes which are used to teach, because they are relatively young and everything is new to them. Education chain innovations are often in particular directed towards first and last year students, which explains why lower secondary students are positively influenced by these innovations. Practical education students are usually the more difficult students, with respect to both learning possibilities and personality. These students might benefit more from a (differentiated and) specific way of teaching and specific profiling of the school than of, for example, a more professionalized teacher.

The positive finding on process innovations might be explained by the fact that students are facilitated by these changes and have an optimal environment to improve their performance. The finding on teachers' professionalization is explained by the large impact of the teacher on the student. If the teacher professionalizes and performs better, this has a positive effect on the attitude and performance of the student. Furthermore, a motivated teacher also motivates his or her students. The finding on education chain innovations can be explained because secondary schools invest in relations with primary education and higher education, which might lead to a selection effect. For primary education, the selection effect might lie in the fact that, because of the good relations between schools, secondary education schools pressure primary schools to lower the level of the type of secondary school recommended for their students. Doing this means that less students will fail, compared with the school level recommended, and that more students will perform above expectations. For the relation with higher education, by allowing the well-

performing students to follow classes at, for example, university, these students stay motivated and interested and will perform better in their final examination than students who are bored.

The overall negative finding on new courses/profiling and pedagogical/didactical innovations can be explained by the fact that schools maybe offer too many different courses and focus too much on having a specific profile (such as sports school, school for highly-gifted students, culture school etc.). They focus on these innovations and do not focus enough on student performance in the main courses. Another explanation is that continuous changes in pedagogy and didactics might confuse the students and make them perform worse instead of better.

To sum up: we find a positive relationship between, on the one hand, process innovations, teachers' professionalization innovations and education chain innovations, and, on the other hand, school productivity. With respect to the positive relationship of process/IT innovations that we find, the literature provides very inconsistent findings concerning this relationship. Although many studies find negative effects or no effects at all, our study corresponds to the findings of the group that find positive effects (e.g. Angrist & Lavy, 2002; Leuven et al., 2007). Finally, the positive finding on teacher professionalization is confirmed by the literature, as it mainly finds positive relations (e.g. Atkinson et al., 2009; Clotfelter et al., 2007). Lastly, our finding on education chain innovations confirms the positive findings of studies by, for example, Bragg and Ruud (2007) and Valentine et al. (2009).

Next, we find a negative relationship between, on the one hand, new courses/profiling innovations and pedagogic/didactic innovations and, on the other hand, school productivity.

The findings of a negative relationship between, on the one hand, changes in the curriculum and new courses, the pedagogical way of teaching and changes in education chain, and transition approaches, and, on the other hand, productivity and

educational performance is different from what we find in literature. The finding on new courses/profiling innovation is, for example, different from the conclusions by Choudhury et al. (2008) and Niedermier et al. (2010). Furthermore, the literature also finds a positive relation between pedagogical/didactical innovations and school performance (e.g. Fuchs & Woessmann, 2007; Queen, 2009).

Productivity development

Figure 5-1 shows the productivity development between 2005 and 2007, using cost shares as inputs and the six types of student achievement as outputs. Figure 5-1 shows that productivity has decreased over these 3 years. This is comparable to what we presented in Chapter 3.

Figure 5-1 – Productivity development between 2005 and 2007

Table 5-5 – Relationship between innovations and student achievement

	<i>Achievement lower secondary education</i>		<i>Achievement practical education</i>		<i>Achievement VMBO-bk</i>		<i>Achievement VMBO-gt</i>		<i>Achievement HAVO</i>		<i>Achievement VWO</i>	
	Estimate	T-value	Estimate	T-value	Estimate	T-value	Estimate	T-value	Estimate	T-value	Estimate	T-value
New courses/profiling innovations Pedagogical/didactical innovations Process innovations Professionalization of the teachers innovations Education chain innovations	-0.285	-4.838	0.251	6.122	0.016	0.929	-0.029	-0.811	0.065	1.672	-0.017	-0.719
	-0.158	-2.349	0.139	2.544	0.009	0.907	-0.016	-0.805	0.036	1.420	-0.009	-0.701
	0.056	0.612	-0.049	-0.611	-0.003	-0.550	0.006	0.500	-0.013	-0.582	0.003	0.466
	0.568	3.988	-0.499	-4.870	-0.032	-0.910	0.058	0.808	-0.129	-1.609	0.034	0.715
	0.071	1.792	-0.062	-1.862	-0.004	-0.819	0.007	0.749	-0.016	-1.251	0.004	0.671

5.6 Conclusion

This chapter has investigated the relationship between innovations and secondary school productivity by a budget-constrained indirect output distance function. This model was applied in a context of maximizing production subject to a budget constraint, similar to what was done in Chapter 3, which fits the case of Dutch secondary education best. We applied different models and tested for the specification of technical change and for the nature of technical change.

Comparing the various types of models showed that the models with explicitly defined innovations provided more reliable estimates than models with only a time trend. This finding is in line with earlier research, for instance, on hospitals (Blank & van Hulst, 2009). Furthermore, the results show that innovations are non-neutral and output-biased, which implies that innovations influence the performance of students, but not the money spent on the different personnel groups and materials. Our results indicate that professionalization of teacher's innovations and education chain innovations are significantly and positively related to productivity and student performance, whereas new courses/profiling innovations, pedagogic/didactic innovations are significantly and negatively related to productivity and student performance. Process innovations are positively related but this finding is not significant. These findings can be explained by the facilitation of better performance of students by IT, more highly motivated professionalized teachers who influence the performance and motivation of their students, and that self-selection of good students into schools that have good relationships with primary schools and higher education. The negative findings might be explained by the fact that schools maybe offer too many different courses and focus too much on having a specific profile. They might focus too much on these innovations and do not focus enough on student performance in the main courses.

Most of these findings are in line with the literature, except for the finding on new courses/profiling and pedagogic/didactic innovations. In the literature, it is found that there is a positive relationship between specific attention for the

transition between primary and secondary education or between secondary and tertiary education and the performance of students (e.g. Valentine et al., 2009). There is much debate in the literature about the relationship between IT use and student performance. There are numerous studies that find a positive relation, numerous studies that find a negative relation, and numerous studies that do not find a relation at all (e.g. Kirkpatrick & Cuban, 1998; Youssef & Dahmani, 2008). The relationship between the professionalization of the teachers and educational performance has been researched in many other studies (e.g. Atkinson et al., 2009; Clotfelter et al., 2007; Croninger et al., 2007), where in most studies a positive relationship is found. With respect to new courses/profiling and pedagogic/didactic innovations, in the literature it is found that there is a positive relationship between changes in the curriculum and educational performance (e.g. Choudhury et al., 2008) and that pedagogical/ didactical changes influence student performance (e.g. Nii & Chin, 1996; Queen, 2009). Unfortunately, these findings cannot be confirmed by our results. However, our findings do not imply that schools should stop investing in these innovations, because we find many studies in the literature with positive relationships between educational performance and new courses/profiling and pedagogic/didactic innovations.

This chapter is a first step towards analysing the relationship between innovations at the school level with educational performance. Although some of the results are promising, more research is needed before decisions are made on whether or not to keep investing in certain types of innovations. Furthermore, we do not have enough information on the innovations to establish a causal relation, something on which it is necessary to base important investment decisions such as these. Therefore, this chapter should just be considered as a first step and should only be used in addition to causal evidence when making decisions on innovations in education.

6 The role of innovations in secondary school efficiency – A conditional efficiency model³⁶

“...cutting our investments in innovation and education is like lightening an overloaded airplane by removing its engine. It may make you feel like you’re flying high at first, but it won’t take long before you feel the impact”

- Barack Obama -

6.1 Introduction

“Evidence-based policy and practice should be the driver of reform in education and training systems.” (European Commission, 2007, p. 8). Despite the attention given to educational innovations, and to evidence-based education research, little is known on the effectiveness of educational innovations. Practical examples of innovations in education are, for example, the change in teaching style, teaching facilities and teaching professionalization.

³⁶ This chapter is based on: C. Haelermans & K. De Witte (2011). The role of innovations in secondary school efficiency – Evidence from a conditional efficiency model. *TIER working paper*.

This chapter 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 correlate with school performance. School performance is measured as the ratio of students' educational achievements to school monetary resources. We take into account both school resources (i.e. the budget constraint that schools are facing), as well as the innovation mix, as we did in Chapter 5. Prior to this thesis, the latter two have never been considered simultaneously. Moreover, the previous literature paid significant attention only to the influence of introducing single innovations, and largely ignored the (existing) mix of innovations. Illustrative studies are those on the use of information technology in schools (e.g. Angrist & 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). As argued in Chapter 5, studying the mix of innovations, compared with one single innovation at a time, better represents the reality of the learning environment in the school, and thereby, its effects on student performance.

Innovation is obviously endogenous to schools. The endogeneity might 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. Parental background may also influence both variables (Houtenville & 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 urbanization and a disadvantaged neighbourhood, 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 (for an extensive discussion of this, see Chapter 4 and Haelermans, 2010). Furthermore, the non-innovation data are collected nationally by the Ministry of

Education. On the other hand, the applied methodology mitigates the influence of atypical 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. However, the previous literature indicated that this is not the case (Reynolds et al., 2000). It has been argued that it is not student ability but competition (Lubienski, 2003), teachers' attitudes (Ghaith & Yaghi, 1997), and teachers' beliefs (Hermans et al., 2008) that can be considered as determinants for innovations. Furthermore, as shown in Chapter 4, school expenditures and student performance are not among the determinants of innovations in education.

As a second contribution (specific to this chapter), 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. The 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 of the present study, 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 & Simar, 2005, 2007; De Witte & Kortelainen, 2008). 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 here, the use of innovations is expected (though not necessarily) to influence the educational attainments (i.e. the output of schools). Previous research has often applied a two-stage model in which the background

variables are regressed on the efficiency scores.³⁷ It is clear that, in a similar setting, the innovations can only explain the efficiency score, but not influence it. The conditional efficiency model avoids this pitfall.

This chapter 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 (see Chapter 4) that we sent out to secondary schools, we distinguish five innovation clusters: (1) new courses/profiling, (2) pedagogic/didactic, (3) process, (4) teacher professionalization, and (5) education chain innovations. The impact of each innovation cluster on school efficiency is examined. In doing so, both endogenous (e.g. size of the school, pupil teacher ratio) and exogenous (e.g. school type and region) differences between schools are accounted for.

Schools innovate in various ways. Some innovations are visible to prospective parents. Illustrative of this 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 teachers' 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 possible other innovations at the school might over- or underestimate this causal effect. This study includes all recent innovations at the school simultaneously. An extensive discussion on the previous literature per innovation cluster can be found in Chapter 5.

The remainder of this chapter is structured as follows: Section 6.2 describes the empirical methodology, and Section 6.3 explains the institutional setting and the data. Section 6.4 presents the results, and relates them to the literature. Section 6.5 relates the findings in this chapter to the findings in Chapter 5, and discusses the

³⁷ Simar and Wilson (2010) extensively criticize the two-step method (see Footnote 5 in Chapter 2).

similarities and differences. The chapter ends with some concluding remarks in Section 6.6.

6.2 Methodology

The 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), and 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, known as 'robust FDH'; Cazals et al., 2002) and exogenous variables (heterogeneity across schools, known as 'robust conditional FDH'; Daraio & Simar, 2005). This conditional efficiency approach limits the reference set of the evaluated units by only comparing like with like. The robust estimates have been shown to possess attractive properties: i.e. they are consistent (i.e. estimate the 'true' inefficiency) (see, e.g., Cazals et al., 2002; Jeong et al., 2010), 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).

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 ($x_1 \dots x_p$) and q heterogeneous and non-negative outputs y ($y_1 \dots y_q$). The sample χ is then denoted by $\chi = \{(X_i, Y_i) = 1, \dots, n\}$. 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 also be possible to produce the output y with more inputs, and to produce

less outputs with a given input set x . Formally: $\forall (x, y) \in \Psi$, if $x \geq x$ and $y \leq y$ then $(x, y) \in \Psi$ [where Ψ denotes the production technology set: $\Psi = \{(x, y) \mid x \in \mathcal{R}_+^p, y \in \mathcal{R}_+^q, (x, y) \text{ is feasible}\}$]. We use the best-practice production set which is defined as a free disposable hull of undominated input-output combinations:

$$\Psi_{FDH} = \{(x, y) \in \mathcal{R}_+^{p+q} \mid x \leq X_i, y \leq Y_i \in \mathcal{X}\}. \quad (6-1)$$

In the present chapter, we evaluate efficiency from an output-biased perspective: With the given resources, what is the output shortfall for a school if it produces as efficiently as the observations on the best-practice frontier? The output-biased inefficiency estimates, $\lambda(x_o, y_o)$, measure the distance to the best-practice frontier (for further details, see Fried et al., 2008):

$$\lambda(x_o, y_o) = \sup\{\lambda \mid x_o, \lambda y_o \in \Psi_{FDH}\}. \quad (6-2)$$

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 lower than 1. The inefficiency $(1 - \lambda)$ indicates the potential percentage increase in output if the observation produces as efficiently as its reference partner.

The robust FDH model

The FDH model in equation 6-2 is deterministic, and may be problematic in the 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 argued in the introduction that the data are relatively free of measurement errors, some small and undetected errors could still be present). Cazals et al. (2002) suggested mitigating the impact of outlying observations in the FDH model, by drawing replacement subsamples of size $m < n$ from among those observations with fewer inputs than the evaluated observation (i.e. from among those Y_i such that $x_o \geq X_i$). 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 present setting, 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 2,000), the robust efficiency scores $\lambda^m(x_o, y_o)$ are obtained by taking the arithmetic average of the B inefficiencies.³⁸

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. If the evaluated observation (x_o, y_o) does not constitute its own reference set in every one of the B drawings, the efficiency score λ^m will be greater than 1. This ‘super-efficiency’ indicates that the evaluated observation is performing better than the average m observations in its reference sample (Daraio & Simar, 2007).

The robust and conditional FDH model

The robust FHD scores can be easily adapted to include heterogeneity among schools (Daraio & Simar, 2005). The exogenous variables, which can – at least in the short run – not be influenced by the school management, are denoted by z ($z_1, \dots, z_r$).³⁹ As argued in the Introduction of this chapter, the benefit of this approach is that it can account for environmental factors in efficiency estimation without assuming the separability condition.

Daraio and Simar (2005) suggested drawing the subsamples of size m with a given probability, which is determined by a Kernel function around the continuous

³⁸ 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.

³⁹ 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 the inputs, outputs, and exogenous variables (for an extensive discussion, see Fried et al., 2008).

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:

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

among those Y_i , such that $X_i \leq x_o$, where $K(\cdot)$ denotes a Kernel function and h the appropriate bandwidth (estimated by cross-validation). 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 (2008) study the possibility of also including discrete variables in the model, compared with having only continuous environmental variables. They propose a standard multivariate product kernel for continuous, ordered discrete and unordered discrete variables, in order to smooth the 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.

The conditional efficiency estimates 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 & Simar, 2005,

2007). Daraio and Simar (2005) use a smooth nonparametric kernel regression to estimate the regression model. This approach allows us 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 it is possible to 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 favourable (unfavourable) 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 the other variables are kept at a fixed value.

Based on the work of Li and Racine (2007), De Witte and Kortelainen (2008) also present a nonparametric bootstrap procedure to obtain statistical inference on the direction of the influence. They propose a local linear regression for the 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.

6.3 Institutional Setting and Data

The empirical application exploits a data set of 119 Dutch secondary education schools in 2007. These schools teach about 204,000 students, which is around 22 per cent of the total number of students in secondary education in 2007. The 119

schools comprise around 20 per cent of the Dutch secondary schools. The data are the same as the data used in Chapters 4 and 5.

Schools are currently 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 6-1, reveal that the expenses per student significantly vary among schools. Average expenses amounted to 7,150 euros per student in 2007. However, this figure ranges from about 5,250 to about 16,400 euros per student. The expenses per student depict the financial possibilities of a school. This serves in the analysis as an input (denoted before by x).

In this chapter, we use the two possible, very comparable, output variables of Dutch secondary education: the average national examination grades per school and the average student achievement per school. These variables are explained and described in both Chapters 3 and 5. Table 6-1 indicates that the average national examination grade in 2007 was 6.4, and the average student achievement was 0.86. There is significant variation among schools.

