UNIVERSIDAD DE MÁLAGA
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES
PROGRAMA DE DOCTORADO EN ECONOMÍA Y EMPRESA

ESSAYS ON ECONOMICS OF EDUCATION FROM THE PERSPECTIVE OF MICROECONOMETRICS AND OPTIMIZATION

TESIS DOCTORAL
(PhD Dissertation)

Estudiante de doctorado (PhD Student): Luis Alejandro López Agudo

Director y Tutor (Supervisor and Tutor): Óscar David Marcenaro Gutiérrez

Codirector (Co-Supervisor): Mariano Luque Gallego

Año (Year) 2017

UNIVERSIDAD DE MÁLAGA

AUTOR: Luis Alejandro López Agudo

http://orcid.org/0000-0002-0906-3206

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

PhD dissertation

"Essays on Economics of Education from the perspective of microeconometrics and optimization"

PhD Student: Luis Alejandro López Agudo

PhD Program in Economics and Business [*Programa de Doctorado en Economía y Empresa de la Universidad de Malaga*]. Faculty of Economics and Business. University of Málaga. 2017

Supervisor and Tutor: Óscar David Marcenaro Gutiérrez

University of Málaga. Faculty of Economics and Business. Department of Applied Economics (Statistics and Econometrics). Plaza de El Ejido s/n, 29013, Málaga (España); e–mail: odmarcenaro@uma.es. Tel: +34 952137003, Fax: +34 952137262.

Co-Supervisor: Mariano Luque Gallego

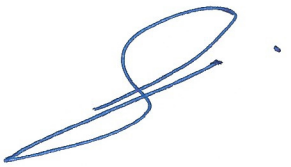
University of Málaga. Faculty of Economics and Business. Department of Applied Economics (Mathematics). Plaza de El Ejido s/n, 29013, Málaga (España); e–mail: mluque@uma.es. Tel.: +34 952131173; Fax: +34 952132061.

INFORME DEL DIRECTOR DE TESIS PARA LA AUTORIZACIÓN DE DEFENSA DE TESIS DOCTORAL

D. Oscar David Marcenaro Gutiérrez, Director de la tesis doctoral de D. Luis Alejandro López Agudo, **informa favorablemente** la solicitud de autorización de defensa de la tesis doctoral con el título "*Essays on Economics of Education from the perspective of microeconometrics and optimization*" presentada por dicho doctorando y que opta a la mención internacional.

La citada tesis, realizada por compendio de artículos, reúne todos los requisitos exigidos por el Programa de Doctorado en Economía y Empresa de la Universidad de Málaga, tanto en lo referido a su redacción, desarrollos metodológicos y aplicaciones en el campo concreto de la Economía de la Educación. En tal sentido supone una contribución de indudable relevancia al ámbito científico.

Y para que así conste y tenga los efectos oportunos, en Málaga a 15 de Septiembre de 2017.



Fdo. Dr. Oscar David Marcenaro Gutiérrez
Profesor Titular del Departamento de Economía Aplicada (Estadística y Econometría)
Universidad de Málaga

INFORME DEL CODIRECTOR DE TESIS PARA LA AUTORIZACIÓN DE DEFENSA DE TESIS DOCTORAL

D. Mariano Luque Gallego, Codirector de la tesis doctoral de D. Luis Alejandro López Agudo, **informa favorablemente** la solicitud de autorización de defensa de la tesis doctoral con el título "*Essays on Economics of Education from the perspective of microeconometrics and optimization*" presentada por dicho doctorando y que opta a la mención internacional.

Esta tesis, realizada por compendio de artículos, reúne todos los requisitos exigidos por el Programa de Doctorado en Economía y Empresa de la Universidad de Málaga, tanto en lo referido a su redacción, desarrollos metodológicos y aplicaciones en el campo concreto de la Economía de la Educación. En tal sentido supone una contribución de indudable relevancia en el avance científico.

Y para que así conste y tenga los efectos oportunos, en Málaga a 15 de Septiembre de 2017.

Fdo. Dr. Mariano Luque Gallego
Catedrático de Universidad del Departamento de Economía Aplicada (Matemáticas)
Universidad de Málaga

AUTORIZACIÓN DE DEFENSA DE LA TESIS DOCTORAL POR EL TUTOR

D. Oscar David Marcenaro Gutiérrez, Tutor de D. Luis Alejandro López Agudo, en el programa de doctorado de Economía y Empresa de la Universidad de Málaga, a la vista de las razones expuestas en el informe de los Directores de la tesis, **ratifica el referido informe para la autorización de la defensa doctoral presentada por dicho doctorando.**

Y para que así conste y tenga los efectos oportunos, en Málaga a 15 de Septiembre de 2017.

Fdo. Dr. Oscar David Marcenaro Gutiérrez
Profesor Titular del Departamento de Economía Aplicada (Estadística y Econometría)
Universidad de Málaga

Acknowledgements

This thesis would not have been possible without the support of many people, who I am afraid I would not be able to thank enough for their help during this big part of my life. First and foremost, I offer my sincerest gratitude to my PhD supervisors: Oscar Marcenaro, for sharing all his expertise in Economics of Education with me, putting all his effort, commitment and infinite patience in helping me to mature in the researcher and personal way; and Mariano Luque, who has enriched my formation with his great Mathematical knowledge, which he has provided to me in the easiest way, also supporting and guiding my academic and personal formation during all this PhD stage.

I would also like to acknowledge that it has been a pleasure to work with a great expert in Economics as Miguel Ángel Ropero, coauthor in one of the works of this thesis, and with John Jerrim and Nikki Shure, who have contributed a lot to my formation as researcher and made me feel that London was my second home.

I would also like to thank the Department of Applied Economics (Statistics and Econometrics) for all its support. I would also like to express my sincere gratitude to the *Ministerio of Education, Cultura y Deporte*, for believing in my work and supporting my PhD research, and to the excellent training received from the University of Malaga PhD Program in Economy and Business [*Programa de Doctorado en Economía y Empresa de la Universidad de Malaga*].

Special appreciations to my family, beginning with my parents, for always supporting me, including when I looked like the furniture of my room, and to my brother, who made me feel I was never working alone. I also thank my grandparents, who always believed in me and understood my decisions, especially when I went abroad in my internship for some time. I would also like to thank my school and faculty friends, who have always been there, even when they did not see me for a long time.

To all of them, thank you from the bottom of my heart.

Funding and additional acknowledgements

The works in this thesis have been partly supported by the Spanish Ministry of Economy and Competitiveness (Research Project ECO2014-56397-P); by the Andalusian Regional Ministry of Innovation, Science and Enterprise (PAI group SEJ-532); the scholarship FPU2014 04518 of the Ministry of Education, Culture and Sports [*Ministerio de Educación, Cultura y Deporte*]; and the research internships at the Institute of Education of the University College London (UCL) in 2016 and 2017 funded by the Ministry of Education, Culture and Sports [*Ministerio de Educación, Cultura y Deporte*] for the scholarship FPU2014 04518.

Index

Introduction

I.1. The current research

This thesis belongs to the specific field of Economics of Education, an area of research within Economics which is based on the application of econometric techniques to analyse the education reality. Particularly, this thesis intends to evaluate different aspects of the teaching-learning process in primary and secondary education of the Spanish education system, with special emphasis on the region of Andalusia. Once this evaluation is performed, the next aim of this thesis is to provide proposals on the design of policies – for these education systems – to accomplish the following education-related objectives: to overcome the existent gender differences in academic achievement, to help students get a balanced academic achievement, to provide education systems with effective and satisfied teachers – who can enhance students' learning and engagement – and to improve the knowledge of education agents and researchers about the implications of using data from new education assessment methods.

I.2. Why is the quality of education systems important?

The focus of this thesis has been placed on education due to its importance in society: as it has been highlighted by Hanushek and Woessmann (2007)[1], it plays an essential role in providing well-being to the population, what places the quality of education in a relevant position in the political debate. This quality of education is so closely related to the development of population's cognitive skills that authors as Hanushek and Woessmann (2007) used them as its proxy. To the extent that these skills affect wages and, hence, individuals' well-being, they will have a relevant influence on economic growth (Klassen, 2002; OECD, 2015). Improving the quality of education also increases the population's human capital, hence providing higher productivity and higher levels of outputs, being finally translated into economic growth (Mankiw, Romer, & Weil, 1992). In addition, this increase in the quality of education supposes the development of better technologies, what will also provide economic growth (Aghion & Howitt, 1998). Furthermore, cognitive skills and educational performance are relevant in other aspects of individual's lives as, e.g., they are a good predictor of the career progression of men and women during their adult lives (see e.g. Dolton, Makepeace, & Marcenaro-Gutierrez, 2005; Vignoles, De Coulon, & Marcenaro-Gutierrez, 2011); they can also significantly affect the fertility rates of a country (Basu, 2002) due to the job security, flexibility and salary that adults can obtain (thanks to their level of education), what may increase their willingness to have more children.

However, Hanushek and Woessmann (2007) highlighted that, when analysing education quality, it is essential to have into account particular characteristics of the country under analysis, as ignoring them can suppose reaching to misleading conclusions about the relationship between education and economic outcomes. This problem may be the consequence

[1] These authors have made a great contribution to the literature on Economics of Education.

of differences in terms of returns of education between countries, being these returns higher in low income countries and with a lower schooling initial level (Psacharopoulos & Patrinos, 2004). Thus, this thesis keeps the focus on the economic and education contexts of the country – Spain – and region – Andalusia – in which the analyses are being performed.

I.3. The increasing awareness about the relevance of education

From the latter, it seems very obvious that education is essential in the socio-cultural and economic development of a country. However, it was not until the appearance in the middle nineties of PISA[2] and TIMSS[3]/PIRLS[4] reports[5], conducted by organizations as the OECD[6] or IEA[7], respectively, that participating countries – and those willing to participate – developed a high competitive spirit in terms of performance in these reports. Each cycle of these and other ILSA tests has mainly evaluated student competencies in subjects as reading, mathematics and science – among others as, e.g., problem solving or financial competences – in different levels of compulsory education.