Schools also differ on various other aspects. In a sense, the innovations are in fact 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 who determine the ‘school of thought’ of its represented schools. This determination of the course can be very strict, including 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 negative effects of a centralized bureaucracy. The 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 relationship between bureaucracy and school performance depends on the way performance is measured (e.g. test scores give a negative relationship whereas attendance gives a positive relationship) (Smith & Larimer, 2004). Table 6-1 shows that in 2007 Dutch schools had an average of 2 locations and belonged to a governing body with on average 5 schools.

Second, we control for school location (Ainsworth, 2002) and the degree of urbanization (Naper, 2010; Oliveira & Santos, 2005), which have been indicated to play a role in educational performance. Population density is used as proxy for the degree of urbanization. It is measured by Statistics Netherlands as the percentage of large and very large urbanized squared kilometres within one area; it ranges from 0 to 97 per cent. Following common practice in the Netherlands (Education Inspectorate, 2011), school location is proxied by the four large Dutch regions: the North, West (“Randstad”), South, or East region of the Netherlands.

Third, the number of students per teacher 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 student per teacher per school and is on average 13.6.

Finally, we account for the observation that disadvantaged students obtain lower educational attainments (Becker & Luthar, 2002; Gaziel, 1997), and for the existence of different school types (Conceição et al., 2001; Lubienski et al., 2008). Table 6-1 shows that the share of disadvantaged students in 2007 was on average

2.7 per cent, but ranges from 0 to about 20 per cent. We observe eight different school types, which range from specialized schools for either the highest or the lowest level of education, or comprehensive schools that offer (almost) all types of education. In particular, there are four main school types (i.e. practical pre-vocational; theoretical pre-vocational; higher general; and pre-academic education) and some combinations of them (e.g. practical pre-vocational and theoretical pre-vocational education together; both pre-vocational types and general higher education together).

Table 6-1 – Descriptive statistics 2007 (n=119)

	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				
New courses/profiling innovations	5.23	2.81	0	12
Pedagogic/didactic innovations	14.60	4.94	2	26
Process innovations	12.87	4.08	3	23
Professionalization of the teacher innovations	7.20	2.87	0	14
Education chain innovations	4.39	2.15	0	10
Locations	2.06	1.55	1	8
Schools per governing body	5.63	7.90	1	31
School type (scale 1-10)	7.05	1.53	3	10
Region dummy (scale 1-4)	2.84	0.98	1	4
Percentage of students from disadvantaged neighbourhoods (in %)	2.7	3.7	0.00	19.6
Degree of urbanization (in %)	45.0	33.0	0.00	97.0
Student/teacher ratio	13.62	2.77	7.90	29.48

6.4 Results

The results are summarized in Table 6-2. The first column presents the efficiency results when we do not account for heterogeneity among schools (known as ‘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 per cent. In other words, with a given budget, Dutch secondary schools can increase their student performance by 28.6 per cent if they perform equally efficiently as the best practice. However, there is a 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 greater 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. Table 6-3 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 times bootstrap samples (De Witte & Kortelainen, 2008). Note that, due to the structure of the non-parametric bootstrap, we only present whether the exogenous variable is significantly (un)favourable correlated with efficiency. The marginal coefficient on the median is less meaningful (see De Witte & Kortelainen, 2008).

Table 6-2 – 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

Model 1 includes the five innovation clusters. The average efficiency score in Model 1 does not dramatically change compared with the unconditional model. The average efficiency amounts to 77 per cent, 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 favourable and significant. This is in line with the previous literature (see Chapter 5). The finding on new courses/profiling, pedagogic/didactic and education chain innovations is similar to the findings in Chapter 5 (Haelermans & Blank, 2012). However, this chapter is very different from Chapter 5, because in Chapter 5 we use a parametric model and panel data, whereas the present chapter we use a nonparametric model and one year of data. More discussion on the differences and similarities with Chapter 5 can be found in Section 6.5. 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 teacher professionalization innovations are also in line with the previous literature.

In Models 2 and 3, 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

Table 6-3 – Influence of environment on educational performance

	Model 1	Model 2	Model 3
New courses/Profiling innovations	favourable (0.042) **	favourable (0.106)	favourable (0.000) ***
Pedagogic/didactic innovations	favourable (0.038) **	favourable (0.000) ***	favourable (0.000) ***
Process innovations	favourable (0.032) **	favourable (0.000) ***	favourable (0.002) ***
Professionalization of the teacher innovations	favourable (0.032) **	favourable (0.000) ***	favourable (0.240)
Education chain innovations	favourable (0.048) **	favourable (0.000) ***	favourable (0.012) **
Number of locations per school			unfavourable (0.008) ***
Number of schools per governing body			unfavourable (0.000) ***
School type		(0.240)	(0.032) **
Region		(0.004) ***	(0.000) ***
Percentage of students from disadvantaged neighbourhood		unfavourable (0.432)	unfavourable (1.000)
Degree of urbanization		unfavourable (0.996)	unfavourable (1.000)
Student/teacher ratio			favourable (1.000)

Note: n = 119; Bootstrapped p-values between brackets; ***,** and * denote significance at the 1%, 5% and 10%-level, respectively.

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. We see that the school type variable is significantly related to efficiency, which is similar to what is found in the literature. Furthermore, Model 2 includes the share of disadvantaged students, the degree of urbanization and the region. The first two variables are unfavourable and insignificant, whereas the latter is just insignificant. The negative finding on the share of disadvantaged students is in line with the literature presented in Section 6-2 (e.g. Becker & Luthar, 2002). The negative sign on the degree of urbanization is similar to that found in other studies (e.g. Naper, 2010). Urbanization has an insignificant sign, indicating that, controlling 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 capture all heterogeneity which underlies 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 bureaucracy part of the literature (e.g. Marlow, 2001), but not with the literature on scale (e.g. Chakraborty et al., 2000). In our 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 number of students at the school. This contrasts with the previous literature on scale economies (Chakraborty et al., 2000; Johnes, 2006), as commonly positive scale effects are observed. However, as an additional location also leads to an additional administrative burden (due to the larger scale), the finding is in line with this stream of literature. Furthermore, an additional location may require teachers to travel from one location to the other during the course of the day, leaving less time to actually

6. The Role of Innovations in Secondary School Efficiency

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 relation 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 underlined by the literature on school types that we discussed (e.g. Lubienski et al., 2008).

The partial regression plots on the innovation clusters from Model 3 are presented in Figure 6-1. The horizontal axis presents the number of innovations within one innovation cluster. The vertical axis presents the ratio of conditional to unconditional efficiency. As described above, an upward sloping line indicates a positive relation, and a downward sloping line indicates a negative relation. As can be seen from Figure 6-1, it is not always a clear straight line. However, one can see the direction of the positive relation (as mentioned in Table 6-3) from the individual plots in Figure 6-1.

Figure 6-1 – Partial regression plots of the five innovation clusters in Model 3

6.5 Comparison of Chapter 5 and Chapter 6

Chapters 5 and 6 both analyse the relation between the five innovation clusters and productivity/efficiency. In their aim and results, these chapters are very similar, yet in their approach they are very different. Similarities between the two chapters are: that they both study the influence of innovations; that both chapters use models which allow for multiple outputs at the same time; and that both chapters use a productivity model with which efficiency scores can be calculated. However, there are more differences between the two chapters than similarities.

The first, and most important, difference is the choice of a parametric model (Chapter 5) versus a non-parametric FDH-model (Chapter 6). In Chapter 5, an assumption is made on the function form of the model (*ex post*), whereas here in Chapter 6 we let the data speak for themselves (*ex ante*), using a very robust model. Furthermore, in Chapter 5 we had to make an assumption about the separability condition, whereas in Chapter 6 no separability condition is needed. Note also that in Chapter 5 it was not possible to use the preferred estimation method (because of the complexity of the model we could not use Full Information Maximum Likelihood), which leads to possibly bias in the results.

Furthermore, another reason for differences in results might come from the fact that Chapter 5 uses 3 years of data (2005-2007) and lagged values of innovations, whereas Chapter 6 uses only one year of data (2007). Even though in Chapter 5 fixed effects are used to take into account the panel structure and incorporate the heterogeneity between schools, the results could still partly be dependent on one another, as the fixed effects only represent region and years, and not all the individual characteristics of the schools. Using the latter was impossible due to the complex structure of the model and the estimation difficulties that are related to this.

The last difference is the direct inclusion of innovations in the productivity model in Chapter 5, compared with the ‘conditional’ status of innovations in Chapter 6 (where schools are compared conditional upon their number of innovations). In

the specific indirect model in Chapter 5, it is hard to determine the influence of innovations on outcomes (instead of on productivity). However, the interaction terms with outputs and inputs solves this problem. Note that we could not use the two-stage approach to determine the influence of innovations, on account of the problems as specified by Simar and Wilson (2007). Chapter 6 does not directly include innovations in the model, but uses them as an endogenous variable based upon which schools are compared with one another. In Chapter 6 we use the difference between the efficiency scores of the unconditional and the conditional model to determine the influence of the innovations and other variables on the efficiency of schools. A non-parametric bootstrap method is then used to determine the statistical inference of these influences (a method which avoids the problems described by Simar and Wilson).

One cannot say that one method is better than the other, because they are two very different methods (as described here). However, since the results are very similar it is not necessary to choose one method over the other in this case. The results of one model makes the other model more robust, and vice versa.

Three clusters have a positive and significant influence in both chapters. Professionalization of teacher's innovations was shown to be negative and significant in Chapter 5, but is positive and not significant in Chapter 6. Considering the previous literature on the professionalization of teachers, the positive result of Chapter 6 is the more likely result. The results of the process innovations in the two chapters contradict each other, as the results are significant in both chapters but show a different sign. This difference is possibly caused by the fact that the nonparametric analysis only takes into account the year 2007, whereas the parametric analysis takes into account the years 2005-2007, and uses lagged values of the innovations. For the other types of innovations this does not change the results, but for process innovations (and professionalization of the teacher innovations) this could mean that these innovations have only recently become effective for schools. The discrepancy in the effect of process and IT innovations on educational performance is also found in the literature.

6.6 Conclusion

This chapter has developed a tailored, fully nonparametric conditional efficiency model to study secondary school efficiency in the Netherlands. The specific setting of the model allows us to study 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 (2008). 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 the 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 the school budget can be relatively freely allocated within the school. The efficiency scores are linked to a questionnaire on the innovation-mix in schools. The results indicate that innovation is positively related to efficiency. We distinguish five innovation clusters. New courses/profiling, pedagogic/didactic, 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. These findings are in line with the literature. Next, we find that school type and region are significantly related to efficiency, and that the number of locations per school and the number of schools per governing body are significantly and negatively related to school efficiency. These findings are in line with the literature.

The findings suggest that the effort of governments to stimulate innovations in education, and the effort of schools to implement these innovations, has not been in vain. Innovations seem to positively influence school performance, taking into account other exogenous and endogenous factors which influence performance. This positive relation confirms the assumptions made on innovations in education, and could be the basis for further development and stimulation of innovations in schools

6. The Role of Innovations in Secondary School Efficiency

and the research on this. Note that this chapter does not allow for a causal interpretation of the results. Causal effects still need to be studied, but this chapter provides some indications for future studies on the relationship between innovations and school performance, and the direction of this relationship.

7 Conclusions and Discussion

“Education is not the answer to the question – Education is the means to the answer to all questions”

- William Allin -

7.1 Conclusions

This thesis has explored the efficiency and productivity of (secondary) education and the role of innovations. It includes a meta-analysis on the efficiency and productivity of education, and empirical analyses of the efficiency of Dutch secondary education, the determinants of innovations in education, and the relationship between innovation and educational performance.

Different estimation techniques such as meta-regression, logistic regression, least squares, and efficiency analysis (COLS and Free Disposal Hull) are used to estimate the productivity and efficiency of education. The findings and conclusions from the various chapters are combined in this final chapter, and are discussed as concluding research findings.

1. Productivity in Dutch secondary education has decreased over the years.

Several measures of productivity have shown that productivity in Dutch secondary education has decreased between 2002 and 2007. This decrease in productivity is due to the increased investments in education relative to a decreased quality of education and decreased value of diplomas (Kuyper & van der Werf, 2007; Lange de & Dronkers, 2007). In other words, performance in education has decreased, causing productivity to decrease. We need to deal with the source of the productivity decrease first, in order to enable productivity to grow again in the future.

2. The international average efficiency scores of educational institutes is around 80 per cent, whereas the analyses in this research show that the efficiency of Dutch secondary education varies between 70-99 per cent.

In the literature, the average efficiency score of a decision-making unit usually lies between 80 and 85 per cent. The meta-analysis, on all efficiency studies published in international peer-reviewed journals in English, shows an average efficiency score of 80 per cent for educational institutes. Studies presenting efficiency scores with an average of 90 per cent and above are rare. If a sector has an average efficiency score of 90 per cent it is performing quite well and the institutes within this sector are rather homogeneous.

Table 7-1 shows that: the lowest efficiency score is presented for the full sample parametric analysis (average of 70 per cent); an efficiency score of 75 per cent is found in the subsample parametric analysis; and the highest efficiency score is found in the nonparametric analysis (99 per cent).

Compared with the average in the meta-analysis and the average finding with respect to efficiency, Dutch secondary schools are rather homogeneous, and have a high average efficiency score. The pattern that parametric studies yield lower scores than nonparametric studies is confirmed by the literature. There is a small difference between the average efficiency score of Dutch secondary education schools in the

full sample (Chapter 3) and the representative subsample of schools that provided innovation information (Chapter 5), which could be due to the sample size and the presence of innovations. However, the average efficiency score of the nonparametric analysis in Chapter 6 is considerably higher. Apart from the type of analysis (parametric vs. nonparametric) and the type of model specification with respect to inputs and outputs, the explanation for this difference is likely to be that the conditional efficiency analysis specifically compares ‘like’ with ‘like’. Another explanation for this difference is that in the bootstrapping part only subsamples of very similar schools are drawn. The efficiency score is determined by comparing an individual school with the average of the group of schools with which it is compared.

In the conditional efficiency analysis, it is even possible to obtain an efficiency score higher than 1. Chapters 3 and 5 do not make comparisons with the average school performance but with the best performing school.

Chapter 3 uses a parametric approach and compares the schools with the complete sample based on the best-performing schools (instead of comparing them with average performance of a subsample of comparable schools). Therefore, the results from the parametric analyses are the results upon which the remainder of these conclusions are drawn.

Table 7-1 – Overview of efficiency scores in the whole thesis

Chapter	Type analysis	Innovations included	Average efficiency score
2	Meta-analysis	N/a	80.4%
3	Full sample parametric	No	70.1%
5	Sub sample parametric	Yes	75.8%
6	Nonparametric (FDH)	Yes	99.7%

3. The Dutch secondary education sector is performing rather well with respect to efficiency, although there are large differences between schools.

Based on expenditures and school performance (corrected for school and environmental variables, such as region, school type, and school size), the efficiency analyses show that Dutch secondary schools have an average efficiency of between 70 and 95 per cent, which shows they are rather homogeneous when compared with each other. The average efficiency score usually is around 80 to 85 per cent and the meta-analysis on data from studies of different countries also shows an average efficiency score of 80 per cent. Although in theory these schools could improve their efficiency score by 15 to 20 per cent, in practice it is almost impossible to increase performance above an efficiency score of approximately 90 per cent.

However, 25 per cent of Dutch secondary education schools have an efficiency score of below 70 per cent, meaning that they could improve by (more than) 30 per cent.

More than a 25 per cent of the schools have between 3 and 15 per cent too few teachers, while the top 25 per cent quarter of the schools have between 3 and 12 per cent too many teachers. Half of all these schools belong to the 25 per cent that have an efficiency score below 70 per cent. In line with studies on allocative efficiency, this research shows that these schools could improve their efficiency by using a different mix of their resources, given their prices.

The conclusion is that, on average, Dutch schools are performing rather well, but that some schools could improve their efficiency. Schools in the bottom 25 per cent could shift some of the tasks from teachers to, for example, computer use, or other schools could hire more teachers instead of support or management personnel.

4. The estimates of the efficiency of education are influenced by methodological choices and data differences (meta-analysis) and differ by innovations, region, school size, and the type of education (conditional efficiency analysis).

Both the meta-analysis and the non-parametric analysis on Dutch secondary education show that the technical efficiency of education is influenced by type of education and location (region).

Innovativeness and school type also influence the efficiency, as do some methodological choices, such as type of analysis and choice of inputs and outputs. This means that the magnitude of the efficiency scores is partly determined by the researchers' methodological choices, in which the active behaviour of the sector does not play a role. However, it is also partly determined by factors that the sector *can* influence. Taking that into consideration, this research shows large similarities in the average efficiency scores of Dutch secondary schools for the different (non)parametric analyses. This conclusion indicates that schools can actively influence some of the variables that affect their efficiency.

5. Large schools in secondary education located in areas with more competition that use a progressive teaching method are more innovative than other schools.

The number of innovations in a school is mainly determined by the size of the school (i.e. number of students and number of locations), competition, school type, and whether the school uses a traditional or a progressive teaching method.

Student performance and school expenditure do not influence the innovativeness of Dutch secondary education. This means that larger schools with more competition (the number of schools within the municipality compared with the number of inhabitants) and a progressive way of teaching innovate more, and that this can be actively influenced (if desired) by either school management or the national government.

School size positively influences innovativeness, because a large scale is often related to the professionalization of an organization. This means that these organizations have more information available on new potential innovations and the

expected costs of innovations. Larger organizations also frequently have more resources, both with respect to money and with respect to personnel, which enables them to implement more innovations.

Schools that face more competition often want to become more attractive and distinguish themselves from the other schools that their potential students can choose. Schools with a more traditional teaching method mainly teach in a traditional way, in a traditional classroom setting, with one-directional communication. Traditional schools innovate less because many innovations in innovative schools are, for example, related to changing classrooms, a changing teacher role, and changing class materials.

6. Secondary schools that have implemented more innovations perform better than schools that have fewer innovations.

Both the parametric and the nonparametric analyses show that process innovations, teacher professionalization innovations, and education chain innovations are positively related to student performance, productivity and efficiency. Only the results with respect to new courses/profiling innovations and pedagogic/didactic innovations are ambiguous. The parametric analysis finds a significant negative relationship whereas the nonparametric analysis finds a significant positive relationship.

This difference is likely to be due to the fact that the nonparametric analysis only takes into account the year 2007, whereas the parametric analysis takes into account 2005 to 2007 and uses lagged values on innovations. There is not much literature on courses/profiling innovations and pedagogic/didactic innovations, but the findings are mainly positive. Student performance and school expenditures do not influence the number of innovations in a school, but most innovations do positively and significantly influence student performance.

Hence, the conclusion is that innovations are beneficial for school performance. The explanation for this is that, in order to distinguish themselves from

other schools, most schools have a focus on a few specific innovations within one innovation cluster.

However, when choosing a school, students often take the diversification between schools into account, and specifically choose a school which offers the innovative direction in which they are interested, resulting in very motivated students who will perform better. This possible selection effect could not be taken into account because no data are available on the motivation of students when they are choosing a school.

Overall, the productivity of Dutch secondary education has decreased, mainly due to decreasing student performance. This has not influenced the innovativeness and efficiency of the sector, as Dutch secondary schools are applying many different types of innovations, and have rather high average efficiency scores. Innovations are positively related to educational performance and efficiency.

7.2 Policy Implications

The findings of our research have several policy implications. Some of the conclusions can be translated into relevant advice to national government, whereas other conclusions may have implications for secondary schools. For the government, two types of policy conclusions can be derived from this thesis. These refer to: 1) possible ways to stimulate innovations in schools; and 2) possible efficiency gains for about 25 per cent of the schools in Dutch secondary education, and the role innovations play in this. For these 25 per cent of the secondary schools, policy implications also concern the possible efficiency gains that could be obtained by investing in innovations and by reallocating resources.

Policy implications at the national level

School efficiency and student performance are positively influenced by innovations. Therefore, the government should improve the efficiency and student performance of those secondary schools that have relatively low efficiency scores.

Chapters 5 and 6 showed that efficiency and student performance are significantly influenced by innovations, and by student, school, and environmental characteristics. First, with respect to innovations, the analyses showed that process innovations, teacher professionalization and education chain innovations are positively related to educational performance and efficiency. The nonparametric analysis also showed a positive relationship for new courses/profiling and pedagogic/didactic innovations.

Where policy implications are concerned, it is advisable to keep investing in education chain innovations and process innovations, and specifically to stimulate the professionalization of the teachers' innovations, as both the literature and the empirical research find positive relationships. This does not mean that the government should discourage investment in new course/profiling innovations and pedagogic/didactic types of innovations. The empirical evidence on these is mixed. It is not clear what the effects of these types of innovations are, and more research is necessary here.

Additionally, we find significant relations between the number of locations, the number of schools per governing body, school type, and region, on the one hand, and student performance and efficiency, on the other. These factors are also the determinants of innovations. The first two are measurements of scale and bureaucracy. The effects of bureaucracy are confirmed in both the literature and the analysis of our research. Policy makers and schools could introduce measures to decrease the amount of bureaucracy in schools and in the governing bodies of schools. Differences between school types and regions have to be taken into account when changing policies with respect to the performance of students and teachers.

Scale, competition, school type, and teaching methods are the main determinants of innovations. Because of the performance advantages of innovating, it is advisable to stimulate smaller schools to innovate: for example, by introducing specific innovation programmes and subsidies for small schools. Another option is to stimulate mergers, or to stimulate collaboration between schools and governing bodies. Examples of this are the community schools and network schools.

Competition stimulates innovation. It is advisable to stimulate the entry of new schools, and to stimulate schools to diversify. It is important to note that there is some tension between policies which stimulate an increase in competition and policies which stimulate an increase in scale. Larger schools mostly coincide with fewer schools per square metre and, therefore, with less competition. A policy geared towards economies of scale could undo the effect of competition, and vice versa, leaving no effect at all. One should be careful just increasing the size or competition, since this could have the undesired side effect of decreasing the innovation rate. One way of dealing with this is to demand new schools of a certain size.

Policy implications at the secondary school management level

Most secondary schools in the Netherlands are performing rather well, but about a quarter of the schools still have scope to achieve (high) efficiency gains. If a school belongs to the bottom quarter of schools with respect to efficiency, the management of these schools can actively work on improving their efficiency score and the performance of their students.

Apart from innovation-stimulating policies by the government, the school management can also decide to invest in innovations. The school management can use the technical efficiency and misallocation scores of their own school to adjust the way they are distributing the lump-sum budget over their activities. The share of schools that employ too many teachers and too little supporting personnel could

consider changing their hiring policies towards more supporting personnel, and their teaching policies towards using more innovative study materials.

7.3 Study Limitations and Further Research

There are several limitations to this research. 1) A large share of Dutch secondary schools did not fill out the questionnaire on innovations. Although the quarter of schools that did do this is a representative sample of the whole population, we do not have any knowledge on the innovativeness of the schools that did not participate; 2) There may be innovations in Dutch secondary education that are not included in the list of innovations used for the research. Schools that participated in the research had the possibility to add new innovations to the list. However, schools that did not participate may be working on different innovations compared with those in the participating schools and on the list of innovations; 3) Due to the complexity of the models in Chapters 3 and 5, it was not possible to estimate these models using maximum likelihood. Instead, least squares is used, which may lead to potential bias of the results; 4) Owing to data limitations and a lack of insight into the motivation of the management of the school in the decision to innovate, no causal relations can be analysed.

For future research, it would be advisable to: 1) set up a research design such that the causal effects of aggregated innovations on educational performance can be studied; 2) perform more research on the relationship between innovations with respect to teachers and educational performance; 3) study innovations in education in other countries as well; 4) extend the research on innovations, (allocative) efficiency, and productivity to primary education, both in the Netherlands and in other countries; and 5) relate the efficiency scores to the quality reports of the education inspectorate, and see if, and how, they match.

Bibliography

- Abbott, M., & Doucouliagos, C. (2002). A data envelopment analysis of the efficiency of Victorian TAFE institutes. *The Australian Economic Review*, 35(1), 55-69.
- Abbott, M., & Doucouliagos, C. (2009). Competition and efficiency: overseas students and technical efficiency in Australian and New Zealand Universities. *Education Economics*, 17(1), 31-57.
- Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from US universities, 1981-1999. *Research Policy*, 34, 259-285.
- Afonso, A., & St. Aubyn, M. (2006). Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. *Economic Modelling*, 23, 476-491.
- Agasisti, T., & Dal Bianco, A. (2009). Reforming the university sector: effects on teaching efficiency - Evidence from Italy. *Higher Education*, 57, 477-498.
- Agasisti, T., & Johnes, G. (2009). Beyond frontiers: comparing the efficiency of higher education decision-making units across more than one country. *Education Economics*, 17(1), 59-79.
- Agasisti, T., & Johnes, G. (2010). Heterogeneity and the evaluation of efficiency: the case of Italian universities. *Applied Economics*, 42, 1365-1375.
- Agasisti, T., & Perez-Esparrells, C. (2010). Comparing efficiency in a cross-country perspective: the case of Italian and Spanish State universities. *Higher Education*, 59, 85-103.
- Agasisti, T., & Salerno, C. (2007). Assessing the Cost Efficiency of Italian Universities. *Education Economics*, 15(4), 455-471.
- Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *The Quarterly Journal of Economics*, 120(2), 701-728.

- 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.
- Arrow, K. J. (1962). Economic welfare and the allocation of resources for invention. In Nelson (Ed.), *The Rate and Direction of Inventive Activity*. Princeton: Princeton University Press.
- Ashenfelter, O., Harmon, C., & Oosterbeek, H. (1999). A review of estimates of the schooling/earnings relationship, with test for publication bias. *Labour Economics*(6), 453-470.
- Athanassopoulos, A. D., & Shale, E. (1997). Assessing the Comparative Efficiency of Higher Education Institutions in the UK by Means of Data Envelopment Analysis. *Education Economics*, 5(2), 117-134.
- 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.
- Baker, B. D. (2001). Can flexible non-linear modeling tell us anything new about educational productivity? *Economics of Education Review*, 20, 81-92.
- Ball, C. G., & Goldman, S. (1997). Improving education's productivity: Reexamining the system to get the schools we need. *Phi Delta Kappan*, 78(3), 228-232.
- Banker, R. D., Janakiraman, S., & Natarajan, R. (2004). Analysis of trends in technical and allocative efficiency: An application to Texas public school districts. *European Journal of Operational Research*, 154, 477-491.
- Barber, M., & Mourhsed, M. (2007). *How the world's best-performing school systems come out on top*. Paris: OECD.
- Barbetta, G. P., & Turati, G. (2003). Efficiency of junior high schools and the role of proprietary structure. *Annals of Public and Cooperative Economics*, 74(4), 529-551.
- Barnett, R. R., Glass, C. J., Snowdon, R. I., & Stringer, K. S. (2002). Size, Performance and Effectiveness: Cost-Constrained Measures of Best-Practice Performance and Secondary-School Size. *Education Economics*, 10(3), 291-311.
- Barro, R. T. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2), 407-443.
- 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.
- Becker, G. (1964). *Human Capital*. New York: Columbia University Press.
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50, 1088-1101.
- Berg van der, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, 34, 145-154.

- Bettencourt, L. M. A., Lobo, J., & Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy*, 36, 107-120.
- Blank, J. L. T. (2009). Non-maximizing output behavior for firms with a cost-constrained technology. *Journal of Productivity Analysis*, 31(1), 27-32.
- Blank, J. L. T. (2010). *Principes van productiviteitsmeting. Elementaire handleiding voor kwantitatief onderzoek naar de productiviteit, doelmatigheid, effectiviteit en kwaliteit van de publieke sector*. Maastricht: Shaker Publishing B.V.
- Blank, J. L. T., Koot, P. M., & van Hulst, B. L. (2007a). *Basisonderwijs en bureaucratie - Een empirisch onderzoek naar de allocatie van middelen in basisscholen*. Delft/Rotterdam: IPSE Studies / ECORYS.
- Blank, J. L. T., Koot, P. M., & van Hulst, B. L. (2007b). *Overhead of Onderwijs - Een benchmark van de allocatie van middelen in onderwijsinstellingen voor voortgezet onderwijs*. Delft/Rotterdam: IPSE Studies / ECORYS.
- Blank, J. L. T., & van Hulst, B. L. (2009). Productive Innovations in Hospitals: An Empirical Research on the Relation between Technology and Productivity in the Dutch Hospital Industry. *Health Economics*, 18(3), 665-679.
- Blank, J. L. T., van Hulst, B. L., & Urlings, T. (2011). Does bureaucracy in secondary education increase with time and average school size? A time series and cross sectional analysis. *Bildung und Erziehung* 64(3).
- Blatchford, P., & Sumpner, C. (1998). What do we know about breaktime? Results from a national survey of breaktime and lunchtime in primary and secondary schools. *British Educational Research Journal*, 24(1), 79-94.
- Bohte, J. (2001). School Bureaucracy and Student Performance at the Local level. *Public Administration Review*, 61(1), 92-99.
- Borge, L. E., & Naper, L. R. (2006). Efficiency potential and efficiency variation in Norwegian lower secondary schools. *FinanzArchiv*, 62(2), 221-249.
- Boyne, G. A., Gould-Williams, J. S., Law, J., & Walker, R. M. (2005). Explaining the adoption of innovation: An empirical analysis of public management reform. *Environment and Planning: Government and Policy*, 23(3), 419-435.
- Bradley, S., Johnes, G., & Millington, J. (1999). School choice, competition and the efficiency of secondary schools in England. Lancaster University Management School Working Paper 003.
- 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*. Champaign, IL: University of Illinois.
- Brons, M., Nijkamp, P., Pels, E., & Rietveld, P. (2005). Efficiency of urban public transit: A meta analysis. *Transportation*, 2005(32), 1-21.
- Büchel, F. (2002). The effects of overeducation on productivity in Germany - The firms' viewpoint. *Economics of Education Review*, 21(263-275).
- Busman, L., Horsmans, E., Klein, T., & Oomen, C. (2007). *Beweging in beeld II - Feiten en verhalen over innovatie in het voortgezet onderwijs 2007*. Utrecht: Oberon & VO-raad.

- Busman, L., Klein, T., & Oomen, C. (2006). *Beweging in beeld - Feiten en cijfers over innovatie in het voortgezet onderwijs 2006*. Utrecht: Oberon & VO-Raad.
- Camisón-Zornoza, C., Lapiedra-Alcamí, R., Segarra-Ciprés, M., & Boronat-Navarro, M. (2004). A meta-analysis of innovation and organizational size. *Organization Studies*, 25(3), 331-361.
- Cantwell, J. (2006). Innovation and Competitiveness. In J. Fagerberg, D. C. Mowery & R. R. Nelson (Eds.), *The Oxford Handbook of Innovation*.
- Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics* (Vol. 3, pp. 1801-1863).
- Carlin, W., Schaffer, M., & Seabright, P. (2004). A minimum of rivalry: Evidence from transition economies on the importance of competition for innovation and growth. *Contributions to Economic Analysis & Policy*, 3(1), 1-43.
- Cazals, C., Florens, J. P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1), 1-25.
- CentER data/ECORYS/QQQ Delft. (2006). *De toekomstige arbeidsmarkt voor onderwijspersoneel tot 2015*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Centre for Educational Research and Innovation. (2008). *Innovating to learn, learn to innovate*. Paris: OECD.
- Chakraborty, K., Biswas, B., & Lewis, W. C. (2000). Economies of scale in public education: an econometric analysis. *Contemporary Economic Policy*, 18(2), 238-247.
- Chakraborty, K., Biswas, B., & Lewis, W. C. (2001). Measurement of technical efficiency in public education: A stochastic and nonstochastic production function approach. *Southern Economic Journal*, 67(4), 889-905.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444.
- Cherchye, L., & Vanden Abeele, P. (2005). On research efficiency. A micro-analysis of Dutch university research in Economics and Business Management. *Research Policy*, 34, 495-516.
- 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), May 2008.
- 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.
- 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.
- Cokgezen, M. (2009). Technical efficiencies of faculties of economics in Turkey. *Education Economics*, 17(1), 81-94.
- Conceição, M., Portela, A. S., & Thanassoulis, E. (2001). Decomposing school and school-type efficiency. *European Journal of Operational Research*, 132, 357-373.