The special focus on these three competences can be found in the relevance that they present for the personal development of individuals and their ability to deal with real-life issues: in the case of the reading competence, it greatly influences students' academic achievement in other subjects, as it involves the understanding of texts; in addition, it is also relevant for adults, as it increases their participation in society and help them to obtain average or higher wages (Cunningham & Stanovich, 1997; Smith, Mikulecky, Kibby, Dreher, & Dole, 2000). Furthermore, a good level of reading competences not only supposes social advantages, but also a better well-being (Friedman, 2005). In the case of the mathematics competence, it is essential to describe, explain and predict phenomena (OECD, 2014a). As indicated by OECD (2010a), daily real-world experiences as shopping, cooking, dealing with personal finances, etc., require some kind of mathematical capacities. In addition, dealing with media content supposes the understanding of documents which include tables, charts, etc., while dealing with timetables or making transactions require a certain level of mathematic competences in order to interpret their content. Finally, science competence is also relevant, due to its influence in the preparation of students for life in modern society, as it provides citizens with the ability to participate in decisions which are related to science (OECD, 2014a) as, e.g., those which imply the use of technology, the development of certain biological procedures, etc. Hence, due to the relevance

[2] Programme for International Student Assessment.

[3] Trends in International Mathematics and Science Study.

[4] Progress in International Reading Literacy Study.

[5] This kind of studies which are oriented to the assessment of students' competences and academic achievement with an international focus are known as International Large Scale Assessments (ILSA) tests.

[6] Organization for Economic Co-operation and Development.

[7] International Association for the Evaluation of Educational Achievement.

that these competences have in developing the cognitive skills of the population, this thesis will deal the three of them[8].

I.4. The situation of the Spanish education system

Moving into the results obtained by the country which is the focus of this thesis – Spain – in the most recent cycles of these ILSA tests, they have shown an alarming reality, which has motivated the election of the case study of this country: Spanish students obtain a low performance in these three competences, compared to the students of the best performing countries. Considering that PISA scores – for tenth grade students – have a mean of 500 and a standard deviation of 100, in the case of PISA 2015 (OECD, 2016) Spain presents 486 points in mathematics, which is overcome by the 564 points (a difference of 0.78 standard deviations) of Singapore and 548 (0.62 standard deviations) of Hong-Kong (China), being the average of the OECD 490 (0.04 standard deviations higher); Spain presents 496 points in reading, 0.39 standard deviations under Singapore and 0.31 under Hong-Kong (China), being the average of the OECD 0.03 standard deviations lower; and 493 in science, being that of Singapore 0.63 standard deviations higher and 0.45 standard deviations higher in the case of Japan, while the average of the OECD is the same as in Spain. In the case of TIMSS and PIRLS, considering that both have also scores with a mean of 500 and a standard deviation of 100, Spain obtained in PIRLS 2011 – for fourth grade students – 513 points in reading, compared to 571 points (0.58 standard deviations higher) of Hong Kong and 568 (0.55 standard deviations higher) of Russian Federation (Mullis, Martin, Foy, & Drucker, 2012). For TIMSS 2015 (fourth grade), Spain obtained 505 points in mathematics, compared to Singapore, which performed 1.13 standard deviations higher, and Hong-Kong, performing 1.10 standard deviations higher (Mullis, Martin, Foy, & Hooper, 2016); in science, Spain obtained 518 points, while Singapore performed 0.72 standard deviations higher and the Republic of Korea 0.71 standard deviations higher (Martin, Mullis, Foy, & Hooper, 2016).

In addition to these figures, OECD (2016) provides information about the classification of students in the performance distribution according to their scores in PISA[9]. In this sense, 16.2% of Spanish students who participated in PISA 2015 were considered as low performers – not achieving basic standards of learning – in reading, while only 5.5% scored in the top of the distribution. In the case of mathematics, 22.2% of Spanish students obtained low scores, while only 7.3% got top results. Finally, in the science subject, 18.3% of students scored in the low part of the distribution and 5% in the top. Hence, these results are denoting that Spanish

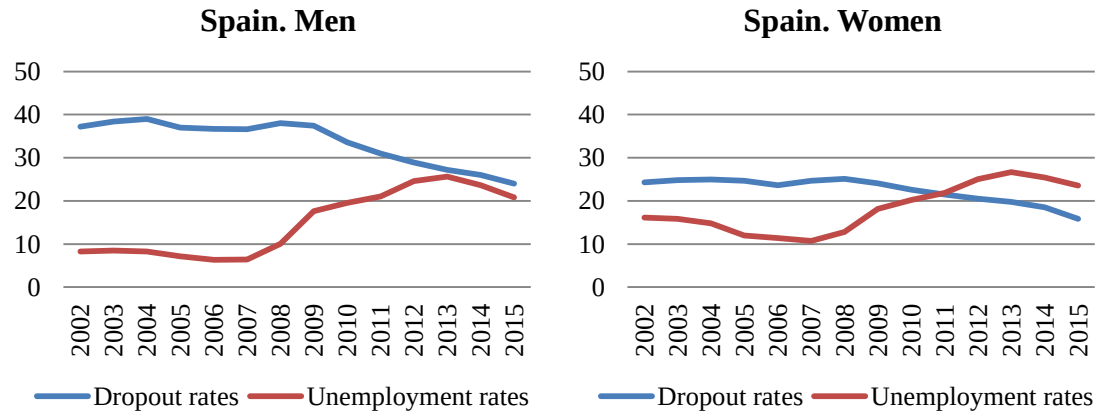
[8] Due to the high correlation that the science competence has with mathematics (Rivkin & Schiman, 2015), the former competence occupies a secondary role in this thesis, so it is only studied in the last section of the last chapter.

[9] Concretely, the cut point for lowest performers (level 2 cut point) in reading is a score of 407, while for mathematics it is 420 and for science 410. In the case of top performers (level 5 cut point), it is a score of 626 in reading, 607 in mathematics and 633 in science.

students, besides from performing lower than students of other countries, have a distribution of performance which is skewed to the left, i.e., Spanish students are low performers on average.

This trend of low academic achievement in Spain is not only reflected in ILSA test results, but also in the figures of early leaving education, which is one of the main indicators of educational achievement. This early leaving education presented a rate of 20% (24% for male students and 15.8% for female students) for Spanish students in 2015 (IECA, 2017a), which was higher than in the rest of the EU-28 countries in 2015 (a total of 11%, 12.4% for males and 9.2% for females; Eurostat, 2017). Although these premature dropout rates are still high, they are not so much as in 2008, when they reached a total of 31.7%. These figures have begun to decrease in this economic context of crisis we are living nowadays, as this crisis has drastically reduced the employment opportunities (according to the EPA, 2017, the unemployment rate increased from 13.79% in the fourth term of 2008 to 20.90% in the fourth term of 2015). Thus, these high unemployment rates have, at the same time, reduced the opportunity cost of staying at the education system, meaning that the Spanish population is seeking to solve their lack of required qualifications increasing their education level, in order to facilitate their incorporation to the labour market. In this sense, Figure 1 shows the relationship between annual dropout and unemployment rates for the period 2002-2015, from where it can be appreciated the increase in unemployment rates and the indicated fall in dropout rates (being this effect even higher for men). This importance of the education level in employment finding can be appreciated, for example, in the figures of Ivie (2013) report, which offers information about the relationship between the unemployment rates and the highest obtained level of education of Spanish workers: those workers who presented in 2013 a primary or lower level of education had an unemployment rate of 39.5%, which was reduced while education level increased, until reaching to 12.8% for workers with a University degree.

Figure 1. Annual Spanish dropout and unemployment rates from 2002 to 2015

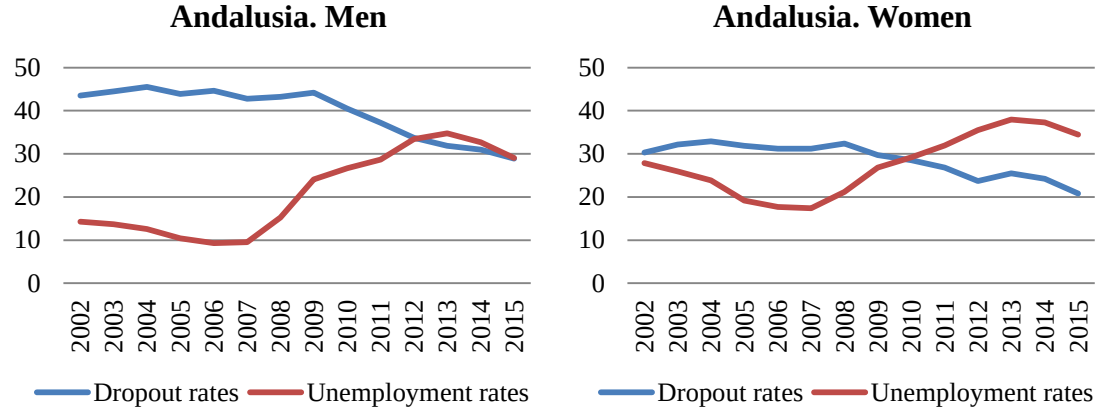


Source: Author's own elaboration from IECA (2017a) and EPA (2017).

I.5. The situation of the Andalusian education system

Hence, all these figures seem to indicate that the Spanish education system needs a deep exploration to determine which factors may be driving to these poor results; once these factors are identified, the education authorities will be able to implement policies to improve the Spanish situation. However, within Spain, there is one region – Autonomous Community – in which this thesis puts special interest, as its situation is even worse compared to that of Spain in general; this is, concretely, Andalusia. In this sense, there are many reasons which make necessary the study of this region: it has obtained scores under the average of Spain in the three competencies evaluated by PISA 2015 – reading, mathematics and science – (OECD, 2016); specifically, it is one of the two worst performing regions in Spain in terms of these subjects, as measured by PISA 2015 test scores (479 in reading, 0.17 standard deviations under Spain; 466 points in mathematics, 0.20 standard deviations under Spain; and 473 in sciences, 0.20 standard deviations under Spain; OECD, 2016). In addition, the early dropout rates of Andalusia are even higher than those of Spain, with figures of 28.9% and 20.8% for both male and female students in 2015, respectively (a total dropout rate of 24.9%, overcoming Spanish dropout rates in 4.9%; IECA, 2017a). Figure 2 makes a similar analysis to that of Figure 1 but for the case of Andalusia, showing the relationship between dropout and unemployment rates for the period 2002-2015. It presents a recent increase in unemployment rates and a fall in dropout rates (being this effect even higher for men), as it happened for Spain.