- Conroy, S. J., & Arguea, N. M. (2008). An estimation of technical efficiency for Florida Public elementary schools. *Economics of Education Review*, 27, 655-663.
- Conti, G., Heckman, J., & Urzua, S. (2010). The education-health gradient. *The American Economic Review*, 100(2), 234-246.
- Cooley Nichols, S. M., Bicard, S. C., Bicard, D. F., & Casey, L. B. (2008). A field at risk: The teacher shortage in special education. *Phi Delta Kappan*, 89(8), 597-600.
- Cordero-Ferrera, J. M., Pedraja-Chaparro, F., & Salinas-Jiménez, J. (2008). Measuring efficiency in education: an analysis of different approaches for incorporating non-discretionary inputs. *Applied Economics*, 40, 1323-1339.
- Council of the European Union. (2007). Strategic report on the renewed Lisbon strategy for growth and jobs: launching the new cycle (2008-2010). Keeping up the pace of change. Retrieved 22 July, 2009, from http://ec.europa.eu/growthandjobs/pdf/european-dimension-200712-annual-progress-report/200712-annual-report_en.pdf
- 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.
- Damanpour, F. (1990). Organizational Size and Innovation. *Organization Studies*, 13, 375-402.
- Damanpour, F., & Schneider, M. (2006). Phases of the adoption of innovation in organizations: Effects of environment, organization and top managers. *British Journal of Management*, 17, 215-236.
- 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*. Heidelberg: Springer.
- Dawson, J., van Steen, J., & van der Meulen, B. (2009). *Science systems compared: a first description of governance innovation in six science systems*. Den Haag: Rathenau Instituut.
- De Witte, K., Groot, W., & Maassen van den Brink, H. (2010). The efficiency of education in generating literacy: a stochastic frontier approach. TIER Working Paper Series 10/17.
- De Witte, K., & Kortelainen, M. (2008). Blaming the exogenous environment? Conditional efficiency estimation with continuous and discrete environmental variables. CES Discussion Paper Series DPS 08.33; MPRA Paper 14034.
- Dee, T. S., & Keys, B. J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management*, 23(3), 471-488.
- Denaux, Z. S. (2009). Determinants of technical efficiency: Urban and rural public schools in the state of Georgia. *Southwestern Economic Review*, 26(1), 105-115.
- Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor efficiency in post offices. In M. Marchand, P. Pestieau & H. Tulkens (Eds.), *The performance of public*

- enterprises: Concept and measurements* (pp. 243-267). Amsterdam North-Holland.
- Dijkgraaf, E., van der Geest, S. A., & de Jong, J. M. (2008). *Effect van concurrentie op de kwaliteit van het HAVO en VWO*. Rotterdam: SEOR.
- Dobbins, K. (2009). Feeding innovation with learning lunches: contextualising academic innovation in higher education. *Journal of Further and Higher Education*, 33(4), 411-422.
- Dosi, G. (2000). The research on innovation diffusion: an assessment. In G. Dosi (Ed.), *Innovation, Organization and Economic Dynamics*. Cheltenham, Glos, USA: Edward Elgar Publishing Limited.
- Doucouliafos, C., & Laroche, P. (2003). What do unions do to productivity? A Meta-Analysis. *Industrial Relations*, 42(4), 650-691.
- Doucouliafos, C., & Stanley, T. D. (2009). Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of Industrial Relations*, 47(2), 406-428.
- Education Inspectorate. (2011). *De staat van het onderwijs - Onderwijsverslag 09/10*. Utrecht: Inspectie van het Onderwijs.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629-634.
- European Commission. (2006). *Progress towards the Lisbon objectives in education and training*. Brussels: European Commission.
- European Commission. (2007). *Towards more knowledge-based policy and practice in education and training*. Brussels: Commission of the European Communities. 28.08.2007.
- European Trade Union Committee for Education. (2011). *Teacher shortage in Europe*. Brussels: European Trade Union Committee for Education.
- European Union. (2009). Lisbon Strategy. Retrieved 22 July 2009, from http://europa.eu/scadplus/glossary/lisbon_strategy_en.htm
- Executive Agency of Education Audiovisual and Culture. (2009a). *National summary sheets on educational system in Europe and ongoing reforms - The Netherlands*. Brussels: European Commission.
- Executive Agency of Education Audiovisual and Culture. (2009b). *Organisation of the education system in the Netherlands*. Brussels: European Commission.
- Färe, R., & Primont, D. (1995). *Multi-Output Production and Duality: Theory and applications*. Dordrecht: Kluwer Academic Publishers.
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), 253-290.
- Ferrari, G., & Laureti, T. (2005). Evaluating technical efficiency of human capital formation in the Italian university: Evidence from Florence. *Statistical Methods & Applications*, 2005(14), 243-270.
- Financial Times. (2010). Business school rankings. from <http://rankings.ft.com/businessschoolrankings/global-mba-rankings>
- Flegg, A. T., & Allen, D. O. (2007). Does expansion cause congestion? The case of older British universities, 1994-2004. *Education Economics*, 15(1), 75-102.

- 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.
- Foreman-Peck, J., & Foreman-Peck, L. (2006). Should Schools Be Smaller? The Size-Performance Relationship for Welsh Schools. *Economics of Education Review*, 25(2), 157-171.
- Freeman, C. (1982). *The economics of industrial innovation* (2 ed.). London: Pinter Publishers.
- Fried, H. O., Lovell, C. A. K., & Schmidt, S. S. (2008). *The measurement of productive efficiency and productivity growth*. New York: Oxford University Press.
- 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.
- Giles, C., & Hargreaves, A. (2006). The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform. *Educational Administration Quarterly*, 42(1), 124-156.
- Gimenez, V. M., & Martinez, J. L. (2005). Cost efficiency in the university: A departmental evaluation model. *Economics of Education Review*, 2006(25), 543-553.
- Glass, J. C., & McKillop, D. G. (1995). The achievement of scale efficiency in UK universities: A multiple-input multiple-output analysis. *Education Economics*, 3(3), 249-264.
- Goolsbee, A., & Guryan, J. (2006). The Impact of Internet Subsidies in Public Schools. *Review of Economics and Statistics*, 88(2), 336-347.
- Greene, W. H. (2008). *Econometric Analysis*. New Jersey: Prentice-Hall International.
- Grinsven van, L., & Tillema, H. (2006). Learning opportunities to support student self-regulation: comparing different instructional formats. *Educational Research*, 48(1), 77-91.
- Groom, B., & Rose, R. (2005). Supporting the inclusion of pupils with social, emotional and behavioural difficulties in the primary school: the role of teaching assistants. *Journal of Research in Special Educational Needs*, 5(1), 20-30.
- Groot, T., & García-Valderrama, T. (2006). Research quality and efficiency. An analysis of assessments and management issues in Dutch economics and business research programs. *Research Policy*, 35, 1362-1376.
- Groot, W., & Maassen van den Brink, H. (2000). Overeducation in the labor market: a meta-analysis. *Economics of Education Review*, 19, 149-158.
- Groot, W., & Maassen van den Brink, H. (2010). The effects of education on crime. *Applied Economics*, 42(3), 279-289.

- Groot, W., & Maassen van den Brink, H. (2011). *Onderwijs en productiviteit*. Maastricht/Amsterdam: TIER.
- Grosskopf, S., Hayes, K. J., & Taylor, L. L. (2009). The relative efficiency of charter schools. *Annals of Public and Cooperative Economics*, 80(1), 67-87.
- Grosskopf, S., Hayes, K. J., Taylor, L. L., & Weber, W. L. (1997). Budget-Constrained Frontier Measures of Fiscal Equality and Efficiency in Schooling. *Review of Economics and Statistics*, 79(1), 116-124.
- Grosskopf, S., Hayes, K. J., Taylor, L. L., & Weber, W. L. (2001). On the determinants of school district efficiency: Competition and monitoring. *Journal of Urban Economics*, 49(3), 453-478.
- Grupp, H. (1998). *Foundations of the Economics of Innovation*. Cheltenham, UK: Edward Elgar Publishing Limited
- Gunday, G., Ulusov, G., Kilic, K., & Alpan, L. (May 2011). Effects of innovation types on firm performance. *International Journal of Production Economics*, doi:10.1016/j.ijpe.2011.05.014
- 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.
- Haelermans, C., & Borghans, L. (2012). Wage effects of on-the-job training: A meta-analysis. *British Journal of Industrial Relations*, doi: 10.1111/j.1467-8543.2012.00890.x.
- Haelermans, C., De Witte, K., & Blank, J. L. T. On the allocation of resources for secondary schools. *Economics of Education Review*, (2012), doi:2010.1016/j.econedurev.2012.2002.2007.
- Hanushek, E. A., & Woessmann, L. (2009). Do better schools lead to more growth? Cognitive skills, economic outcomes and causation. IZA Discussion Paper No. 4575.
- Hargreaves, L., Cunningham, M., Everton, T., Hansen, A., Hopper, B., McIntyre, D., et al. (2006). *The status of teachers and the teaching profession: Views from inside and outside the profession (Interim findings from the teacher status report)*: Department for Education and Skills.
- Hargreaves, L., Cunningham, M., Hansen, A., McIntyre, D., Oliver, C., & Pell, T. (2007). *The status of teachers and the teaching profession: Views from inside and outside the profession (Final report of the teacher status report)*: Department for Education and Skills.
- Hedges, L. V. (1992). Modeling Publication Selection Effects in Meta-Analysis. *Statistical Science*, 7(2), 246-256.
- 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.

- High Level Group chaired by Wim Kok. (2004). *Facing the challenge - The Lisbon strategy for growth and employment*. Brussels: European Council.
- Hofman, W. H. A., Hofman, R. H., Dijkstra, B. J., de Boom, J., & Meeuwisse, M. (2007). *Innovaties in het voortgezet onderwijs - Een verkenning van innovaties en effecten in het voortgezet onderwijs*. Groningen / Rotterdam: UOCG & GION / RISBO Contract Research BV.
- Houtenville, A. J., & Smith Conway, K. (2008). Parental effort, school resources and student achievement. *The Journal of Human Resources*, XLIII(2), 437-453.
- Hoxby, C. (2002). School choice and school productivity (or Could school choice be a tide that lifts all boats?). National Bureau of Economic Research, NBER Working Papers: 8873.
- Hoxby, C. (2003). School choice and school competition: evidence from the United States. *Swedish Economic Policy Review*, 10, 9-65.
- Hoxby, C. M. (2004). Productivity in education: the quintessential upstream industry. *Southern Economic Journal*, 71(2), 208-231.
- Hoyle, J. R., & Collier, V. (2006). Urban CEO Superintendents' Alternative Strategies in Reducing School Dropout. *Education and Urban Society*, 39(1), 69-90.
- Hsu, M. K., Chao, G. H., & James, M. L. (2009). An efficiency comparison of MBA programs: Top 10 vs non-top 10. *Journal of Education and Business*, May/June, 269-274.
- Hu, Y., Zhang, Z., & Liang, W. (2009). Efficiency of primary schools in Beijing, China: an evaluation by data envelopment analysis. *International Journal of Educational Management*, 23(1), 34-50.
- Huang, J., Maassen van den Brink, H., & Groot, W. (2009). A meta-analysis of the effect of education on social capital. *Economics of Education Review*, 28, 454-464.
- Ingersoll, R. M., & Perda, D. (2010). Is the supply of mathematics and science teachers sufficient? *American Educational Research Journal*, 47(3), 563-594.
- Izadi, H., Johnes, G., Oskrochi, R., & Crouchley, R. (2002). Stochastic Frontier Estimation of a CES cost function: the case of higher education in Britain. *Economics of Education Review*, 21, 63-71.
- Jacoby, D. (2006). Effects of part-time faculty employment on community college graduates rates. *The Journal of Higher Education*, 77(6), 1081-1103.
- Jarne, G., Sanchez-Choliz, J., & Fatas-Villafranca, F. (2007). S-shaped curves in economic growth. A theoretical contribution and an applications. *Evolutionary and Institutional Economics Review*, 3(2), 239-259.
- Jeong, S., Park, B., & Simar, L. (2010). Nonparametric conditional efficiency measures: asymptotic properties. *Annals of Operations Research*, 173(1), 105-122.
- Johansson, B., & Forslund, U. (2008). The analysis of location, colocation and urbanization economies. In C. Karlsson (Ed.), *Handbook of research on cluster theory*. Cheltenham, UK: Edward Elgar Publishing.
- Johnes, G. (1995). Scale and technical efficiency in the production of economic research. *Applied Economics Letters*, 2, 7-11.

- Johnes, G., & Johnes, J. (2009). Higher education institutions' costs and efficiency: taking the decomposition a further step. *Economics of Education Review*, 28, 107-113.
- Johnes, G., Johnes, J., Thanassoulis, E., Lenton, P., & Emrouznejad, A. (2005). *An Exploratory Analysis of the Cost Structure of Higher Education in England*. Lancaster: Lancaster University Management School.
- Johnes, G., & Salas-Velasco, M. (2007). The determinants of costs and efficiencies where producers are heterogeneous: the case of Spanish universities. *Economics Bulletin*, 4(15), 1-9.
- 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.
- Johnes, J. (2008). Efficiency and production change in the English higher education sector from 1996/97 to 2004/05. *The Manchester School*, 76(6), 653-674.
- Kantabutra, S., & Tang, J. C. S. (2006). Urban-rural and size effects on school efficiency: The case of northern Thailand. *Leadership and Policy in Schools*, 5, 355-377.
- Kempkes, G., & Pohl, C. (2008). Do institutions matter for university cost efficiency? Evidence from Germany. Ifo Institute for Economic Research.
- Kempkes, G., & Pohl, C. (2010). The efficiency of German universities - some evidence from nonparametric and parametric methods. *Applied Economics*, 42, 2063-2079.
- Ken, Y., Tsai, T.-Y., & Ou, Y.-K. (2008). *Study of the time lag effect of patent impact on profitability of US pharmaceutical industry from innovation to profit*. Paper presented at the Portland International Conference on Engineering and Technology, Portland.
- 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.
- Kirkpatrick, H., & Cuban, L. (1998). Computers make kids smarter -- Right? *TECHNOS Quarterly for Education and Technology*, 7(2), 1-11.
- Kirschner, P. A., Hendriks, M., Paas, F., Wopereis, I., & Cordewener, B. (2004). *Determinants for Success and Failure of Educational Innovation Projects*. Paper presented at the EDUCAUSE Annual Conferences.
- Klein, T., Oomen, C., & van der Linden, J. (2008). *Beweging in Beeld III - feiten en verhalen over innovatie in het voortgezet onderwijs 2008*. Utrecht: Oberon & VO-Raad.
- Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. G. (2007). Teachers' perspectives on innovations: Implications for educational design. *Teaching and Teacher Education*, 23, 985-997.
- Krueger, A. (2003). Economic considerations and class size. *The Economic Journal*, 113(485), F34-F63.
- Krueger, A. B., & Lindahl, M. (2001). Education for growth: Why and for whom. *Journal of Economic Literature*, 39(4), 1101-1136.

- Kumbhakar, S. C., & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. New York: Cambridge University Press.
- Kuyper, H., & van der Werf, M. P. C. (2007). *De resultaten van VOCL'89, VOCL'93 en VOCL'99 leerlingen. Vergelijkende analyse van prestaties en rendement*. Groningen: GION.
- Lachenmeier, S., & Rottmann, H. (2011). Effects of innovation on employment: a dynamic panel analysis. *International Journal of Industrial Organization*, 29, 210-220.
- Lange de, M., & Dronkers, J. (2007). *Groeide de ongelijkwaardigheid van het eindexamen tussen scholen verder in 2005? Discrepanties tussen de cijfers voor het schoolonderzoek en het centraal examen in het voortgezet onderwijs: 2005 versus 1997-2004*. Paper presented at the ORD 2007, Groningen.
- Lavy, V. (2009). Performance pay and teachers' effort, productivity and grading ethics. *American Economic Review*, 99(5), 1979-2011.
- Leigh, A., & Ryan, C. (2008). *How has school productivity changed in Australia*.
- 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*.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. USA: Sage Publications. Inc.
- Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests and self-reports. *American Economic Review*, 94(1), 155-189.
- Lovell, C. A. K. (2000). Measuring efficiency in the public sector. In J. L. T. Blank (Ed.), *Public provision and performance: Contributions from efficiency and productivity measurement*. Amsterdam: Elsevier.
- 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.
- Mangrubang, F. R. (2005). Issues and trends in science education: The shortage of qualified science teachers *American Annals of the Deaf*, 150(1), 42-46.
- Marlow, M. L. (2001). Bureaucracy and student performance in US public schools. *Applied Economics*, 33, 1341-1350.

- Mayston, D. J. (2003). Measuring and Managing Educational Performance. *The Journal of Operational Research Society*, 54(7), 679-691.
- McMillan, M. L., & Chan, W. H. (2006). University Efficiency: A Comparison and Consolidation of results from Stochastic and Non-Stochastic Methods. *Education Economics*, 14(1), 1-30.
- 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.
- Millimet, D. L., & Collier, T. (2008). Efficiency in public schools: does competition matter? *Journal of Econometrics*, 145, 134-157.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4), 281-302.
- Ministry of Education. (2009). *OnderwijsBewijs (Educational Evidence)*.
- Ministry of Education. (2010). *InnovatieImpuls Onderwijs (innovation impulse education)*.
- Ministry of Education. (2011a). Innovatie en ICT in het onderwijs. Retrieved June 21, 2011, from <http://www.rijksoverheid.nl/onderwerpen/innovatie-en-ict-in-het-onderwijs>
- Ministry of Education. (2011b). *Kerncijfers 2006-2010*. Den Haag: Ministry of Education, Culture and Science.
- Ministry of Education. (2011c). *Nota werken in het onderwijs 2011*. Den Haag: Ministry of Education, Science and Culture.
- Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *The American Economic Review*, 94(3), 656-690.
- Morone, P., & Testa, G. (2008). Firms growth, size and innovation - an investigation into the Italian manufacturing sector. *Economics of Innovation and New Technology*, 17(4), 311-329.
- Municipality of Amsterdam. (2003). *Convenant personeelstekort in het Amsterdams onderwijs*. Amsterdam: Sectorbestuur Onderwijsarbeidsmarkt.
- Naper, L. R. (2010). Teaching hiring practices and educational efficiency. *Economics of Education Review*, 29, 658-668.
- National Centre for Education Statistics. Trends in International Mathematics and Science Study (TIMSS). <http://nces.ed.gov/timss/>
- Nelis, H. (2004). Leerlingen geven oplossingen voor het lerarentekort. *VV Magazine*, 4, 10-11.
- NG, Y. C., & Li, S.-K. (2009). Efficiency and productivity growth in Chinese universities during the post-reform period. *China Economic Review*, 20, 183-192.
- Ni, Y. (2009). The impact of charter schools on the efficiency of traditional public schools: Evidence from Michigan. *Economics of Education Review*, 28, 571-584.
- 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.
- Noailly, J., & Koning, P. (2009). Schoolkeuze, concurrentie en kwaliteit in het basisonderwijs. *Economisch Statistische Berichten*, 94(554), 118-120.
- 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.
- O'Mahony, M., & Stevens, P. (2009). Output and productivity growth in the education sector: comparisons for the US and UK. *Journal of Productivity Analysis*, 31, 177-194.
- Odenthal, L., Poelmann, M., van der Ven, A., & van der Weide, J. (2007). *De kloof tussen management en docenten*. Amersfoort: CPS.
- OECD. Programme for International Student Assessment (PISA). <http://www.pisa.oecd.org/>
- OECD. (2009). *Education at a Glance*. Paris: OECD.
- OECD. (2010). *Education at a Glance*. Paris: OECD.
- Ojen van, Q. H. J. M., de Zwart, B. C. H., & Veerman, T. J. (2011). *Arbeidsproductiviteit en CAO's. een verkennend onderzoek naar de aanjagende rol van de cao-afspraken op arbeidsproductiviteit in de publieke sector*. Leiden: Astri.
- 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*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2006a). *Versteving van kennis in het onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2006b). *Waardering voor het leraarschap*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2007). *Leraarschap is eigenaarschap*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2008). *Onderwijs en open leermiddelen*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2009). *Tien jaar adviezen over leraren en lerarenopleidingen: overzicht en balans*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2011). *Excellente leraren als voorbeeld*. Den Haag: Onderwijsraad.
- Orozco-Gómez, G. (2006). Can we be more creative in thinking about how to scale up educational innovation? *Journal of Educational Change*, 7, 345-349.
- Ostrom, E. (2005). *Understandig Institutional Diversity*. New Jersey: Princeton University Press.
- Ouellette, P., & Vierstraete, V. (2005). An evaluation of the efficiency of Québec's school boards using the data envelopment analysis method. *Applied Economics*, 37, 1643-1653.
- Pereira, P. T., & Martins, P. S. (2004). Returns to education and wage equations. *Applied Economics*, 36, 525-531.
- Phillips, J. M. (1994). Farmer education and farmer efficiency: a meta-analysis. *Economic Development and Cultural Change*, 43(1), 149-165.

- Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 41(1), 116-145.
- 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*, doi:10.1016/j.econedurev.2011.07.016.
- Primont, D. F., & Domazlicky, B. (2006). Student achievement and efficiency in Missouri schools and the No Child Left Behind Act. *Economics of Education Review*, 25, 77-90.
- Punie, Y., Zinnbauer, D., & Cabrera, M. (2006). *A review of the impact of ICT on learning*. Brussels: European Commission.
- Queen, S. (2009). *Effect of cooperative learning and traditional strategies on academic performance in middle school language arts*. Walden University, Minneapolis, Minnesota.
- Rahman, S. U. (1992). Alternative representations of technological change, learning by use, and energy substitution in Canadian manufacturing industries. *Open Access Dissertations and Theses*, Paper 3843.
- Reid, K. (2003). A strategic approach to tackling school absenteeism and truancy: the PSSC scheme *Educational Studies*, 29(4), 351-371.
- Reynolds, D., Teddlie, C., Hopkins, D., & Stringfield, S. (2000). Linking school effectiveness and school improvement. In D. Reynolds & C. Teddlie (Eds.), *The international handbook of school effectiveness research*. Suffolk: RefineCatch Limited.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica*, 72(2), 417-458.
- Rodrigues, M. J. (2009). *Europe, globalization and the Lisbon agenda*. Cheltenham, UK: Edward Elgar.
- Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: Free Press.
- 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.
- Ruggiero, J. (2007). Measuring the cost of meeting minimum educational standards: an application of data envelopment analysis. *Education Economics*, 15(1), 1-13.
- Santiago, P. (2002). Teacher demand and supply: Improving teacher quality and addressing teacher shortages. OECD.
- SBO. (2011). *Arbeidsproductiviteit in het onderwijs - Notitie ten behoeve van de expertmeeting 'Arbeidsproductiviteit in het onderwijs' georganiseerd door SBO op 31 maart 2011*. Den Haag: CAOP.
- Schalock, M., Cowart, B., & Staebler, B. (1993). Teacher productivity revisited: definition, theory, measurement, and applications. *Journal of Personnel Evaluation in Education*, 7, 179-196.

- Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. New York: Harpers Collins.
- Sectorbestuur Onderwijsarbeidsmarkt. (2009). *Tekort schoolleiders in het primair onderwijs*. Den Haag: SBO.
- Shanghai Ranking. (2011). Academic ranking of world universities. *from* <http://www.shanghairanking.com>.
- Shefer, D., & Frenkel, A. (2005). R&D, firm size and innovation: an empirical analysis. *Technovation*, 25, 25-32.
- Shymansky, J. A., Kyle, W. C., & Alport, J. M. (1983). The effects of new science curricula on student performance. *Journal of Research in Science Teaching*, 20(5), 387-404.
- Silles, M. A. (2009). The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education Review*, 28, 122-128.
- Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1), 31-64.
- Simar, L., & Wilson, P. W. (2010). Inferences from cross-sectional, stochastic frontier models. *Econometric Reviews*, 29(1), 62-98.
- Simkins, T. (2000). Education reform and managerialism: Comparing the experience of schools and colleges. *Journal of Education Policy*, 15(3), 317-332.
- Simpson, H. (2009). Productivity in public services. *Journal of Economic Surveys*, 23(2), 250-276.
- Smet, M., & Nonneman, W. (1998). Economies of scale and scope in Flemish secondary schools. *Applied Economics*, 30, 1251-1258.
- Smith, A., Courvisanos, J., Tuck, J., & McEachern, S. (2011). *Building innovation capacity: the role of human capital formation in enterprises - a review of the literature*. Adelaide, Australia: National Centre for Vocational Education Research (NCVER).
- Smith, K. B., & Larimer, C. W. (2004). A mixed relationship: bureaucracy and school performance. *Public Administration Review*, 64(6), 27-36.
- Soares de Mello, J. C. C. B., Goncalves Gomes, E., Angulo Meza, L., Soares de Mello, M. H. C., & Soares de Mello, A. J. R. (2006). Engineering post-graduate programmes: a quality and productivity analysis. *Studies in Educational Evaluation*, 32, 136-152.
- Sonis, M. (2009). Innovation diffusion theory: 100 years of development. In M. Sonis & G. J. D. Hewings (Eds.), *Tool Kits in Regional Sciences*. Berlin Heidelberg: Springer-Verlag.
- 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.
- Stb. 1963; 40. *Wet op het voortgezet onderwijs*, Stb. 1963, 40. *Laatste wijziging Stb. 2010*, 80.
- Stevens, P. A. (2005). A Stochastic Frontier Analysis of English and Welsh Universities. *Education Economics*, 13(4), 355-374.

- Stijnen, P. J. J. (2003). *Leraar worden: 'under construction'?* , Open University Netherlands, Heerlen.
- Stoneman, P. (2001). *The Economics of Technological Diffusion*. Oxford.
- Story, M., Sherwood, N. E., Himes, J. H., Davis, M., Jacobs, D. R. J., Cartwright, 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 Pepe, M. (2004). *The statistical evaluation of medical tests for classification and prediction*. Oxford, UK: Oxford University Press.
- Tang, J. (2006). Competition and innovation behaviour. *Research Policy*, 35(1), 68-82.
- Tarde, G. (1903). *The laws of Imitation* (E. C. Parsons, Trans.). New York: Henry Holt and Company.
- Tauer, L. W., Fried, H. O., & Fry, W. E. (2007). Measuring efficiencies of academic departments within a college. *Education Economics*, 15(4), 473-489.
- The Economist. (2010). Full-time MBA ranking. from <http://www.economist.com/whichmba>
- The Netherlands Institute for Social Research. (2007). *Publieke prestaties in perspectief*. Den Haag: SCP.
- The Netherlands Institute for Social Research. (2010). *Publieke dienstverlening in perspectief*. Den Haag: SCP.
- The Netherlands Institute for Social Research. (2012). *Waar voor ons belastinggeld? Prijs en kwaliteit van publieke diensten*. Den Haag: SCP.
- Thiam, A., Bravo-Ureta, B. E., & Rivas, T. E. (2001). Technical efficiency in developing country agriculture: a meta-analysis. *Agricultural Economics*, 25, 235-243.
- Tyagi, P., Yadav, S. P., & Singh, S. P. (2009). Relative performance of academic departments using DEA with sensitivity analysis. *Evaluation and Program Planning*, 32, 168-177.
- US Department of Education. (2010). *Postsecondary education transition: A summary of the findings from two literature reviews*. Jessup, MD: Office of Vocational and Adult Education.
- 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.
- Van Beers, C., Kleinknecht, A., Ortt, R., & Verburg, R. (2008). *Determinants of Innovative Behaviour - A firm's internal practices and its external environment*. New York: Palgrave Macmillan.
- Van den Broek, J., Forsund, F. R., Hjalmarsson, L., & Meeusen, W. (1980). On the estimation of deterministic and stochastic frontier production functions. *Journal of econometrics*, 13, 117-138.
- Van Velden, W. (2002). Schoolleiders gezocht. *VV Magazine*, 1, 8-9.
- Vignoles, A., Levacic, R., Walker, J., Machin, S., & Reynolds, D. (2000). *The relationship between resources allocation and pupil attainment: A review*. London.
- VO-Raad. (2007). *Expeditie Durven Delen Doen (Dare, Share, Do)*.

- Waldo, S. (2007). On the use of student data in efficiency analysis - Technical efficiency in Swedish upper secondary school. *Economics of Education Review*, 26, 173-185.
- Waslander, S. (2007). *Leren over Innoveren*. Utrecht: VO-Raad.
- Waslander, S., & van der Weide, M. (2009). *Politiek, Pers en Praktijk - Over de context waarbinnen vo-scholen innoveren*. Utrecht: VO-raad.
- Wenger, J. W. (2000). What do schools produce? Implications of multiple outputs in education. *Contemporary Economic Policy*, 18(1), 27-36.
- Wong, P. K., Lee, L., & Foo, M. D. (2008). Occupational Choice: The Influence of Product vs. Process Innovation. *Small Business Economics*, 30, 267-281.
- Wooldridge, J. (2002). *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press.
- Worthington, A. C., & Lee, B. L. (2008). Efficiency, technology and productivity change in Australian universities, 1998-2003. *Economics of Education Review*, 27, 285-298.
- 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.
- Youssef, A. B., & Dahmani, M. (2008). The impact of ICT on student performance in higher education: direct effects, indirect effects and organizational change. *In: The Economics of E-learning (online monograph)*. *Revista de Universidad y Sociedad del Conocimiento (RUSC)*, 5(1), 46-56.

Appendix A

A1 – Productivity model Chapter 3

To estimate productive efficiency, we estimate an indirect-output distance function and a number of cost share equations. The model is algebraically presented as:

$$\begin{aligned}
 \ln(OD) = & a_0 + \sum_{i=1}^m b_i \ln(Y_i) + \sum_{i=1}^n c_i \ln\left(\frac{W_i}{C}\right) + \sum_{i=1}^{n'} d_i \ln(Z_i) + \\
 & \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m b_{ij} \ln(Y_i) \ln(Y_j) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} \ln\left(\frac{W_i}{C}\right) \ln\left(\frac{W_j}{C}\right) + \frac{1}{2} \sum_{i=1}^{n'} \sum_{j=1}^{n'} d_{ij} \ln(Z_i) \ln(Z_j) + \\
 & \sum_{i=1}^m \sum_{j=1}^n e_{ij} \ln(Y_i) \ln\left(\frac{W_j}{C}\right) + \sum_{i=1}^{n'} \sum_{j=1}^n f_{ij} \ln(Z_i) \ln\left(\frac{W_j}{C}\right) + \sum_{i=1}^{n'} \sum_{j=1}^m g_{ij} \ln(Z_i) \ln(Y_j) +
 \end{aligned}
 \tag{A-1}$$

$FE + \mu$
,

where

IOD = Indirect-Output Distance function;

C = total costs;
 Y_i = output/production-indicator i ($i = 1, \dots, m$);
 W_i = price of used resource i ($i = 1, \dots, n$);
 Z_i = environmental variable i ($i = 1, \dots, n'$);
 FE = fixed effects on school type, year and region; and
 μ = randomly distributed error term.
 $a_o, b_i, c_i, d_i, b_{ij}, c_{ij}, d_{ij}, e_{ij}, f_{ij}$ and g_i are the estimated parameters.

Roy's identity is used to derive the optimal cost shares (see Färe & Primont, 1995, p. 92). The optimal cost shares are the following:

$$S_i^o = \left\{ \frac{T_i}{\sum_{j=1}^n T_j} \right\}, \quad (A-2)$$

with:

$$T_i = c_i + \sum_{j=1}^n c_{ij} \ln\left(\frac{W_j}{C}\right) + \sum_{j=1}^m e_{ij} \ln(Y_j) + \sum_{j=1}^{n'} f_{ij} \ln(Z_j) + \varepsilon. \quad (A-3)$$

The output distance function is specified as a Translog function and the share equations are derived from it. Homogeneity of degree 1 in prices and symmetry is imposed by putting constraints on some of the parameters to be estimated. These constraints are the following: $\sum_{i=1}^n b_i = 1$, $\sum_{i=i}^n b_{ij} = 0, \forall j$ and $\sum_{i=i}^n e_{ij} = 0, \forall j$.