Figure 2. Annual Andalusian dropout and unemployment rates from 2002 to 2015



Source: Author's own elaboration from IECA (2017a; 2017b) and EPA (2017).

Furthermore, according to Fundación Foessa (2016), Andalusia was the Spanish region with higher income inequality in 2014[10] (Gini Income Index of 34.69). Last, Andalusia was the second region with the lowest GDP per capita in Spain in 2014 (INE, 2015) and its unemployment rates are also much higher than the Spanish ones – a total of 26.94% for

[10] This high inequality affects many areas of society – as education – having an impact on economic growth through the level of preparation of workers who access to the labour market (Lipman, 2003).

Andalusia, 24.22% for men and 30.25% for women, compared to a total of 18.75% for Spain, 17.22% for men and 20.51% for women, for the first term of 2017 (IECA, 2017b). These figures become even more alarming when having into account that Andalusia is the most populated region of Spain, meaning that this poor situation will affect many children.

Besides from the demographic and socio-economic characteristics of this region, another reason to study the case of the Andalusian education system in this thesis is the novelty that presents the availability of a rich database for this region. Concretely, it is the Social Survey 2010: Education and Housing – ESOC10 – (IECA, 2010), which was focused on Andalusia and was conducted among 11-12 and 15-16 year old students and their families – constructed with information on a wide set of personal, family and school environment characteristics. This ESOC10 was linked to the results from the administrative records (SENECA) of teacher-based scores, provided by the *Consejería de Educación de la Junta de Andalucía*. This combined database is going to be referred from now on as ESOC10-SEN. Thus, the mixture of a measurement of the Andalusian educational output of the teaching-learning process with the wide collection of socio-economic characteristics contained in ESOC10, make ESOC10-SEN an essential element to go into depth on the study of the teaching-learning process in Andalusia.

I.6. The education production function

The presented evidence and figures have pointed out the necessity to improve the unfavourable situation of the Spanish and Andalusian education systems. Thus, this thesis has as general objective the analysis of the education characteristics of this country and region, to contribute to the optimal design of education policies, which utterly aim at improving their academic results. In order to face this ambitious and also demanding objective, this thesis works with a tool called "the education production function", so every methodology and approach in this thesis is focused on studying specific aspects of this function, from different perspectives. But, what does this education production function mean? Overall, the concept of production function has been widely used in Economics in recent years (Solow, 1957) and it is understood as a way to relate the amount of outputs produced per unit of time to the amount of inputs used in producing those outputs. In the field of Economics of Education, the education production function is used as the instrument to model the teaching-learning process and it is usually defined as a function which assigns quantities of measured inputs of school, teacher and student characteristics, to some measure of educational outputs, like students' test scores (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, & York, 1966; Bowles, 1970; Hanushek, 1979).

The analysis of this education production function began with the "Coleman Report", which was released in 1966 by the US government (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, & York, 1966). In this study, the importance of the distribution of students' performance – understood as the output of the educational process – was highlighted, and also

the relationship between this academic achievement and school inputs. Hence, it set the basis in the definition of relevant inputs which compose the specification of the education production function as it is reflected in the current literature. Among these inputs, students' characteristics and family background can be found, which are usually integrated by socio-demographic characteristics such as, e.g., family structure and socio-economic and cultural levels. When classmates' inputs are included – which could cause the so-called "peer effects" – they are typically aggregates of students' socio-demographic characteristics or academic achievement for a school or classroom (average cultural level of students' parents, proportion of immigrant students in the classroom, proportion of girls in a class, etc.). In the case of school inputs, they usually include teacher background (sex, experience, education level, etc.), school organization (class size, time of lessons, materials, etc.) and district or community factors (average expenditure levels, population, etc.).

However, the large literature on the education production function has generally shown conflicting or weak results, what makes more difficult to define adequate policy interventions. These weak results come from many sources of uncertainty as, e.g., the proper specification of the education production function in terms of educational outputs (due to their multidimensional character) or an adequate choice of relevant control variables, functional form and estimation methods for the education production function – although some specifications show consistently significant effects. Many of these specification issues appear due to the potential existence of unobservable variables at the student, school or teacher level. Because of these difficulties, every use of the education production function made in this thesis is argued and supported on previous literature, so reliable results can be obtained.

I.7. The structure of the thesis

At this point, the education systems in which this thesis focuses on and the instrument employed to analyse them have been defined. Nevertheless, the education reality is so broad and composed by lots of agents – which also interact between them – that trying to approach several education issues at the same time could be problematic. Thus, to accomplish the objectives indicated for this thesis, many of the outputs and inputs which compound the structure of the education production function, in Andalusia and Spain, are grouped and studied separately in more detail, focusing the attention in a particular objective in each study. To do this, three elements of the teaching-learning process are analysed, so each one of the three chapters in this thesis is devoted to one of these elements, beginning with students, passing through their teachers, until reaching to the assessment methods used by teachers and institutions to evaluate students. Each one of these three chapters is integrated by two lines of research – devoting one section of the chapter to each line, with the exception of the third chapter, composed by one line of research – oriented to a specific casuistry of the corresponding element. In the following paragraphs a brief outline on the content of these three chapters is presented.

To begin with this three-element structure, the first chapter is dedicated to those who are receiving education and learning in every education system: students. The first line of research in this chapter is focused on the subject of gender differences in academic achievement. Particularly, the potential existence of a differential influence of expectations – and some other background variables – on students' academic achievement by gender for secondary education Andalusian students is analysed. Once gender differences have been considered, the second line of this chapter is devoted to study the way to obtain balanced results in three subjects, by primary and secondary education Andalusian students, also incorporating the distinction between boys and girls in this analysis.

Right after, in the teachers' chapter, the first line of research is dedicated to define and measure teachers' effectiveness and to highlight the relevance of teachers' efficiency on this effectiveness. In addition, evidence on how the interrelation of these two concepts fosters fourth grade Spanish students' engagement in lessons is provided, together with the advantages of using teacher effectiveness as a complement for raw score teachers' rankings. However, the capacity of teachers to make their students learn is not only related to teachers' effectiveness in doing so, but also to other more subjective factors, like their own happiness with their job or their state of mind. Because of that, in the second line of research, teachers' satisfaction is proposed as an education production output. Hence, the optimal balance between teachers' satisfaction and the results obtained by their students is analysed for fourth grade Spanish students.

Finally, in the third and last chapter, the assessment methods used to measure students' academic achievement in secondary education are analysed: concretely, the line of research in this chapter studies the relationship between traditional – paper and pencil – and new digital – computer – assessment methods employed in external evaluations. The objective of this line is to provide an insight into the differences in the measures of students' academic achievement which can arise when using one assessment method or the other, due to potential differences in the kind of students' ability which they are measuring.

After this brief outline on the lines of research which compound this thesis, their content is developed in more detail in what follows.

I.7.1. The students

Beginning with the first chapter, it is devoted to the analysis of the education production function from the perspective of students. Students are firstly analysed because they are the most important resources that every education system possesses and their learning is the ultimate objective of these systems. Only once some clues on the way students learn are obtained, the current research can move on to analyse how to provide this learning – by the study of those agents who provide it, i.e., teachers. In this sense, the first issue under analysis is the different way in which boys and girls learn and how this can affect their academic achievement. The existence of differences in boys' and girls' academic achievement has been recently put into focus by International Large Scale Assessment tests (PISA, TIMSS, PIRLS, etc.). One of the most robust outcomes across ILSA tests – for most of the countries and cycles – is girls' superior performance in reading, compared to boys. On the contrary, it is frequently found that boys outperform girls in mathematics, although this result seems to be more country-specific. These differences can be appreciated, e.g., in the case of PISA 2009 report (OECD, 2010b), in which girls achieve higher average scores in reading comprehension than boys in the sixty five participating countries, but boys' advantage in mathematics is only present in half of the countries. A similar situation has remained in PISA 2012 (OECD, 2014b) and PISA 2015 (OECD, 2016). This also happens in countries which get high results in these assessments. For example, in the case of Finland, which is usually taken as a paradigmatic model of high educational performance, it presented the largest reading gender gap in PISA 2012, with girls outscoring boys by 0.62 standard deviations, although it did not present any differences in mathematics (OECD, 2015, p. 152). In the case of Spain in PISA 2012, the gap is of 0.29 standard deviations in favour of girls in reading and 0.16 standard deviations in favour of boys in mathematics – these trends of gender differences are kept in PISA 2015 (OECD, 2016).

In this context, this perpetuation of gender differences in academic achievement seems to require further analysis, in order to find out which factors can be conditioning this situation. The study of these factors in compulsory education is essential to the better understanding of the roots of women's disadvantaged position in the labour market, what could help to foster their entrepreneurship and representation in companies' management positions[11]. Furthermore, the differentiated academic performance of students by gender can significantly affect the economic growth of a country as, according to OECD (2015), the overall increase of academic achievement in OECD countries in the past 50 years accounts for 50% of economic growth, being more than half of that growth due to women's higher achievement. Because of that, the first line of research in this students' chapter is based on analysing gender differences in

[11] There is also evidence that entrepreneur women are less discriminatory to hire other women for managerial positions, what could mean that promoting female entrepreneurship helps to reduce labor market gender discrimination (see Marques, 2015).

academic achievement. The used approach is focused on social theories which try to explain these gender differences, in contrast with other biological theories[12]. Particularly, this research is focused in one particular variable which may be causing these gender differences, i.e., students' expectations, based on the assumption that children's expectations are closely related to their final academic achievement. To the extent that these expectations have been found to be different for boys and girls (Fortin, Oreopoulos, & Phipps, 2015, for the case of the United States) this could mean that girls are expecting to attend college to a greater extent compared to boys, which might explain why girls are leaving boys behind at school. Because of that, the objective of this section is to check the following hypothesis: Andalusian boys and girls present differences in academic achievement and the dissimilarities in their expectations about their future level of education – and other student and family variables as, e.g., socio-economic status or misbehaviour – are conditioning these differences in academic achievement in a significant way.

In spite of the relevance of studying expectations, their use can be subject to critic, due to their potential endogeneity. However, the Spanish case is different from that of other countries, as the average tuition fees paid by students are, e.g., five times lower than those paid in the United States (OECD, 2014c)[13]. Applying Spence (1973) theory[14], having a university degree could be a good signal of high productivity, to the extent that low-ability workers find it very costly to finish an undergraduate program, so they prefer not going to the university. Because of that, only the most skilled students will find it profitable to study at the university, so they can compensate the cost of their studies with their future salary. This theory can be applied to the American case; however, finishing a degree in Spain does not seem to signal a greater ability, mainly because less skilled students find doing it profitable – even when spending more years studying – as the annual tuition fees they have to pay are much lower than those in the United States. For this reason, children's expectations in Spain may not depend so much on their abilities as in the United States (where students access to University due to their high academic achievement), what reduces expectations' potential endogeneity in Spain. Thus, this situation would make the Spanish case analysis – and, concretely, the Andalusian case – especially interesting, as it would help to avoid endogeneity and thus get reliable results.

Nevertheless, in spite of the relevance of this subject, the scarce information on children's expectations about their future level of education has prevented more researchers from analysing the effect of this variable on children's academic results. Fortunately, the ESOC10-SEN database contains rich information about the number of years of schooling that boys and girls expect to complete in the future (reported directly by the student). This

[12] A review of the literature of both theories can be found in Francis and Skelton (2005).

[13] These figures are expressed in real terms.

[14] This theory states that workers use their level of education to signal their abilities in the labour market.

information, together with the methodology employed, i.e., an improvement of Oaxaca decomposition technique developed by Fortin, Oreopoulos, and Phipps (2015)[15], let to measure the effect of gender differences in children's expectations about the future on the gender gap in academic achievement. Hence, the particular situation of Andalusia, which reduces expectations' endogeneity, together with the combination of both, a rich and new database and an advanced methodology, make this research to contribute in a relevant way to the existent literature on gender differences in academic achievement.

As it has been highlighted, the existence of gender differences in academic achievement remarks that it is necessary to consider the different characteristics that boys and girls present in the way they learn, so this is accounted for in the following research lines. The next issue is focused on analysing which factors can be contributing – or making more difficult – that students reach the best academic achievement that they possibly can. However, aiming at making all students get the highest performance in every subject is rather unrealistic, as students have different learning skills and abilities. Because of that, in the second line of research in this students' chapter, making students obtain an optimal and balanced academic achievement in many competences is the main objective, through the measurement of the relationship between educational inputs and outputs in the educational production function. The relevance of this objective is based on the assumption that students are not equally skilled in terms of learning abilities in different areas, so they cannot reach the highest scores in all subjects and, hence, some kind of trade-off exists between the subjects they are studying – in terms of academic achievement. Thus, the hypothesis to contrast in this section is the existence of a set of educational inputs over which it is possible to act in order to achieve a balanced achievement in three specific subjects – reading (Spanish), mathematics and English language (by primary and secondary education Andalusian students), being the influence of these inputs different by sex.

This hypothesis is contrasted by the optimization of the education production function built by econometric techniques, which represents the relationship between education outputs – in this case, academic achievement in three subjects – and many student inputs as, e.g., socio-economic characteristics, which can condition these outputs. This supposes the combination of econometric techniques with mathematic programming methodologies – concretely, multiobjective programming. The latter methodology is employed due to its capacity to generate efficient solutions through the simultaneous optimization of many objective functions (one for each one of the considered subjects) by using a reference point approach (Miettinen, 1999; Luque, Miettinen, Eskelinen, & Ruiz, 2009; Luque, Ruiz, & Steuer, 2010). This last technique has been employed in many fields (since it presents great mathematic properties), although the

[15] This methodology presents the novelty that it lets to obtain the educational response and composition effects for gender differences in academic achievement.

proposed combination of econometrics and mathematic programming techniques has not been extensively studied in the literature, which has into its seminal works those of Wallenius, Wallenius, and Vartia (1978) – who proposed a multiobjective optimization model for formulating macroeconomics policy decisions in Finland – or Marcenaro-Gutierrez, Luque, and Ruiz (2010) – for the study of labour satisfaction of Spanish workers. Thus, its application to the Andalusian education system would be a novelty, which could suppose an important help in making political decisions focused on obtaining a balanced academic achievement from primary and secondary education Andalusian students. In addition, a sensitivity analysis is performed, in order to measure the robustness of the results to variations in certain factors or to the relative importance given to each subject. The ESOC10-SEN database is employed to determine this profile of the most successful and "balanced" primary and secondary education Andalusian student in terms of these three-subjects' academic achievement, due to the previously indicated rich information that it possesses on students' and their families' characteristics.

I.7.2. The teachers

Although students presented those characteristics which could help them to obtain a balanced academic achievement, it would not be until they get to the school and start their learning process that they could develop their knowledge and competences. In this process, teachers are responsible for transmitting knowledge to students and making them learn, what denotes the important role that teachers play in education systems; this is why teachers are studied in the second chapter. Following Cochran-Smith and Lytle (1999), teacher research has been focused on improving teachers' curriculum and practices to enhance their quality, as it is the key to the improvement of student outcomes (Goldhaber, Brewer, & Anderson, 1999; Rivkin, Hanushek, & Kain, 2005; Goldhaber, Gross, & Player, 2010). However, although many pieces of research have been focused on teacher quality, they fail in delimiting the different aspects which compose it. This is necessary to the extent that a better knowledge of the teaching-learning process would improve the application of education policies, as they would be more focused.

In order to shed light on this issue of delimiting teacher quality determinants, the first line of research in this chapter defines the concepts of teacher effectiveness and efficiency. In spite of the relevance of these concepts, they are usually confused in the Economics of Education literature (Lockheed & Hanushek, 1994). The current research intends to empirically recognise the factors and teaching methodologies which could condition teachers' effectiveness, in which teachers' efficient use of classroom resources might occupy a relevant place. This relationship between teacher effectiveness and efficiency has not been clearly stated – to the best of my knowledge – in the previous literature, so this research would suppose a new insight into the delimitation of these concepts. This delimitation is discussed in the following, together with the relevance of effective teachers in education systems.

The training of effective teachers has been one of the main aims of education systems. Traditionally, the literature has indicated that students' academic achievement is a relevant element to measure the effectiveness of a particular teacher. However, the current analysis focuses on the capacity of teachers to engage students in their lessons as an alternative to student achievement, following the definition of effective teachers proposed by Creemers and Reezigt (1996) and Kyriakides (2008) "effective teachers are expected (…) to maximize engagement rates" (as cited in Muijs, Kyriakides, Van der Werf, Creemers, Timperley, & Earl, 2014, p. 233). Effective teachers are then expected to increase their students' engagement and implication in their studies. This definition of effective teachers can be complemented by that given by Braga, Paccagnella, and Pellizzari (2014), who defined effective teachers as "(…) those who provide their students with knowledge that is useful in future learning, presumably require their students to exert effort by paying attention and being concentrated in class and by doing demanding homework". Thus, an effective teacher has the ability to reach out to the students and make them learn by the use of certain practices, their behaviour in the classroom and the way they conduct their lessons.

With regard to the proposed inclusion of the concept of teacher efficiency[16] in teacher effectiveness in the current line of research, this was also highlighted in the teacher effectiveness definition of Creemers and Reezigt (1996) and Kyriakides (2008), who indicated that effective teachers are expected "(…) to organize and manage the classroom environment as an efficient learning environment" (as cited in Muijs, Kyriakides, Van der Werf, Creemers, Timperley, & Earl, 2014, p. 233). This could mean that teachers' efficiency may be determined by variables related to the way that teachers use classroom environment inputs whose availability is not directly controlled by them, such as the time they spend teaching, the use of the material resources they have, how they cope with a large number of students in the classroom, etc., being the output their students' achievement. This lets to define teachers' efficiency as their ability to perform fruitful work in an organized way, using available time and resources without wasting any of them.

Hence, once delimited these two concepts, the hypothesis to contrast in this section for Spanish teachers is twofold: firstly, effective teachers are also efficient, being this effectiveness a relevant way of measuring teachers' performance; secondly, there are factors which can be classified, both theoretically and empirically, in teacher effectiveness or efficiency factors.

The identification of effective teachers is a relevant issue to the extent that "league tables", which are created by the use of students' average score for each teacher, do not take into account sampling variability and other sources of error (Leckie & Goldstein, 2009). The

[16] It basically means getting the maximum output from the amount of inputs that teachers use in the teaching-learning process, or achieving the same output with a minimum quantity of inputs. This is an own adaptation of Lockheed and Hanushek (1994) – and references therein – general definition of efficient system.

use of a proper classification of teachers in terms of effectiveness is particularly important, so that effective teachers can be recognized and rewarded based on their performance fostering higher levels of student engagement and learning. In this sense, some research has highlighted the importance of acknowledging teachers for their work; e.g. Dolton and Marcenaro-Gutierrez (2011) studied panel data of 39 countries, reaching to the conclusion that higher teacher salaries are positively correlated with higher student outcomes. This has also been found by authors as Loeb and Page (2000) or Lavy (2009), who showed the positive relationship between teacher performance-related salary and their students' performance – together with lower dropout rates. Hence, teachers' rankings based on their effectiveness are going to be proposed as a better way of sorting them according to their performance, in comparison to only using their students' average scores.