A2 – Estimation Results Final Model Chapter 3

	Parameter	Estimate	Standard error	T-statistic	P-value
A0	Constant	0.008	0.012	0.677	[.498]
B1	production indicator 1	0.497	0.114	4.373	[.000]
B2	production indicator 2	0.503	0.114	4.424	[.000]
B11	production indicator 1 * production indicator 1	8.555	1.519	5.631	[.000]
B12	production indicator 1 * production indicator 2	-8.555	1.519	-5.631	[.000]
B22	production indicator 2 * production indicator 2	8.555	1.519	5.631	[.000]
C1	price management personnel / costs	0.029	0.001	33.958	[.000]
C2	price teaching personnel / costs	0.313	0.009	36.117	[.000]
C3	price supporting personnel / costs	0.023	0.001	35.387	[.000]
C4	price material / costs	0.075	0.002	35.359	[.000]
C11	price management personnel / costs * price management personnel / costs	-0.004	0.004	-1.047	[.295]
C12	price management personnel / costs * price teaching personnel / costs	-0.016	0.004	-3.662	[.000]
C13	price management personnel / costs * price supporting personnel / costs	0.030	0.002	14.801	[.000]
C14	price management personnel / costs * price material / costs	-0.003	0.002	-1.696	[.090]
C22	price teaching personnel / costs * price teaching personnel / costs	0.081	0.011	7.296	[.000]
C23	price teaching personnel / costs * price supporting personnel / costs	-0.028	0.005	-6.058	[.000]
C24	price teaching personnel / costs * price material / costs	0.026	0.004	6.554	[.000]
C33	price supporting personnel / costs * price supporting personnel / costs	-0.004	0.004	-0.934	[.350]
C34	price supporting personnel / costs * price material / costs	0.006	0.001	5.898	[.000]
C44	price material / costs * price material / costs	-0.012	0.003	-3.625	[.000]
E11	production indicator 1 * price management personnel / costs	-0.035	0.010	-3.412	[.001]
E12	production indicator 1 * price teaching personnel / costs	-0.364	0.099	-3.669	[.000]

E13	production indicator 1 * price supporting personnel / costs	-0.029	0.007	-3.884	[.000]
E14	production indicator 1 * price material / costs	-0.061	0.026	-2.391	[.017]
E21	production indicator 2 * price management personnel / costs	0.035	0.010	3.412	[.001]
E22	production indicator 2 * price teaching personnel / costs	0.364	0.099	3.669	[.000]
E23	production indicator 2 * price supporting personnel / costs	0.029	0.007	3.884	[.000]
E24	production indicator 2 * price material / costs	0.061	0.026	2.391	[.017]
D1	competition	-0.038	0.008	-4.748	[.000]
D2	number of schools per governing body	0.012	0.004	2.835	[.005]
D3	percentage disadvantaged students	0.032	0.003	9.704	[.000]
D4	percentage urbanization	0.008	0.009	0.927	[.354]
D5	number of locations per school	-0.077	0.011	-7.111	[.000]
D6	average experience teachers	0.078	0.021	3.700	[.000]
D7	share of part-time workers	0.068	0.016	4.218	[.000]
D11	competition*competition	-0.006	0.009	-0.670	[.503]
D12	competition*number of schools per governing body	-0.010	0.004	-2.517	[.012]
D13	competition*percentage disadvantaged students	-0.006	0.002	-2.850	[.004]
D14	competition*percentage urbanization	-0.005	0.005	-0.901	[.368]
D15	competition*number of locations per school	0.036	0.009	3.830	[.000]
D16	competition*average experience teachers	-0.016	0.017	-0.933	[.351]
D17	competition*share of part-time workers	-0.004	0.014	-0.313	[.754]
D22	number of schools per governing body*number of schools per governing body	-0.022	0.005	-4.168	[.000]
D23	number of schools per governing body*percentage disadvantaged students	0.002	0.001	1.585	[.113]
D24	number of schools per governing body*percentage urbanization	0.004	0.002	1.526	[.127]
D25	number of schools per governing body*number of locations per school	0.001	0.005	0.188	[.851]
D26	number of schools per governing body*average experience teachers	0.015	0.009	1.672	[.095]
D27	number of schools per governing body*share of part-time workers	0.019	0.009	2.144	[.032]

D33	percentage disadvantaged students*percentage disadvantaged students	0.006	0.001	7.247	[.000]
D34	percentage disadvantaged students*percentage urbanization	0.000	0.001	-0.112	[.911]
D37	percentage disadvantaged students*number of locations per school	-0.010	0.003	-3.691	[.000]
D38	percentage disadvantaged students*average experience teachers	-0.004	0.006	-0.620	[.535]
D39	percentage disadvantaged students*share of part-time workers	0.000	0.004	0.074	[.941]
D44	percentage urbanization	0.003	0.006	0.489	[.625]
D47	percentage urbanization*number of locations per school	0.006	0.007	0.852	[.394]
D48	percentage urbanization*average experience teachers	-0.004	0.015	-0.294	[.769]
D49	percentage urbanization*share of part-time workers	-0.004	0.008	-0.503	[.615]
D77	number of locations per school	-0.098	0.027	-3.609	[.000]
D78	number of locations per school*average experience teachers	-0.104	0.028	-3.748	[.000]
D79	number of locations per school*share of part-time workers	-0.052	0.022	-2.341	[.019]
D88	average experience teachers	0.020	0.022	0.919	[.358]
D89	average experience teachers*share of part-time workers	0.000	0.028	-0.001	[.999]
D99	share of part-time workers*share of part-time workers	0.092	0.038	2.429	[.015]
G11	production indicator 1 *competition	-0.164	0.097	-1.690	[.091]
G12	production indicator 2 *competition	0.164	0.097	1.690	[.091]
G21	production indicator 1 *number of schools per governing body	0.109	0.055	1.974	[.048]
G22	production indicator 2 *number of schools per governing body	-0.109	0.055	-1.974	[.048]
G31	production indicator 1 *percentage disadvantaged students	0.004	0.031	0.139	[.889]
G32	production indicator 2 *percentage disadvantaged students	0.039	0.031	1.257	[.209]
G41	production indicator 1 *percentage urbanization	0.024	0.069	0.354	[.723]
G42	production indicator 2 *percentage urbanization	-0.065	0.068	-0.955	[.340]
G51	production indicator 1 *number of locations per school	-0.407	0.151	-2.693	[.007]

G52	production indicator 2 *number of locations per school	0.585	0.150	3.907	[.000]
G61	production indicator 1 *average experience teachers	0.423	0.240	1.763	[.078]
G62	production indicator 2 *average experience teachers	-0.214	0.238	-0.901	[.367]
G71	production indicator 1 *share of part-time workers	0.548	0.208	2.639	[.008]
G72	production indicator 2 *share of part-time workers	-0.405	0.204	-1.987	[.047]
F11	price management personnel / costs*competition	0.003	0.001	5.475	[.000]
F12	price teaching personnel / costs*competition	0.041	0.006	7.087	[.000]
F13	price supporting personnel / costs*competition	0.003	0.000	6.182	[.000]
F14	price material / costs* competition	0.008	0.001	5.330	[.000]
F21	price management personnel / costs*number of schools per governing body	0.000	0.000	-1.099	[.272]
F22	price teaching personnel / costs*number of schools per governing body	-0.005	0.003	-1.306	[.192]
F23	price supporting personnel / costs*number of schools per governing body	0.000	0.000	-1.257	[.209]
F24	price material / costs*number of schools per governing body	-0.002	0.001	-2.460	[.014]
F31	price management personnel / costs *percentage disadvantaged students	0.001	0.000	5.653	[.000]
F32	price teaching personnel / costs *percentage disadvantaged students	0.014	0.002	6.433	[.000]
F33	price supporting personnel / costs *percentage disadvantaged students	0.001	0.000	6.891	[.000]
F34	price material / costs *percentage disadvantaged students	0.003	0.001	6.169	[.000]
F41	price management personnel / costs*percentage urbanization	-0.001	0.001	-2.027	[.043]
F42	price teaching personnel / costs*percentage urbanization	-0.016	0.007	-2.280	[.023]
F43	price supporting personnel / costs*percentage urbanization	-0.001	0.000	-1.432	[.152]
F44	price material / costs*percentage urbanization	-0.003	0.002	-1.501	[.133]
F51	price management personnel / costs*number of locations per school	0.012	0.001	9.549	[.000]

F52	price teaching personnel / costs*number of locations per school	0.113	0.013	8.625	[.000]
F53	price supporting personnel / costs*number of locations per school	0.009	0.001	9.086	[.000]
F54	price material / costs*number of locations per school	0.029	0.003	8.686	[.000]
F61	price management personnel / costs*average experience teachers	0.007	0.001	9.744	[.000]
F62	price teaching personnel / costs*average experience teachers	0.068	0.008	8.552	[.000]
F63	price supporting personnel / costs*average experience teachers	0.004	0.001	5.181	[.000]
F64	price material / costs*average experience teachers	0.016	0.002	6.950	[.000]
F71	price management personnel / costs *share of part-time workers	0.003	0.001	2.181	[.029]
F72	price teaching personnel / costs *share of part-time workers	0.029	0.011	2.551	[.011]
F73	price supporting personnel / costs *share of part-time workers	0.003	0.001	2.760	[.006]
F74	price material / costs *share of part-time workers	0.009	0.003	3.048	[.002]
R1	school type 1	0.176	0.035	5.045	[.000]
R2	school type 2	-0.053	0.025	-2.146	[.032]
R3	school type 3	0.076	0.016	4.710	[.000]
R4	school type 4	0.071	0.029	2.443	[.015]
R5	school type 5	0.033	0.009	3.651	[.000]
R7	school type 7	-0.132	0.011	-12.326	[.000]
R8	school type 8	-0.119	0.014	-8.428	[.000]
R9	school type 0	-0.088	0.017	-5.286	[.000]
Q2	year=2003	0.007	0.008	0.799	[.424]
Q3	year=2004	0.008	0.008	0.914	[.361]
Q4	year=2005	0.033	0.009	3.884	[.000]
Q5	year=2006	0.055	0.009	6.291	[.000]
Q6	year=2007	0.090	0.009	10.153	[.000]
S1	region=1 (North)	0.043	0.011	3.900	[.000]
S2	region=2 (East)	-0.010	0.009	-1.148	[.251]
S3	region=3 (West)	0.031	0.008	3.700	[.000]

A3 – Results second-order conditions

Table A3-1 – Eigenvalues of the matrix of global concavity

Lambda	Eigenvalues
λ_1	0.172
λ_2	-4.830
λ_3	-8.167
λ_4	-30.572

Table A3-2 – Eigenvalue of the Hessian matrix

	Mean	Minimum	Maximum
Minimum eigenvalues	-3.665+12	-4.398+12	-27.484
Maximum eigenvalues	0.029	0.000033	0.19703
Maximum eigenvalues negative	0.833	0	1

A4 – Rank-order correlations

Table A4-1 – Rank-order correlations for separate years

	2007	2006	2005	2004	2003	2002
2007	1					
2006	.840	1				
2005	.724	.848	1			
2004	.482	.517	.721	1		
2003	.444	.493	.648	.855	1	
2002	.339	.401	.538	.701	.873	1

Table A4-2 – Rank order correlations efficiency scores averaged over 3 years

	2002-2004	2003-2005	2004-2006	2005-2007
2002-2004	1			
2003-2005	.905	1		
2004-2006	.726	.920	1	
2005-2007	.559	.770	.932	1

Appendix B

B1 – Productivity model Chapter 5

The budget constrained indirect-output distance function is defined as follows:

$$\begin{aligned}
 \ln(IOD) = & a_0 + \sum_{m=1}^M b_m \ln(Y_m) + \sum_{n=1}^N c_n \ln\left(\frac{W_n}{C}\right) + \frac{1}{2} \sum_{m=1}^M \sum_{m'=1}^{M'} b_{mm'} \ln(Y_m) \ln(Y_{m'}) + \\
 & \frac{1}{2} \sum_{n=1}^N \sum_{n'=1}^{N'} c_{nn'} \ln\left(\frac{W_n}{C}\right) \ln\left(\frac{W_{n'}}{C}\right) + \sum_{m=1}^M \sum_{n=1}^N e_{mn} \ln(Y_m) \ln\left(\frac{W_n}{C}\right) + \\
 & h_0 * A(\ln(TC)) + \sum_{m=1}^M j_m * \ln(Y_m) * A(\ln(TC)) + \\
 & \sum_{n=1}^N j_n * \ln\left(\frac{W_n}{C}\right) * A(\ln(TC)) + FE + v
 \end{aligned} \tag{B-1}$$

$$\text{with: } Y_m = achievement_m, \tag{B-2}$$

where $achievement_m$ is the achievement for the students of education type m .

IOD = Indirect-Output Distance Function;

C = Total costs;

- Y_m = Output/production-indicator m ($m = 1, \dots, M$);
 W_n = Price used resource n ($n = 1, \dots, N$);
 TC = Technical change;
 FE = School type and year dummy fixed effects;
 v = Randomly distributed error term.
 $\alpha_o, b_m, c_n, b_{mm'}, c_{nn'}, e_{mn}, h_0, i_m$ and j_n are the estimated parameters.

Based on Roy's identity (see Färe & Primont, 1995, p. 92) we derive the following cost share equations:

$$S_i^o = \frac{T_i}{\sum_{j=1}^n T_j} \quad , \quad (B-3)$$

With:

$$T_i = c_i + \sum_{n=1}^N c_{nn'} \ln\left(\frac{W_j}{C}\right) + \sum_{m=1}^M e_{mn} \ln(Y_j) + \sum_{n=1}^N j_n A(\ln(TC)) \quad . \quad (B-4)$$

The output distance function is specified as a translog function and the share equations are derived from it. Homogeneity of degree 1 in prices and symmetry is imposed by putting constraints on some of the parameters to be estimated. These

constraints are as follows: $\sum_{i=m}^M b_m = 1$, $\sum_{i=m}^M b_{mm'} = 0, \forall j$ and $\sum_{i=m}^M e_{mn} = 0, \forall j$.

B2 – Full Estimation Results Final Model Chapter 5

	Parameter	Estimate	Standard error	T-statistic	P-value
A0	Constant	0.376	0.038	9.960	[.000]
A1	new courses/profiling innovations	0.473	0.067	7.054	[.000]
A2	pedagogical/didactical innovations	0.263	0.106	2.472	[.013]
A3	process innovations	-0.092	0.152	-0.608	[.543]
A4	professionalization of the teachers innovations	-0.942	0.177	-5.314	[.000]
A5	education chain innovations	-0.118	0.062	-1.893	[.058]
B1	production indicator 1	0.537	0.035	15.371	[.000]
B2	production indicator 2	0.241	0.020	12.140	[.000]
B3	production indicator 3	0.062	0.021	2.978	[.003]
B4	production indicator 4	0.056	0.036	1.556	[.120]
B5	production indicator 5	0.047	0.049	0.959	[.337]
B6	production indicator 6	0.057	0.023	2.473	[.013]
B11	production indicator 1 *	-0.176	0.058	-3.037	[.002]
B12	production indicator 1 *	0.198	0.027	7.442	[.000]
B13	production indicator 2 *	0.012	0.019	0.614	[.539]
B14	production indicator 3 *	-0.065	0.031	-2.084	[.037]
B15	production indicator 4 *	0.040	0.025	1.615	[.106]
B16	production indicator 5 *	-0.008	0.016	-0.513	[.608]
B22	production indicator 6 *	-0.190	0.012	-16.383	[.000]
B23	production indicator 2 *	-0.020	0.013	-1.538	[.124]
B24	production indicator 3 *	0.004	0.021	0.206	[.836]
B25	production indicator 4 *	0.026	0.021	1.237	[.216]
B26	production indicator 5 *	-0.019	0.011	-1.639	[.101]
B33	production indicator 6 *	0.042	0.014	2.940	[.003]
B34	production indicator 3 *	-0.008	0.013	-0.620	[.535]
	production indicator 4				