The methodology used for this analysis is a two-step procedure which – to the best of my knowledge – supposes a novelty in the study of teacher effectiveness. The first step measures teachers' efficiency by using stochastic frontier analysis (SFA) – a methodology applied by, e.g., Battese and Coelli (1988) or Jürges and Schneider (2007) – employing as output students' academic achievement in reading and mathematics and, as inputs, available school resources for teachers. In the second step, the efficiency scores obtained from the first procedure have been employed, together with teachers' learning practices in the classroom, to explain students' engagement in reading and mathematics lessons, using multilevel regression models – as Afshartous and Wolf (2007) or Grilli and Rampichini (2009) for the case of schools. The databases used for this study are PIRLS and TIMSS for the fourth grade (9/10-year-old) 2011 wave of Spanish students. The reason to use these databases is that they contain a higher amount of teacher effectiveness and efficiency related variables for Spain than other databases. Fourth grade students were chosen as they are more malleable at this age (Thompson-Schill, Ramscar, & Chrysikou, 2009) so they might show a better reflection of teachers' procedures. In addition, Spanish primary school students have the same teacher in each cycle – being one cycle a group of two grades in primary education – and for all subjects, what may suppose that students' engagement could be related to a particular teacher. Furthermore, 2011 is the first year in which the same students have been tested in PIRLS (which evaluates reading) and TIMSS (which assesses mathematics and science), what lets to use information on reading and mathematics for this analysis – as science is highly correlated with mathematics, it is not used (Rivkin & Schiman, 2015).

Furthermore, although effective teachers are required in any education system, it is not only important to find and hire them, but also to attract and retain potential effective teachers into the profession. However, the status that teachers have in Spain is not very high (Dolton & Marcenaro-Gutierrez, 2011), what makes teaching an unattractive career choice. In this sense, teachers' work satisfaction plays a relevant role in motivating them, so that they perform their

work in the better way they can (Klassen & Chiu, 2010). This satisfaction has received little attention in the previous literature and it has been considered as part of the education systems' output, following the arguments of many authors as Kumar (2014), who found a positive relationship between teaching effectiveness at primary level and job satisfaction, meaning that teachers who were highly satisfied with their jobs were the most effective ones in teaching. Allinder (1994) highlighted that those teachers with high levels of planning and organization were open to new ideas and more willing to experiment with new methods, in order to better meet the needs of their students (Cousins & Walker, 1995). They also exhibit enthusiasm for teaching (Allinder, 1994), what can have a positive influence on students' academic achievement and their own sense of efficacy (Podell & Soodak, 1993; Tschannen-Moran & Hoy, 2001). As it has been highlighted, students' academic achievement is not the only relevant output of education systems, but also their capacity to make teachers to be satisfied with their work. In addition, as previously highlighted, the distribution of Spanish students' academic achievement is skewed to the left, something which denotes that Spanish students are poor achievers on average – as they there is a high proportion who does not get basic levels of academic achievement OECD (2016). Therefore, in the second line of research of this teachers' chapter, alternative ways of measuring education outputs, different from academic achievement, are defined and studied simultaneously to academic achievement, i.e., the percentage of students who did not reach basic learning standards and teachers' working satisfaction. Hence, these two outputs, together with students' academic achievement, are going to be analysed for the Spanish education system, in order to get a balance between them, as requiring a maximum amount for all of them – minimum in the case of students not reaching a basic level of academic achievement – would be rather unrealistic.

Thus, the hypothesis to contrast in this section is twofold: there are factors which contribute in a simultaneous way to maximize "good outputs" (average students' performance and teachers' satisfaction with their work) and to minimize "bad outputs" (percentage of students not achieving basic standards of learning) throughout the teaching-learning process in Spain. The second one is that Spanish teachers' satisfaction is a relevant output to consider as it lets to improve academic achievement in the education system. The conclusions of this analysis let to delimit in a more precise way the outputs of education systems, far from the commonly used students' academic achievement. In addition, education policies oriented to the improvement of teachers' satisfaction should be fostered if it is found that the most satisfied teachers are those who make their students get higher academic achievement or reduce the amount of students not reaching a basic level of academic achievement.

To perform this study, in a first stage the links between teachers' characteristics with their satisfaction and pupil educational outcomes were estimated by econometric methods and, as a second stage, mathematic programming techniques were applied, concretely, multiobjective

programming. Microdata from TIMSS 2011 for Spain in fourth grade were used, following the argument provided in the previous line of research in this chapter. As three outputs are going to be considered simultaneously, the mathematics subject is going to be the focus (students' achievement in mathematics, the percentage of students under a basic level of achievement in mathematics and teachers' satisfaction), so reading scores and the percentage of students under basic level of achievement in reading in PIRLS were not employed, as this would reduce the feasibility of the multiobjective problem – due to a higher bounding of the solutions' space.

I.7.3. The assessment methods

Once boys and girls present a balanced performance and teachers are effective and satisfied with their work, another important thing for an education system is how to measure students' performance, so that any negative deviation in academic achievement can be quickly and accurately identified, and hence solved. Because of that, the third chapter of this thesis is devoted to assessment methods. Concretely, the line of research in this chapter intends to analyse recently developed academic assessment procedures, as they must be objective measurements of educational achievement in educational systems (Battauz, Bellio, & Gori, 2011; Marcenaro-Gutierrez & Vignoles, 2015); otherwise, these assessment procedures may potentially condition the relationship between inputs and outputs. Paper and pencil have been the basic tools for filling out assessment tests in different subjects (Paper and Pencil Assessment, PPA), either through exams held regularly as part of the subjects' program or through standardized tests like PISA. However, with the development of new technologies, assessment tests administered by the use of computers are gaining an increasing importance nowadays. These tests may have many positive facts, as their lower potential cost in the medium term (Poggio, Glasnapp, Yang, & Poggio, 2005) – since there is no need for printed copies – and easier assessments through a more flexible questionnaire design. In addition, they can gather a greater quantity of information in less time, as well as facilitate its storage. This greater dynamism is particularly relevant in the context of international assessment programs where a large number of countries are involved. However, although computer-based assessment (CBA) presents many positive facts, this evaluation method has also risks of distorting the objective measurement of students' competencies. This would happen if the results obtained depended on the ability in using computers of the student being assessed and not on his/her true competencies in terms of the subject under assessment.

Focusing on the comparability of scores in PPA and CBA, there are no clear results when analysing this subject. Clariana and Wallace (2002) allude to the existence of a significant gap between scores in PPA and CBA – in favour of the later – in the group of students with higher academic performance, although this difference is not maintained when analysing the most disadvantaged students, as Ryan (2012) states. In addition, Bennett, Braswell, Oranje, Sandene, Kaplan, and Yan (2008) also highlighted that students obtained a lower performance

in terms of PPA compared with CBA in mathematics, and Butters and Walstad (2011) found that the highest students' academic achievement in Florida was obtained in CBA. However, in general, the existing evidence seems to opt for the superiority of the results of the written tests over the tests developed by computer (Sim & Horton, 2005; Jeong, 2014). On the other hand, many studies as Keng, McClarty, and Davis (2008) did not find any differences between these evaluation methods for reading and mathematics in the fields of item position and the competence they intended to measure, although they found differences favouring PPA in long reading passages and items that required graphic and geometric manipulation, or involved scrolling in the test. To conclude, Kingston (2009) claimed that the methodology employed in evaluating competencies does not affect the results of students in the United States.

Consequently, due to these not confluent results, this section intends to determine the effects of using CBA as assessment method instead of PPA. In the context of a digital era, PISA 2012 applied a Computer-Based Assessment for the 40% of the Spanish sample (as in other countries) in the traditional competencies of reading and mathematics[17]. Thus, this line of research makes use of the analysis of the differences between the results obtained in PPA and CBA by Spanish students for reading and mathematics competences, employing PISA 2012 database. The hypothesis to contrast is that scores in CBA are different from that obtained in PPA, due to the influence on CBA scores of students' skills using digital media. To accomplish this aim, the characteristics of students and their learning environment were examined – both at home and at school – by a multilevel modelling methodology. Concretely, those characteristics which could cause the potential differences in students' academic achievement measurement by both types of tests were analysed, and the degree to which some of those characteristics could increase or reduce the gap between them. The results of this section have a high relevance, especially in the context of PISA, as the OECD seeks to consolidate CBA as the main methodology of its assessments, starting from the PISA 2015 wave. In the event that the detected differences between the two types of assessment were significant, even though they were not large, they could have great influence in many contexts (e.g., when results of CBA are regarded as a reference level to pass a subject), so it would be necessary to know the effects of this situation in the objective measurement of the evaluated competencies. Moreover, it will not be possible to compare the results from PISA 2012 – and previous cycles – performed in PPA with those obtained in subsequent cycles done by CBA. This could suppose the raising of a new PISA shock (due to great changes in results caused only by the change in the assessment methodology), as it happened with the appearance of PISA in year 2000, which could suppose unnecessary policy reforms – and the expenses related to them.

[17] This 40% also took PPA tests.

Once the structure of the thesis has been outlined, in the following, the works in these three chapters are developed, to conclude with a summary of the obtained results and their discussion, a final conclusion and future lines of research derived from this thesis.

1. Chapter 1: Students in the education production function

1.1. Gender differences in adolescents' academic achievement

Marcenaro-Gutierrez, O. D., Lopez-Agudo, L. A., & Ropero-García, M. A. (2017). Gender differences in adolescents' academic achievement. *YOUNG*, in press.

Abstract

Recent empirical literature has highlighted that adolescents show gender differences in academic performance. The present study intends to disentangle the contribution of some less well-known factors to that gender difference in the fourth year of secondary education. To this aim, we use recent methodological advances in decomposition techniques. We observe that girls are less likely to get low scores than boys. More interestingly, gender differences in the returns to expectations about the future have been found to explain most of this advantage for girls, while boys rely more on their initial learning skills to pass. Additionally, we found that boys are more prone to misbehaviour than girls, whereas boys' academic results are more sensitive to changes in their family socio-economic status, which also explains a significant portion of the gender differences in academic achievement.

DOI: 10.1177/1103308817715163

1.2. On the potential balance among compulsory education outcomes through econometric and multiobjective programming analysis

Luque, M., Marcenaro-Gutierrez, O. D., & Lopez-Agudo, L. A. (2015). On the potential balance among compulsory education outcomes through econometric and multiobjective programming analysis. *European Journal of Operational Research*, *241*(2), 527–540.

Abstract

This paper sheds new light on the relationship between inputs and outputs in the framework of the educational production function. In particular, it is geared at gaining a better understanding of which factors may be affected in order to achieve an optimal educational output level. With this objective in mind, we analyze teacher-based assessments (actual marks) in three different subjects using a multiobjective schema. For much of the analysis we use data from a recent (2010) Survey – ESOC10, linked with the results from an educational assessment program conducted among 11 and 15 year-old students and with the administrative records on teacher-based scores. Following the statistical and econometric analysis of these data, they are used to build a multiobjective mixed integer model. A reference point approach is used to determine the profile of, potentially, the most "successful and balanced" students in terms of educational outcomes. This kind of methodology in multiobjective programming allows generating "very balanced" solutions in terms of the objective values (subjects). Finally, a sensitivity analysis is used to determine policies that can be carried out in order to improve the performance levels of primary and secondary education students. Particularly, policy makers should be more concerned with the need to promote some cultural habits – such as reading –, from both the students' and parents' side. Additionally, policy efforts should be focused on making the vocational pathways available to Spanish youth more appealing, with the aim of taking advantage of the particular skills of students not succeeding in the academic track.

DOI: 10.1016/j.ejor.2014.09.007

2. Chapter 2: Teachers in the education production function

2.1. Engaging children in lessons: the role of efficient and effective teachers

Lopez-Agudo, L. A., & Marcenaro-Gutierrez, O. D. (2017). Engaging children in lessons: the role of efficient and effective teachers. *School Effectiveness and School Improvement. An International Journal of Research, Policy and Practice, 28*(4), 650–669.

Abstract

The training of effective teachers has been one of the main aims of educational systems, in so far as it could be an important tool to improve the education performance of children and, consequently, contribute to their career progression and foster social mobility towards a more meritocratic society. The present study intends to identify these teachers by their ability to engage young students in lessons, which may reflect the capacity of their teaching practices and efficient use of available resources to increase students' learning. We focus on 4th-grade reading and mathematics teachers in Spain – using TIMSS and PIRLS 2011 data – and we propose a 2-step procedure for this analysis: the 1st step obtains teachers' efficiency scores, which are later employed in a 2nd step – together with teachers' practices – to explain children's engagement in lessons.

DOI: 10.1080/09243453.2017.1364272

2.2. Balancing teachers' math satisfaction and other indicators of the education system's performance

Marcenaro-Gutierrez, O. D., Luque, M., & Lopez-Agudo, L. A. (2016). Balancing teachers' math satisfaction and other indicators of the education system's performance. *Social Indicators Research, 129*(3), 1319–1348.

Abstract

We investigate the potential balance between some teacher characteristics, particularly teachers' satisfaction and different measures of pupils' performance (average students' test scores and percentage of students achieving basic standards of learning in mathematics), in order to optimize the outputs of the Spanish education system. Our contribution to the existing literature is twofold: on the one side, we provide estimates on the balance between teachers' effectiveness and fourth grade students' performance in Spain, by using recent survey data from the program on Trends in International Mathematics and Science Study; on the other side, we implement a novel methodology which allows optimizing simultaneously a set of indicators of the educational system outputs, to the extent that our empirical approach revealed the existence of some degree of conflict among the outputs under scrutiny. These analyses provide empirical evidence of the importance of simultaneously analysing different indicators of the performance of the education system and the need to invest in teachers' satisfaction as a motivation mechanism for improving national educational achievement, at least in primary education. Additionally they bring attention into the harmful impact of some education policy measures taken as a consequence of the economic crisis, as, e.g. the delay in the replacement of teachers, which entails the lack of teachers at schools to cover all lessons.

DOI: 10.1007/s11205-015-1162-8

3. Chapter 3: On the difficulty of accurately measuring the output of the education production function

3.1. Mind the gap: analysing the factors behind the gap in students' performance between pencil and computer based assessment methods

Marcenaro-Gutierrez, O. D., & Lopez-Agudo, L. A. (2016). Mind the gap: analysing the factors behind the gap in students' performance between pencil and computer based assessment methods. *Revista de Economía Aplicada, 24*(71), 93–120.

Abstract

The implementation of computer-based assessment (CBA) in PISA and the OECD's intention of relying only on this from 2015 suggests the need to evaluate to what extent it provides us with the same information as traditional paper and pencil assessment (PPA) conducted until 2012. Our results show that there is a significant gap between PPA and CBA for Spanish students that can be explained by factors such as availability and access to ICT or students' socioeconomic and cultural status. This leads us to conclude that educational policy reforms based on PISA scores and on any of these factors which affect them, as well as the comparison of these results along time, should be made in light of the specific assessment method used as reference.

DOI: Not available. Published online at http://www.revecap.com/revista/numeros/71/71_inv04.html

4. Results, discussion, conclusions and future research lines

4.1. Results and discussion

Once all the research works of this thesis have been presented, in the following, the obtained results are gathered and discussed. Beginning with students, it was obtained that secondary education Andalusian students show gender differences in their likelihood of failing a subject, being this higher for boys. Furthermore, differences in expectations of boys and girls may explain almost all the advantage of girls in academic achievement, being this caused in a big amount by girls' educational response – unobserved gender differences in characteristics – to expectations. Thus, when girls expect to complete a greater number of schooling years they present a higher productivity on their work and obtain better results than boys in reading and mathematics, what can explain a big amount of the gender gap in academic achievement; however, the contribution of the composition effect – observed gender differences in characteristics – for Andalusian students is lower than that obtained in other researches for the United States (e.g. Fortin, Oreopoulos, & Phipps, 2015). Other interesting result is that boys use their innate skills more than girls at school, while girls have better conduct, what allows girls to obtain better grades[18]. To conclude, boys have been found to be more sensitive to changes in their families' socio-economic status than girls.

These results suggest the need to perform policy interventions with the objective of helping those boys who do not have good results in comparison to girls, so early attention programmes can be useful to help them when they start to present this low academic achievement or misconducts. Vocational and university orientation programs are also relevant to provide students with proper information to mould their expectations, as the latter have been found to be important in explaining students' academic achievement.

However, students do not present only academic achievement differences due to their gender, but also due to differences in their learning capacities and personal characteristics. In this sense, getting all students to perform at the highest levels by implementing certain policies seems as an unrealistic aim for a government; this is because some students' characteristics may present trade-offs due to their different impacts – and sometimes, opposite – in each subjects' academic achievement. Because of that, getting the profile of the student who obtains the most balanced academic achievement in Andalusia seems to be a more interesting and reachable objective than getting that of the highest performer, so the former profile has been pursued in the current thesis. Additionally, as it was previously described, this profile may vary by gender, so it was obtained separately for boys and girls. The set of obtained educational inputs which help to get a balanced achievement in the three subjects is composed by variables in which policy intervention presents different degrees of influence. This means that there are some variables with high relevance on the academic achievement of both primary and secondary

[18] It was found that this result was not based on a better evaluation of girls by teachers.

education Andalusian students, but with little room for policy intervention as, e.g., parents' education; there are others, as, e.g., students' effort and reading habits, which present a higher capacity for policy intervention.

In this sense, it was found that increasing students' time working on their own at home helps to achieve a balanced performance, being this worktime even more productive than that spent on private tuition. Additionally, fostering reading habits by convincing parents about the importance of this matter through, e.g., advertising campaigns, could be useful to get higher student academic achievement; additionally, the intervention of teachers by informing parents on these issues could be also positive. One of the main results is that related to scholarships, as they have been found to be useful only for boys who achieve, on average, a lower performance, so fostering scholarships could motivate this particular group of students. However, the effect of this variable may not be very high, as the Andalusian education system is already publicly funded. Repeater students should also receive enough attention, as they get lower academic achievement, so policy efforts should intend to make vocational pathways more attractive for Spanish youth, with the objective of helping students who are not performing well in the academic track, so that they can make use of other of their skills in these vocational tracks.

The combination of a multiobjective approach with econometric techniques has been found to be useful to identify the desirable characteristics which define the profile of students who obtain balanced scores in reading (Spanish), mathematics and English. This analysis has shown that this balanced profile differs by gender and that – once gender has been accounted for – there is not any substantial degree of conflict in students' performance on the three subjects under analysis. Thus, the latter means that the solutions for the multiobjective problem are in line with those obtained from the estimation of the educational production function. Hence, this multiobjective programming methodology has helped to detect some issues which may be difficult to identify only with an econometric analysis – as gender differences in this balanced achievement profile. Furthermore, a sensitivity analysis has shown the robustness of the obtained solutions.

Nevertheless, even if students presented all the necessary characteristics to get balanced results in their subjects, the way that they are oriented to get them and to learn is very important. Because of that, teachers' work managing the resources that they are given by the school (in which students are included), together with their teaching practices to make students learn, are essential for the improvement of any education system. Basically, the factors which compound the latter two types of teachers' activities have been delimited, both theoretically and empirically, in efficient and effective teachers' practices – respectively. After delimiting these factors, a relevant conclusion obtained from their analysis is that efficiency is integrated in the concept of effectiveness, so requesting teachers to manage their resources in a proper way would also make their students more interested in learning.

Regarding the analysis of teachers' efficiency variables for Spanish fourth grade teachers, the proper use of the available resources has been found to be essential to engage students in reading lessons and, thus, increase teachers' effectiveness, so direct communication on this subject between head teachers and teachers should be achieved. The number of students per class should be controlled and even using a procedure of splitting classrooms into groups may be advisable, as overcrowded classrooms seem to be a problem for mathematics lessons. It was also found that a higher amount of teaching hours may be beneficial for students, until the threshold of 5 hours a week[19].

In the case of the methodologies employed by effective teachers, interactions between teachers on how to explain a concept in a lesson may be positive, while visiting other teachers' classrooms does not seem to be so. This could be indicating that teachers should organise their lessons with other teachers, but they should conduct the lessons on their own. Furthermore, effective reading teachers have been found to monitor and discuss their students' homework, while effective teachers in mathematics make students take a written quiz in mathematics every or almost every lesson and put some emphasis on national achievement tests of mathematics; hence, these practices should be included in teachers' curriculum.