B35	production indicator 3 *	-0.028	0.015	-1.922	[.055]
	production indicator 5				
B36	production indicator 3 *	0.002	0.009	0.220	[.826]
	production indicator 6				
B44	production indicator 4 *	0.081	0.032	2.513	[.012]
	production indicator 4				
B45	production indicator 4 *	-0.018	0.032	-0.566	[.571]
	production indicator 5				
B46	production indicator 4 *	0.006	0.018	0.326	[.745]
	production indicator 6				
B55	production indicator 5 *	-0.036	0.045	-0.803	[.422]
	production indicator 5				
B56	production indicator 5 *	0.017	0.018	0.901	[.367]
	production indicator 6				
B66	production indicator 6 *	0.002	0.010	0.242	[.809]
	production indicator 6				
C1	price management personnel / costs	0.015	0.001	10.106	[.000]
C2	price teaching personnel / costs	0.186	0.018	10.635	[.000]
C3	price supporting personnel / costs	0.012	0.001	10.466	[.000]
C4	price material / costs	0.047	0.004	10.500	[.000]
C11	price management personnel / costs * price management personnel / costs	-0.001	0.001	-1.035	[.301]
C12	price management personnel / costs * price teaching personnel / costs	-0.004	0.002	-2.463	[.014]
C13	price management personnel / costs * price supporting personnel / costs	0.003	0.001	3.995	[.000]
C14	price management personnel / costs * price material / costs	-0.003	0.001	-3.119	[.002]
C22	price teaching personnel / costs * price teaching personnel / costs	-0.044	0.006	-7.434	[.000]
C23	price teaching personnel / costs * price supporting personnel / costs	-0.011	0.003	-3.902	[.000]
C24	price teaching personnel / costs * price material / costs	-0.015	0.002	-6.406	[.000]
C33	price supporting personnel / costs * price supporting	0.005	0.002	2.267	[.023]

	personnel / costs				
C34	price supporting personnel / costs* price material / costs	-0.002	0.001	-3.492	[.000]
C44	price material / costs* price material / costs	0.002	0.002	1.183	[.237]
E11	production indicator 1 * price management personnel / costs	-0.009	0.001	-8.698	[.000]
E12	production indicator 1 * price teaching personnel / costs	-0.112	0.012	-9.522	[.000]
E13	production indicator 1 * price supporting personnel / costs	-0.007	0.001	-9.132	[.000]
E14	production indicator 1 * price material / costs	-0.028	0.003	-9.321	[.000]
E21	production indicator 2 * price management personnel / costs	0.009	0.001	9.818	[.000]
E22	production indicator 2 * price teaching personnel / costs	0.111	0.010	10.722	[.000]
E23	production indicator 2 * price supporting personnel / costs	0.007	0.001	10.442	[.000]
E24	production indicator 2 * price material / costs	0.028	0.003	10.539	[.000]
E31	production indicator 3 * price management personnel / costs	0.000	0.000	0.999	[.318]
E32	production indicator 3 * price teaching personnel / costs	0.001	0.003	0.462	[.644]
E33	production indicator 3 * price supporting personnel / costs	0.000	0.000	0.350	[.726]
E34	production indicator 3 * price material / costs	0.001	0.001	0.974	[.330]
E41	production indicator 4 * price management personnel / costs	0.000	0.000	-0.966	[.334]
E42	production indicator 4 * price teaching personnel / costs	-0.010	0.005	-2.146	[.032]
E43	production indicator 4 * price supporting personnel / costs	-0.001	0.000	-2.294	[.022]
E44	production indicator 4 * price material / costs	-0.003	0.001	-2.129	[.033]
E51	production indicator 5 * price management personnel / costs	0.001	0.000	2.155	[.031]

E52	production indicator 5 * price teaching personnel / costs	0.014	0.005	2.955	[.003]
E53	production indicator 5 * price supporting personnel / costs	0.001	0.000	2.614	[.009]
E54	production indicator 5 * price material / costs	0.003	0.001	2.638	[.008]
E61	production indicator 6 * price management personnel / costs	0.000	0.000	-0.283	[.777]
E62	production indicator 6 * price teaching personnel / costs	-0.004	0.003	-1.400	[.162]
E63	production indicator 6 * price supporting personnel / costs	0.000	0.000	-1.402	[.161]
E64	production indicator 6 * price material / costs	-0.001	0.001	-1.133	[.257]
i1	innovation * production indicator 1	-0.603	0.107	-5.655	[.000]
i2	innovation * production indicator 2	0.530	0.055	9.586	[.000]
i3	innovation * production indicator 3	0.034	0.037	0.932	[.351]
i4	innovation * production indicator 4	-0.061	0.076	-0.805	[.421]
i5	innovation * production indicator 5	0.137	0.088	1.549	[.121]
i6	innovation * production indicator 6	-0.036	0.052	-0.696	[.486]
r1	school type 1	0.287	0.182	1.573	[.116]
r2	school type 2	0.371	0.168	2.203	[.028]
r3	school type 3	0.363	0.146	2.478	[.013]
r4	school type 4	-0.201	0.083	-2.424	[.015]
r5	school type 5	-0.437	0.032	-13.543	[.000]
r7	school type 7	-0.145	0.058	-2.520	[.012]
r8	school type 8	-0.245	0.145	-1.690	[.091]
r9	school type 9	0.846	0.320	2.646	[.008]
Q4	year=2004	-0.074	0.012	-6.092	[.000]
Q5	year=2005	-0.043	0.012	-3.570	[.000]
S1	region=1 (North)	0.050	0.020	2.505	[.012]
S2	region=2 (East)	0.038	0.016	2.358	[.018]
S3	region=3 (West)	-0.028	0.017	-1.671	[.095]

Appendix C

C1 – Innovation clusters

New courses / profiling	Professionalization of teachers
New courses Course on Nature, Life and Technology Mathematics D Mathematics C Course on leisure activities Course on social skills Course on media (e.g. journalism) Course on gathering, processing and using information Other innovation mentioned by school Profiling Highly gifted pupils profile school Sports school School with science classes Technical PA profiling school Art/culture school Network school Social school PA-plus school Preparatory training conservatoire Preparatory training dance Theatre school Specific PV profile school Other innovation mentioned by school Language Bilingual education Grammar for all languages in one class Specific classes for pupils that lag behind in languages Attending project 'Certi Lingua' Attending project 'DELF Scolaire' Attending Anglia language exams Project for other modern languages Literature course for all languages at once Project on youth literature for lower school Project on classical languages Other innovation mentioned by school	Teaching staff Electronic competency file for teachers Personal development plan for teachers Coaches for (teams of) teachers Flat organizational structure Peercoaching for teachers All teachers have their own laptop Compulsory additional education for teachers School is instruction school for starting teachers Teachers are paid by performance New description of the functions of teachers Teacher develop their own teaching material Classes are given by more than one teacher at the same time Teachers are more coaches than real teachers Teachers work in teams, not in subject groups Teams of teachers can decide the content of the courses they teach all by themselves One plan for a group of teachers that all teach the same subject Other innovation mentioned by school Education chain Connection to practice Extra attention for sector orientation PV pupils Learning and working projects for PV pupils Specialization project of last year carried out at company outside the school Social internship Internationalization projects PV 3rd and 4th year have internships more than 73% of the time Pupils start their own company and run it all by themselves Other innovation mentioned by school Connection to primary education and higher education Last year pupils follow lectures at university already Scientific orientation project Preparation project for pupils for higher education guest classes/project for primary education pupils Offer hard learning primary education pupils opportunity to pre apply for secondary education Other innovation mentioned by school

Process	Pedagogical / Didactical (continuation)
Administratively / organizational Questionnaires among pupils Presence registration system Systematical questionnaires for alumni Same exam for all pupils of same year at same time for same subject Teacher takes care of classes of colleagues that are sick Some tasks are contracted out to employment agencies, who hire pupils Using parent or pupils council groups Adapting schedule to working in modules Working in periods, where certain classes are only taught in certain periods Innovative hours scheduled for managers Emphasis on continuous education line Sponsoring of school activities by third parties In lower school pupils are not divided by level yet Other innovation mentioned by school IT Digital chalkboard School uses digital teaching methods more than 30% of the time Teachers develop digital teaching methods Using pupils as IT coaches for teachers (Digital) educational planner for all pupils Examining on the computer Offering central exam on computer Electronically learning Environment All pupils (of at least one year) have own laptop ICT-rich environment at school Electronic grading system Using educational TV programs Digital system to follow the pupil from year to year Giving parents online access to grades and absence information of child Other innovation mentioned by school Infrastructure Financial cooperation with businesses New school building without classrooms but with learning communities One building together with vocational education school First year pupils completely separate from other pupils Pupils in learning communities, no classrooms anymore One big technical area for PT/PV education and vocational education Separating pupils of different years from one another Using sound chairs and tables Other innovation mentioned by school	Services (school uses services of) Education-in-own-language teachers Psychologist Remedial teacher Dyslexia specialist Physiotherapist Occupational therapist Speech therapist Social expert Didactical expert Medical specialist Creative therapist Other innovation mentioned by school Pupils Pupils are mentor for other pupils Peer mediation between pupils Pupils have in say in whether they continue to the next year Use of self reflection instrument for pupils Future development plan for pupils Question driven education Applied education plan (pupils decides own education process) Other innovation mentioned by school Classes and methods Offering more courses for broadening Debating projects Topicality and news project Projects going beyond course subjects Teaching large groups of pupils at the same time Extra education in courses for weak pupils Independent working moment in every class Homework free school Learning process accompaniment in lower schools Longer or shorter classes School uses week tasks instead of day tasks Offering central examination possibilities several times a year Pedagogical/didactical teaching methods for PV Learning by doing research (learning by trial and error) Educational projects organized and guided by parents Other innovation mentioned by school Beta Beta projects Emphasis on improving technical education Using edulabs (technical courses in modules at computer) Other innovation mentioned by school
Pedagogical / Didactical Health, safety and health care Particular care structure in the school Project on safety in the school Collective participation in sport events Extra care for healthy food in the canteen Other innovation mentioned by school	

Summary

Productivity of education is a topic of intense debate among politicians, teachers, and trade unions and employers in education. Productivity drives our economic development. Therefore, it is important not only to focus on improving productivity but also to measure and monitor the development of productivity. As responsibility, accountability, and transparency are increasingly becoming the norm in education, it is important to gain insight into educational productivity.

Productivity is defined as the ratio of *outputs* (e.g. student performance) to *inputs* (e.g. costs per student). The different *inputs* and *outputs* that are used define the focus and, in many cases, also the quality of the productivity research. Productivity in education is often measured over time, to gain insights into productivity development. Productivity relative to others in the same year is referred to as efficiency. Efficiency of education is productivity compared with the best-performing school. In order to improve productivity and efficiency of education, many schools are implementing innovations. The concept of innovation is defined as: *'an idea, process or product that is new for a particular organization when it is introduced by that organization'*.

There are many studies on productivity and efficiency in education, but these differ widely, both with respect to content and methodology. Despite the

importance of knowledge about allocative efficiency, there are only few studies which focus on allocative efficiency. Research on allocative efficiency gives insight into the way schools spend their money, and how they can improve the allocation of resources.

Furthermore, only few studies have been done on the relationship between (aggregated) innovations and school performance and efficiency. Most studies on innovations in education do not search for quantitative evidence on the importance or effects of innovations at the aggregate (school) level, nor do they provide an economic analysis of the relationship of innovations to school productivity.

The added value of this thesis is the study of the (allocative) efficiency of Dutch secondary schools and the influence of innovations on efficiency. In order to do so, we use innovative empirical methods and a unique data set on innovations in education, aggregated at the school level.

The main findings of the research are that Dutch secondary schools are quite efficient. This efficiency is positively influenced by innovations. Apart from innovations, efficiency is also influenced by allocation of resources and size. Furthermore, efficiency scores differ by region and school type.

Chapters 2 and 3 of this thesis study the efficiency of education. The meta-analysis in Chapter 2 provides an overview of all types of efficiency studies in education. This chapter includes primary, secondary, and higher education. This chapter shows that the efficiency of education is approximately 0.804, so, on average, educational performance is 20 per cent less than is potentially possible, considering the resources available. We test and correct for missing studies (publication bias) in the meta-analysis data set of educational institutes. The corrected analysis confirms that efficiency scores are different for several study and population characteristics, which implies that there are several possible reasons for differences in efficiency scores. For example, the average efficiency score turns out to be lower for primary and secondary education than for higher education. This can be explained by the fact

that higher education institutes may be more similar to each other compared with primary and secondary education schools. The similarity might lie in the student population (e.g. with respect to brightness and social status), but also in the organizational structure. Furthermore, the average efficiency score is lower for earlier studies compared with more recent studies. This might be because earlier studies very often only use one *input* and/or *output* and increasing the number of *inputs* and/or *outputs* leads to higher efficiency scores. Furthermore, the average efficiency scores are lower for studies from Europe compared with North American studies. One reason for this difference might be that studies using data from North America often use school districts as the unit of observation, whereas studies using European data use individual school data in most cases. Finally, the analysis shows that the unit of observation (districts vs. schools vs. departments) also influences the efficiency scores, which can be explained by the fact that in higher education departments are often used as the unit of analysis, whereas in primary and secondary education mainly schools or districts are used. Higher education institutes in most cases have two outputs, research and teaching, and it is often argued that lower average efficiency scores are due to the trade-off that has to be made between these two outputs.

The case study used for Chapters 3 to 6 of this thesis is secondary education in the Netherlands. In Chapter 3, the optimal allocation of resources – in terms of school management, teachers, supporting employees, and materials – in secondary schools is studied. A parametric production function is used to study the efficiency for 448 Dutch secondary schools between 2002 and 2007. In the production model it is assumed that schools aim to maximize the educational attainments of students under a given budget, which can be freely allocated. In this chapter, we study whether this budget is optimally allocated over managers, teachers, supporting personnel, and materials. We control for differences between schools by correcting for school type, year, and region. The results indicate that the average efficiency score amounts to about 70 per cent, implying that, with the given resources, schools could improve students' educational performance by about 30 percentage point.

Furthermore, we find that on average, schools had about 0.21 percent too many teachers in 2007, while there were too few managers and supporting personnel. In practice this means that the average school could do with one teacher less. This indicates that the average school is close to the optimal amount of teachers. An explanation for this is that schools might have anticipated the teacher shortage, and are doing rather well on the allocation of resources. Finally, Chapter 3 shows that productivity decreased for these schools between 2002 and 2007, when we compare student achievement and national examination grades with expenditures of schools. This productivity decrease can be explained by the decrease of student performance and educational quality, which was recently shown in studies on Dutch secondary schools.

In Chapters 4 to 6, we study innovations in secondary education and the relation between these innovations and school performance and efficiency. In these chapters, we distinguish between five innovation clusters: new courses/profiling innovations, pedagogic/didactic innovations, process innovations, teacher professionalization and education chain innovations. Innovations in the new courses/profiling cluster are related to the introduction of new courses, such as new mathematics courses; language innovations, such as bilingual education; and 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.

In Chapter 4, we analyse the diffusion and determinants of innovations in secondary education. First, the diffusion path of five clusters (new courses/profiling, pedagogic/didactic, process, teacher professionalization and education chain) of innovations in secondary schools is described. In 2002 there were only few schools with innovations, whereas in 2007 most innovations were spread over most schools. Next, the results show that size is one of the main determinants of the share of innovations present at a school. Other determinants of innovativeness are competition, school type, and the teaching method. This means that larger schools with more competition and a progressive way of teaching (no longer the traditional class room teaching) innovate more. School size positively influences innovativeness, because a large scale is often related to the professionalization of an organization, which often means having more information available on innovations and the expected costs of innovations. Furthermore, schools that face more competition often want to become more attractive and distinguish themselves from the other schools, in order to attract more or better students.

In Chapter 5 we examine the relation between innovations and productivity, using 126 Dutch secondary schools. Comparable with Chapter 3, we use a parametric production function to study productivity. The innovation clusters are included in the production model to study the relationship of innovations to school productivity. We correct the innovations for the determinants from Chapter 4 that influence the innovativeness of a school (e.g. scale and competition). Comparable to the model in Chapter 3, we control for differences between schools, with respect to school type, year, and region. Our results indicate that professionalization of teacher's innovations, process innovations and education chain innovations positively influence productivity and student performance. These findings can be explained by the facilitation of better performance of students by IT, more highly motivated professionalized teachers who influence the performance and motivation of their students, and the self-selection of good students into schools that have good relationships with primary schools and higher education. Furthermore, the model shows that innovations are related to student performance, but not to costs shares

of, for example, managers or teachers. Finally, despite the innovations, Chapter 5 also shows that productivity has decreased between 2005 and 2007.

In Chapter 6 a plain nonparametric efficiency model is applied to study secondary school efficiency in the Netherlands. This chapter also includes the five innovation clusters. The results suggest that new courses/profiling, pedagogic/didactic, process and education chain innovations are all positively related to school efficiency. Furthermore, both the number of locations per school and the number of schools per governing body (two measures of scale) are negatively related to school efficiency. We also see that efficiency scores differ per school type and per region. The former could be explained by the increased share of bureaucracy in larger schools and governing bodies. The finding on the positive influence of innovations is similar to what is found elsewhere in the literature.