However, there are other practices, as making students work on problems on their own without guiding them during the process or not paying enough attention to them, which can make students lose interest in the lesson. This could also happen with the methodology of reading aloud to the class, due to students' difficulties to read at their own pace and understand the content and sense of the text.

Overall, these effective teaching methodologies need to be incorporated to teachers' practices and curriculum so that they can improve their students' engagement, independently of students' socio-economic backgrounds, what could suppose the improvement of social mobility and education systems. Additionally, the use of the proposed teachers' classification in rankings, according to their effectiveness, can help to develop educational policies and to set teachers' salaries or compensations in a fairer way than using "raw scores" rankings.

Hence, effective teachers are desirable for every education system, but finding, attracting, hiring and keeping these teachers could be a difficult task in countries as Spain, in which teachers' status is not very high (Dolton & Marcenaro-Gutierrez, 2011). Because of that, teachers' satisfaction has been proposed as a mechanism to attract and retain in the medium and long run the best graduates into the education system. This has driven to suggest teachers' satisfaction as an output and indicator of how good an education system is, together with academic achievement and the percentage of students not reaching basic standards of learning.

[19] This has been applied in Spain with the new education act, i.e., LOMCE (BOE, 2013) – *Ley Orgánica de Mejora de la Calidad Educativa*.

An econometric analysis, combined with a multiobjective approach, has contributed to identify the profile of those teachers' who reach a balanced result in the three outcomes under analysis; this can be used to define policies oriented to make teachers reach this profile and, hence, to improve their performance. The main results of this research suggest that there are factors which contribute in a simultaneous way to maximize "good outputs" (average students' performance and teachers' satisfaction with their work) and to minimize "bad outputs" (the percentage of students not getting basic standards of learning) for fourth grade Spanish teachers, what can be translated into suggestions of intervention on these factors by teachers or, alternatively, by educational authorities. The analysis indicates that teachers should attend specialized training courses to keep improving the way they communicate their knowledge to the students and, also, to congratulate them – as a feedback – when they are performing correctly. In this sense, educational authorities have a relevant role in fostering the inclusion of these practices in the curriculum offered to future teachers by the Institutes of Education. It is also necessary to solve, before the beginning of the course, the potential problem of not having enough teachers for the scheduled class time. This issue is especially important in the Spanish education system, as authorities have delayed the replacement of teachers[20] due to the economic crisis.

Teachers' experience is also an important factor, as it was obtained that teachers' tenure has a significant and positive impact on students' academic achievement until certain threshold; then, this effect is reduced (inverted U shape). However, it does not seem to affect the ratio of students who perform under the basic level and teachers' satisfaction. In the case of Spain, to the extent that the variations in teachers' salary depend – mainly – on their years of experience, the decreasing marginal effect of teachers' experience could be reflecting few incentives to make teaching efforts, resulting from a nearly flat career progression of teaching. As this factor is clearly under control of policy makers, this situation may be improved by creating career progression routes, based on a balanced target for different education outputs.

Furthermore, policy makers should be more concerned with the promotion of reading habits and raise parents' awareness about the importance of reading via, e.g., advertising campaigns. Moreover, teachers play an important role in the promotion of reading habits among students and also by interacting with their parents on this issue. In the case of students with diglossia, policy makers should put greater effort on the integration of pupils from different cultural backgrounds, what will additionally contribute to social cohesion. In addition, the feeling of safety of teachers about their personal integrity and their job conditions has a very

[20] Although education authorities do not provide figures on this, many Spanish newspapers have published news commenting this problem and the strikes it has caused (see, e.g. http://www.diariodesevilla.es/article/sevilla/1495842/la/sustitucion/profesor/tarda/cinco/veces/mas/por/los/recortes.html, or http://ccaa.elpais.com/ccaa/2014/03/27/catalunya/1395933405_126552.html).

positive influence on the achievement of the pursued objectives. Consequently, creating a comfortable working environment should be a key principle in every school, what can be managed by education authorities and head teachers.

Moreover, it is important to highlight that the solutions of the multiobjective approach show that teacher satisfaction is the only indicator whose optimal value exceeds the sample maximum, indicating that more effort has to be placed on it and, at the same time, there is a great possibility for intervention. In the view of these results, it can be stated that improving teacher satisfaction may be a promising area for future research, to identify how other countries – different from Spain – have obtained highly satisfied teachers who can enhance their students' academic achievement.

In order to assess the way that these teachers are performing their work and the level of academic achievement that students are getting thanks to their lessons, objective and precise assessment methods are needed. Due to the new technological advances, assessment methods have evolved into a digital format, so it is necessary to find out if these new digital methodologies are measuring the skills they are intended to. To do this, the differences between the academic achievement presented by students in paper and pencil assessment (PPA) and computer-based assessment (CBA) PISA 2012 tests have been studied, for Spanish students.

Literature on this issue is not only limited but also far from conclusive, so this research has deepened into this subject by examining which students' characteristics and those of their learning environment may increase or reduce the gap between PPA and CBA. Results have shown that conducting the tests in a CBA format "favours" girls over boys, in the sense that it helps girls to improve their relative position in the mathematics competence, in which girls have shown a lower average performance than boys when evaluated by PPA. Furthermore, the advantage which girls already had in the reading competence remained unchanged or even slightly increased in CBA.

It is also interesting to highlight that turning from PPA to CBA would reduce the distance of students with high socio-economic background with respect to the students found in "more disadvantaged" environments, which denotes that differences between PPA and CBA may go beyond the simple assessment mode effect. This fact is also present when observing that the highest gap between CBA and PPA – in favour of CBA – appears in those students who spend more hours studying with the help of a computer, implying the existence of a selection bias among the "less gifted", who are those getting the greatest relative performance in CBA. Because of that, the availability and use of ICT in the school and at home to explain these gaps have obtained an important role.

To conclude, the results reported in this investigation have shown many interesting facts: using CBA could cause that the "less gifted" students "artificially" reduce the relative disadvantage they present when performing PPA-type tests, although this does not mean an

actual better performance in the teaching-learning process. For this reason, education policy initiatives which take the PISA assessment framework as a best-practice model should be carefully considered, as the effect of interventions on some groups of students might be conditioned by the employed assessment method. This is a relevant issue to the extent that CBA will be used to evaluate the performance of educational systems from PISA 2015 onwards, so it is necessary to be aware of the influence of certain factors which may potentially cause differences between PPA and CBA results. Hence, the comparison of the education performance of a country based on PISA, before and after the year 2015, needs to be made with caution, as the size of the gap significantly correlates with ICT or "non-traditional" factors, whose contribution to the education production function may evolve much faster in this digital era than "traditional factors". Thus, policy makers should be informed about these differences, as education policy and budgetary decisions affecting these factors could have an unexpected effect on students' performance in CBA, compared to PPA.

4.2. Final conclusions

This thesis has attempted to analyse the education reality of Spanish and Andalusian education systems, focusing on certain aspects of the three elements of the teaching-learning process under analysis: students, teachers and assessment methods. To do this, several econometric and multiobjective methodologies – and a combination of them – have been employed. In this sense, all the works which compose this thesis present some degree of novelty and interest in the analysis of the education systems under study, filling gaps in their literature.

To accomplish these objectives, this thesis has studied the education production function of Andalusia and Spain. It has focused on the idea that an education system is not a singular element which works well or bad in general, but a combination of many elements which have to work well in their own and, also, when integrated in the whole system, creating synergies between them to reach the ultimate objective of students' learning. Based on this idea, this thesis has departed from the students, highlighting the importance of considering gender differences in the teaching-learning process when teaching students and, also, showing the relevance of getting a balanced academic achievement from all students, and not only a very high one from some of them. However, these important resources that every education system has, i.e. students, need someone who can lead them through the learning path: the teachers. It has been found that teachers not only have to be effective – and, hence, efficient – in making their students learn, but they also have to be satisfied with their work. The latter is because, to the extent that an education system has satisfied teachers, it will attract more effective teachers, keeping a high level of academic achievement in the whole education system. Finally, although making students learn is essential, their progression has to be measured and controlled in some way that potential problems or difficulties can be discovered and solved on time. Because of that, the design of precise assessment methods has been studied.

It is necessary to highlight that, although many policy suggestions have been derived from this thesis, their application requires an active implication and investment by the government and other educational participants – families, teachers, etc. – to make them succeeded. Thus, these agents should manage their resources and trust in the researchers who perform this kind of analysis to improve their education systems; otherwise, any effort performed in this field of Economics of Education will be in vain. On the researchers' side, the use of representative empirical data for the population under study reinforces the confidence that these agents could have in the obtained results; in addition, researchers' adequate understanding of the data under study is essential to get valid conclusions.

Another relevant implication of this thesis it is that the conclusions obtained for the country and region under analysis cannot be directly extended to other countries, as each country presents its own education and cultural casuistries. However, the methodologies and procedures employed in the current thesis could be used as a first approach to deal with the same issues of this thesis, but for other countries, even allowing to perform cross-country comparisons.

To conclude, the worldwide education reality is so huge nowadays that it is impossible to cover every subject in only one thesis. Because of that, in the next and last section, future lines of research following after this thesis are described.

4.3. Future research lines

This thesis has set a first approach to analyse the differences in boys' and girls' academic achievement, which have been found to be highly affected by gender differences in educational expectations. Nevertheless, students form their expectations not only based on the feedback that they receive from the education system (e.g., scores), but also based on their parents' expectations; the latter also form their expectations having into account their children's expectations, configuring a simultaneous process (Hao & Bonstead-Bruns, 1998). Hence, dealing with educational expectations not only requires carefully accounting for their endogeneity when employed to explain students' academic achievement but, also, taking care of the simultaneity in the determination of students' and parental expectations, in the case that both are used in the analysis. Furthermore, not only interesting is analysing the effect of expectations on students' academic achievement, but also how these expectations – and their effect on students' academic achievement – differ according to the sex of the student and the parent; all these issues on expectations will be the focus of future research.