Overall, the conclusion is that Dutch secondary schools are quite efficient, having an average efficiency score of 70 per cent. This means that on average schools can increase student performance by a maximum of 30 percentage point. Furthermore, we can conclude that Dutch secondary schools are increasingly implementing innovations. This study shows that innovations positively influence educational performance and efficiency. Other factors which influence the efficiency scores are the allocations of resources, region, scale, and school type.

About a quarter of Dutch secondary schools could increase their efficiency. This could be realized by, among other things, a different allocation of resources, as about half of the secondary schools have either too few or too many teachers, managers or support personnel, although these numbers are very small for most schools. Furthermore, a higher efficiency score can be achieved by investing in innovations.

Finally, we conclude that several productivity measures have shown that, despite the rather high average efficiency score and the innovations, productivity of Dutch secondary schools has decreased between 2002 and 2007. This decrease in productivity can be explained by a decrease in student performance.

Samenvatting

Productiviteit in het onderwijs is een onderwerp waarover een verhit debat wordt gevoerd tussen politici, docenten, onderwijsbonden en werkgevers in het onderwijs. Productiviteit drijft onze economische groei. Daarom is het niet alleen belangrijk om te focussen op het verhogen van de productiviteit, maar ook om de ontwikkeling van productiviteit te monitoren en te meten. Daarnaast wordt het ook in het onderwijs steeds belangrijker om verantwoording af te leggen en transparant te zijn. Ook om deze redenen is het belangrijk om inzicht te krijgen in de ontwikkeling van de productiviteit in het onderwijs.

Productiviteit is de verhouding tussen *outputs* (bijvoorbeeld leerling-prestaties) en *inputs* (bijvoorbeeld kosten per leerling). De verschillende *inputs* en *outputs* die gebruikt worden bepalen de focus en vaak ook de kwaliteit van de productiviteitsmeting. Vaak wordt de verandering van de productiviteit over de tijd gemeten, om inzicht te krijgen in de ontwikkeling van de productiviteit. Als je productiviteit meet ten opzichte van andere scholen dan noemen we dit efficiëntie. Efficiëntie is productiviteit vergeleken met de best presterende school en is dus een relatieve maat. Veel scholen investeren in innovaties en dit zou de productiviteit en efficiëntie kunnen verhogen. Een innovatie is in dit onderzoek gedefinieerd als: 'een

idee, proces of product dat nieuw is voor de organisatie op het moment dat de organisatie het introduceert'.

Er zijn veel studies naar productiviteit en efficiëntie van scholen. De verschillen tussen deze studies, met betrekking tot zowel inhoud als methode, zijn groot. Ondanks het belang van inzicht in allocatieve efficiëntie is er maar weinig onderzoek naar gedaan. Allocatieve efficiëntie geeft inzicht in de manier waarop scholen hun geld uitgeven en hoe het heralloceren van de beschikbare middelen kan leiden tot hogere prestaties van de school.

Ondanks dat is onderzocht welke factoren de efficiëntie beïnvloeden, is er niet veel literatuur over de relatie tussen innovaties in het onderwijs en de prestaties en efficiëntie van scholen. De meeste studies naar innovaties in het onderwijs onderzoeken geen effecten van innovaties en gebruiken geen kwantitatieve analyses om de relatie tussen innovaties en productiviteit te onderzoeken.

De meerwaarde van dit proefschrift zit in het onderzoeken van de (allocatieve) efficiëntie van Nederlandse voortgezet onderwijs scholen en in het onderzoeken van de invloed van innovaties op deze efficiëntie. Hiervoor gebruiken we innovatieve empirische methoden en een unieke dataset met informatie over innovaties in het onderwijs, geaggregeerd op schoolniveau.

De hoofdbevindingen van het onderzoek zijn dat scholen in het Nederlandse voortgezet onderwijs gemiddeld redelijk efficiënt zijn en dat deze efficiëntie positief wordt beïnvloed door innovaties. Daarnaast beïnvloeden ook de allocatie van middelen en de schaalgrootte de efficiëntie van scholen en verschillen efficiëntiescores per regio en per schooltype.

In hoofdstukken 2 en 3 van dit proefschrift gaan over de efficiëntie van het onderwijs. De meta-analyse in hoofdstuk 2 geeft een overzicht van alle efficiëntiestudies over het onderwijs. Hierin zijn scholen in het basisonderwijs, voortgezet onderwijs en het hoger onderwijs opgenomen. In dit hoofdstuk laten we zien dat de efficiëntie van onderwijsinstellingen ongeveer 0,804 is. Dit betekent dat

scholen gemiddeld 20 procentpunt minder onderwijsprestaties leveren dan potentieel mogelijk zou zijn, gezien de beschikbare middelen. We testen op en corrigeren voor ontbrekende studies (publicatie-bias). De gecorrigeerde analyse bevestigt dat efficiëntiescores afhangen van studie- en populatieverschillen, zoals de gekozen methode, het land waar het onderzoek is uitgevoerd en de gekozen *inputs* en *outputs*. Er zijn diverse mogelijke oorzaken voor verschillen in efficiëntiescores. De gemiddelde efficiëntie blijkt lager te zijn voor het basis- en voortgezet onderwijs, vergeleken met het hoger onderwijs. Een verklaring hiervoor is dat hogeronderwijsinstellingen meer onderling vergelijkbaar zijn, zowel op het gebied van de studentenpopulatie als in organisatiestructuur. De gemiddelde efficiëntiescore is lager voor eerder gepubliceerde studies ten opzichte van recentere studies, omdat eerdere studies vaak minder *inputs* en *outputs* opnemen in de analyses. De gemiddelde efficiëntiescores zijn ook lager in Europese studies, ten opzichte van Noord-Amerikaanse studies. Reden hiervoor kan zijn dat Noord-Amerikaanse studies vaak schooldistricten onderzoeken, terwijl Europese studies vaak naar scholen zelf kijken. Tenslotte laat de analyse zien dat de onderzoekseenheid die gekozen is (districten, scholen of departementen) de efficiëntiescore ook beïnvloedt. Studies naar het hoger onderwijs gebruiken vaak departementen, en vinden vaak hogere efficiëntiescores, omdat er in het hoger onderwijs zowel onderzoek- als onderwijsoutput geleverd wordt en het efficiënter zou zijn om op een van beide te concentreren.

In hoofdstukken 3 tot en met 6 onderzoeken we het Nederlands voortgezet onderwijs. In hoofdstuk 3 bestuderen we de efficiëntie van scholen en de optimale verdeling van het geld over managers, docenten, ondersteunend personeel en materialen in het voortgezet onderwijs. Met het oog op het voorspelde lerarentekort kan dit tot belangrijke inzichten leiden. We gebruiken een parametrische productiefunctie om de efficiëntie te berekenen van 448 scholen tussen 2002 en 2007. In het model nemen we aan dat scholen hun onderwijsprestaties proberen te verhogen met het budget dat ze te besteden hebben, waarbij ze dit geld vrij kunnen

besteden. We kijken of het geld optimaal verdeeld is over managers, docenten, ondersteunend personeel en materialen. Omdat er vele verschillen zijn tussen scholen houden we in de analyse rekening met het schooltype, de regio en het jaar. De resultaten laten zien dat de gemiddelde efficiëntie van scholen in het voortgezet onderwijs ongeveer 70 procent is. Dit betekent dat, gegeven het huidige budget, scholen de prestaties gemiddeld met ongeveer 30 procentpunt kunnen verhogen. Daarnaast zien we in dit hoofdstuk ook dat scholen gemiddeld 0,21 procent teveel docenten in dienst hadden in 2007, terwijl er te weinig management en ondersteunend personeel was. Gemiddeld genomen betekent dit dat scholen met een docent minder zouden kunnen. Ze zitten dus dicht bij het optimum. Een reden hiervoor kan zijn dat scholen geanticipeerd hebben op het voorspelde lerarentekort. Een klein deel van de scholen heeft veel te veel of veel te weinig docenten in dienst. Voor die scholen geldt wel dat ze door hun financiële middelen anders te besteden zelf een deel van het lerarentekort kunnen aanpakken. Hoofdstuk 3 laat tenslotte nog zien dat voor deze groep scholen de productiviteit gedaald is tussen 2002 en 2007, als we kijken naar het gemiddeld rendement en de gemiddelde examencijfers van leerlingen ten opzichte van de gemaakte kosten. Deze productiviteitsdaling kunnen we verklaren door een daling van het Nederlandse onderwijsniveau die recentelijk is vastgesteld in enkele onderzoeken.

In hoofdstukken 4 tot en met 6 onderzoeken we innovaties in het voortgezet onderwijs en de relatie tussen innovaties en onderwijsprestaties. In deze hoofdstukken onderscheiden we vijf innovatiecategorieën: nieuwe vakken/profilering innovaties, pedagogisch/didactische innovaties, procesinnovaties, docent professionalisering innovaties en onderwijsketeninnovaties. Innovaties in de categorie nieuwe vakken/profilering zijn gerelateerd aan de introductie van nieuwe vakken, zoals Wiskunde C en D, en aan het profiel van een school, zoals een LOOT-school of een hoogbegaafdheidsschool. Pedagogische/didactische innovaties zijn de diensten waar een school gebruik van maakt (bijvoorbeeld psycholoog of logopedist), projecten en leerling-innovaties, zoals leerling als coach.

Procesinnovaties zijn ICT, infrastructurele en organisatorische innovaties. De categorie professionalisering van de docenten bevat alle innovaties die gerelateerd zijn aan het docentenkorps. De onderwijsketeninnovaties bestaan uit innovaties gericht op contact met primair of vervolgonderwijs.

In hoofdstuk 4 bestuderen we hoe innovaties zich over scholen hebben verspreid en welke determinanten van innovaties er aan te wijzen zijn. Eerst beschrijven we het verspreidingspatroon van de vijf innovatieclusters. In 2002 waren er maar enkele scholen met innovaties, maar in 2007 zijn de meeste innovaties verspreid over een groot deel van de scholen. Uit de resultaten komt verder naar voren dat de scholen die onderdeel uitmaken van een groter bestuur of met een groter aantal leerlingen meer innoveren dan instellingen die kleiner zijn. Ook blijkt dat de innovativiteit van scholen verschilt per schooltype. Andere determinanten van innovaties zijn: meer concurrentie en een progressieve manier van lesgeven (waarbij het traditionele klassikale doceren wordt losgelaten). Dit betekent dat grotere scholen die op een progressieve manier lesgeven en veel concurrentie hebben, meer innoveren. De reden dat een grotere schaal het aantal innovaties op een school positief beïnvloedt, is omdat schaal vaak gerelateerd is aan een verhoogde mate van professionalisering. Dit betekent over het algemeen dat scholen dan meer informatie beschikbaar hebben over de mogelijke innovaties en de kosten en opbrengsten van deze innovaties. Scholen met meer concurrentie moeten zich extra inzetten om interessanter te zijn voor de leerling dan andere scholen. Dit kunnen ze onder andere doen door te investeren in nieuwe innovaties en zich beter te profileren.

In hoofdstuk 5 bestuderen we de relatie tussen innovaties en de productiviteit van het voortgezet onderwijs. Ook hier gebruiken we een parametrische productiefunctie. De vijf innovatieclusters zijn direct toegevoegd aan het model om de relatie met onderwijsproductiviteit te onderzoeken. De innovaties worden gecorrigeerd voor kenmerken waarvan we in hoofdstuk 4 gevonden hebben dat ze de innovativiteit van scholen beïnvloeden (o.a. schaal en concurrentie). Ook in

hoofdstuk 5 houden we in de analyse per school rekening met het schooltype, de regio en het jaar, vergelijkbaar met hoofdstuk 3. We analyseren 126 scholen. De resultaten laten zien dat professionalisering van docenten en onderwijsketeninnovaties een positieve invloed hebben op productiviteit. Dit kunnen we verklaren doordat gemotiveerde en geprofessionaliseerde docenten ook hun leerlingen beter kunnen motiveren en lesgeven. Daarnaast trekken scholen met veel relaties met het bedrijfsleven en vervolgopleidingen bepaalde typen leerlingen aan, die vaak erg gemotiveerd zijn. Het model laat verder zien dat innovaties wel gerelateerd zijn aan de leerling-prestaties, maar niet aan de aantallen managers en docenten. Echter, ondanks de positieve relatie tussen innovaties en onderwijsproductiviteit is de onderwijsproductiviteit gedaald tussen 2005 en 2007.

In hoofdstuk 6 gebruiken we een eenvoudig, niet-parametrisch model om efficiëntie en innovaties in het voortgezet onderwijs te bestuderen. Ook hier gebruiken we weer de vijf innovatieclusters om efficiëntie te verklaren. De resultaten laten zien dat alle innovatiecategorieën (nieuwe vakken, pedagogisch/didactische, procesinnovaties en onderwijsketeninnovaties) een positieve invloed hebben op efficiëntie. Verder zien we dat het aantal locaties en aantal scholen per bestuur een negatieve invloed hebben op efficiëntie en dat ook schooltype en regio van invloed zijn op de efficiëntie. De negatieve invloed van grotere scholen kunnen we verklaren door de verhoogde mate van bureaucratie die aanwezig is bij grotere scholen en besturen. De positieve invloed van innovaties wordt onderstreept door bevindingen in de literatuur.

De algemene conclusie van het onderzoek is dat Nederlandse scholen een redelijk hoge gemiddelde efficiëntie hebben van 70 procent. Dit betekent dat scholen gemiddeld genomen de leerlingprestaties met maximaal 30 procentpunt kunnen verbeteren. Daarnaast is de conclusie dat scholen in het voortgezet onderwijs in toenemende mate met innovaties bezig zijn. Dit onderzoek laat zien dat innovaties een positieve invloed hebben op onderwijsprestaties en efficiëntie. Andere factoren

die de efficiëntie beïnvloeden zijn de allocatie van middelen, regio, schaal en schooltype.

Voor ongeveer een kwart van de VO-scholen is een grote efficiëntieverbetering mogelijk. Dit is onder andere te realiseren door het personeelsbestand aan te passen, aangezien een kwart van de scholen teveel of te weinig docenten, managers of ondersteuning heeft. Als de overheid of scholen zelf de efficiëntie willen verbeteren dan kan dit zowel door het geld anders uit te geven als door te investeren in innovaties.

Tenslotte concluderen we dat meerdere productiviteitsmaten laten zien dat, ondanks de redelijk hoge gemiddelde efficiëntie van scholen en de innovaties, de productiviteit van het voortgezet onderwijs in Nederland gedaald is tussen 2002 en 2007. Deze productiviteitsdaling is te verklaren door een daling van de leerprestaties.

Biography

Carla Haelermans was born in Swalmen, the Netherlands, on October 20, 1983. She graduated from high school at the Valuascollege in Venlo in 2002. She did a bachelor in International Economics Studies at Maastricht University between 2002 and 2006. In 2005-2006 she studied at Otterbein College, Westerville, Ohio, USA. She graduated from the master International Economic Studies at Maastricht University in July 2007. During her master year she also did an internship at the Ministry of Social Affairs and Employment.

In September 2007 she started working as a PhD-researcher at IPSE Studies (Innovations and Public Sector Efficiency Studies) at Delft University. From September 2010 on, she finished her PhD-studies at the Top Institute for evidence-based Education Research (TIER) at Maastricht University. She continued to be affiliated with IPSE Studies, Delft University. In 2012, during the last months of her PhD-studies, she was a visiting researcher and lecturer at the Polytechnic Institute of Milano, Italy. Currently, she is a post-doctoral researcher at the Top Institute for evidence-based Education Research at Maastricht University.

Top Institute for Evidence Based Education Research (TIER)

The Top Institute for Evidence Based Education Research (TIER) is an inter-university institute that conducts research to develop evidence based education. The institute has three partners: the University of Amsterdam, Maastricht University and the University of Groningen and is located in Amsterdam, Maastricht and Groningen.

TIER contributes to the improvement of the quality of education in the Netherlands by promoting an evidence based approach as a guiding principle in education policy and practice. It accomplishes this by developing (cost) effective education interventions that are grounded in sound scientific research. TIER research is funded by the Ministry of Education, Culture and Science and the participating universities through the NWO and complies with the quality standards and evaluation procedures used by NWO.

The following books recently appeared in the TIER Research Series:

- I. C. Haelermans, *On the productivity and efficiency of education. The role of innovations in Dutch secondary education*