Continuing with the subject of gender differences in education, the current thesis has analysed this issue focusing on individual characteristics of boys and girls. However, certain gender roles and cultural characteristics that countries' societies present have also been found to play a relevant role on the definition of these gender differences (OECD, 2009). Hence, accounting for the latter factors is especially important when performing an international study

of gender differences in academic achievement, as it is necessary to consider the existence of these cultural dissimilarities between countries. This is because there are many aspects and perceptions of the society which may differ by gender and could have a heterogeneous influence on students' academic achievement as, e.g., gender roles or stereotypes. Future works in this line of research will deal with these and other cultural characteristics which may be affecting these gender differences.

Although gender differences in learning are important determinants of the heterogeneous academic achievement presented by students, there are additional individual characteristics which can be causing differences in performance. In this sense, there are some collectives who present certain learning difficulties as, e.g., immigrant students or those with some kind of learning disability (Calero & Escardíbul, 2013; Kavale & Forness, 2000). The works of this thesis have only controlled by immigration or diglossia and did not include students with learning difficulties[21]. These are relevant fields to work as future research lines for the case of Spain, as immigration has increased in recent years (INE, 2017) and the high dropout rates which have been previously highlighted in the Introduction (IECA, 2017a) may be indicating that many students present learning problems.

Additionally, this thesis has highlighted the low academic achievement that Spanish and Andalusian students present. As a way of "helping" those students who did not reach the minimum level of academic achievement required for their grade, Spanish authorities have adopted an education model in which students have to repeat that grade. However, grade repetition is a very controversial topic in Spain, as many studies have shown that repetition is not an adequate practice to improve students' academic achievement, being even harmful for their learning (Agasisti & Cordero-Ferrera, 2013; García-Pérez, Hidalgo-Hidalgo, & Robles-Zurita, 2014). Although relevant, the study of the effect of repetition on students' academic achievement means dealing with endogeneity problems, what makes difficult to get reliable results about its actual effect on students' academic achievement. Because of that, treating this endogeneity problem to get the actual effect of repetition in Spain will be one of the future lines of research derived from this thesis.

Moving from students to the teachers in the education system, this thesis has highlighted that one of the most important resources that teachers have to make students learn is class time. In this sense, the time devoted by students to their learning activities as, e.g., to attend their lessons at school or do their homework in their own or helped by others (parents, teachers, etc.), plays a relevant role in the way that students learn (Dolton, Marcenaro-Gutierrez & Navarro, 2003; Gromada & Shewbridge, 2016). In relation to this class time, there are many issues which can be analysed: the time that students spend in the classroom, the time they

[21] These collectives were not the focus of the research, so their particularities have not been studied in depth.

devote to do homework or private tuition time (as e.g. private lessons, academies, etc.). In addition, the effect of learning time on academic achievement does not only depend on the use that students and teachers make of it, but also on the methodologies employed to approach this learning. In these terms, the increasing use of digital media (TV, videogames, etc.) and ICTs (smart phones, computers, digital blackboards, etc.) that students are making may affect their academic achievement, in the way that they develop their learning skills and the use that they do of their time. Furthermore, the use that teachers make of these resources to incorporate them to their teaching practices is also relevant, so the effect of these methodologies in learning will be analysed.

As it has been highlighted in this thesis, methodology is an important support for every research which intends to deal with education issues. However, although there are many approaches to study the education reality, each one is limited to a particular kind of analysis or configuration/content of the data under analysis. In this thesis, several combinations of techniques have been proposed (econometrics and multiobjective programming, stochastic frontier and multilevel analysis) to answer the proposed questions. As there is always room for improvement, one possible extension of the econometrics and multiobjective programming combination of techniques can be related to the limitations of using average-point coefficients from the econometric estimations. In order to handle the uncertainty presented by these average-point coefficients, a variant of the econometric estimations, based on the confidence intervals of their coefficients, will be studied, so that these intervals can be employed in the optimization phase by the use of multiobjective interval programming models. Furthermore, other works will be focused on developing new combinations of techniques by mixing efficiency procedures as data envelopment analysis (DEA) and multiobjective programming techniques; this approach can be useful, e.g., to provide another insight into the analysis of teachers' efficiency. Additionally, the development of new techniques makes appear a huge lack of literature on the theoretical, methodological and empirical implications of their use; in this sense, this will also be developed in future works.

The current thesis has focused on compulsory education. Nevertheless, the effects of the outcomes obtained by students in the teaching learning-process extend from the very beginning of children's education to their adult lives. In the section devoted to gender differences in academic achievement it was highlighted how expectations were not so endogenous in Spain as in other countries (e.g., the United States) because staying at the University in Spain supposed a low cost for students, as they do not have to pay most of the expense of these studies, so they could stay longer in the system until getting their degree. This is an issue which can make students who are not sufficiently skilled to access and complete these studies; a problem which might be worsened when having into account that, since the academic year 2010/11, the Bologna process has been employed at Spanish Universities. This has supposed many changes

in the way that classes are conducted, the works that students have to perform to pass a subject and the competences included in the curriculum. Enough time has passed to get information about the first graduates of this new study plan, what could help to analyse in future research the effectiveness that this Bologna process had in the learning of these students.

To conclude, these researches can be also extended to other developing or least developed countries, in which the particularities of education systems can be completely different to those of Spain.

5. References

Afshartous, D., & Wolf, M. (2007). Avoiding 'data snooping' in multilevel and mixed effects models. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 170*(4), 1035–1059. doi: 10.1111/j.1467-985X.2007.00494.x

Agasisti, T., & Cordero-Ferrera, J. M. (2013). Educational disparities across regions: A multilevel analysis for Italy and Spain. *Journal of Policy Modeling, 35*(6), 1079–1102. doi: 10.1016/j.jpolmod.2013.07.002

Aghion, P., & Howitt, P. (1998). *Endogenous Growth Theory*. Cambridge, MA: MIT Press.

Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education*, *17*(2), 86−95. doi: 10.1177/088840649401700203

Avvisati, F., & Keslair, F. (2014). *REPEST: Stata module to run estimations with weighted replicate samples and plausible values*. Statistical Software Components S457918, Boston College Department of Economics. Retrieved from: https://ideas.repec.org/c/boc/bocode/s457918.html

Basu, A. M. (2002). Why Does Education Lead to Low Fertility? A critical review of some of the Possibilities. *World Development, 30*(10), 1779–1790. doi: 10.1016/S0305-750X(02)00072-4

Battauz, M., Bellio, R., & Gori, E. (2011). Covariate error adjustment for multilevel models with application to educational data. *Journal of Educational and Behavioral Statistics*, *36*(3), 283–306. doi: 10.3102/1076998610366262

Battese, G. E., & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics, 38*(3), 387–399. doi: 10.1016/0304-4076(88)90053-X

Bennett, R. E., Braswell, J., Oranje, A., Sandene, B., Kaplan, B., & Yan, F. (2008). Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP. *Journal of Technology, Learning, and Assessment*, *6*(9), 1–39.

BOE (2013). *Organic Law 8/2013, 9th December, for the improvement of the education quality (LOMCE)*. Nº 295, 10th December 2013, 97858–97921. Spain.

Bowles, S. (1970). Towards an educational production function. In W. Lee Hansen (Ed.), *Education, income, and human capital* (pp. 11–61). New York: Columbia University Press for the National Bureau of Economic Research.

Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review, 41*, 71–88. doi: 10.1016/j.econedurev.2014.04.002

Butters, R. B., & Walstad, W. B. (2011). Computer Versus Paper Testing in Precollege Economics. *The Journal of Economic Education, 42*(4), 366–374. doi: 10.1080/00220485.2011.606087

Calero, J., & Escardíbul, O. (2013). El rendimiento del alumnado de origen inmigrante en PISA-2012. In INEE (Ed.), *PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Análisis secundario* (pp. 32–66). Madrid: Ministerio de Educación, Cultura y Deporte, Instituto Nacional de Evaluación Educativa.

Caro, D. (2017). *Package 'intsvy': International Assessment Data Manager*. Last time accessed September 2017 from https://cran.r-project.org/web/packages/intsvy/intsvy.pdf

Cattaneo, M. A., Oggenfuss, C., & Wolter, S. C. (2017). The More, the Better? The Impact of Instructional Time on Student Performance. *Education Economics, 25*(5), 433–445. doi: 10.1080/09645292.2017.1315055

Clariana, P., & Wallace, R. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology, 33*(5), 593–602. doi: 10.1111/1467-8535.00294

Cochran-Smith, M., & Lytle, S. L. (1999). The Teacher Research Movement: A Decade later. *Educational Researcher, 28*(7), 15–25. doi: 10.3102/0013189X028007015

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Cousins, J., & Walker, C. (1995). *Personal teacher efficacy as a predictor of teachers' attitudes toward applied educational research*. Paper presented at the annual meeting of the Canadian Association for the Study of Educational Administration, Montreal.

Creemers, B. P. M., & Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 7*(3), 197–228. doi: 10.1080/0924345960070301

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934–945. doi: 10.1037/0012-1649.33.6.934

Dolton, P., Makepeace, G., & Marcenaro-Gutierrez, O. (2005). Career progression: Getting–on, getting–by and going nowhere. *Education economics, 13*(2), 237–255. doi: 10.1080/09645290500031447

Dolton, P., & Marcenaro-Gutierrez, O. (2011). If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance. *Economic Policy, 26*(65), 5–55. doi: 10.1111/j.1468-0327.2010.00257.x

Dolton, P, Marcenaro-Gutierrez, O., & Navarro, L. (2003). The effective use of student time: a stochastic frontier production function case study. *Economics of Education Review, 22*(6), 547–560. doi: 10.1016/S0272-7757(03)00027-X

EPA (2017). Encuesta de Población Activa. Last time accessed September 2017 from http://www.ine.es/dynt3/inebase/es/index.htm?padre=982&capsel=986 ; Last time accessed

September 2017 from http://www.juntadeandalucia.es/institutodeestadisticaycartografia/iea/resultadosConsulta.jsp?CodOper=25&codConsulta=72916#consulta_72916

Eurostat (2017). Early leavers from education and training. Last time accessed September 2017 from http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=t2020_40&plugin=1

Fortin, N. M., Oreopoulos, P., & Phipps, S. (2015). Leaving boys behind. Gender disparities in high academic achievement. *Journal of Human Resources, 50*(3), 549–579. doi: 10.3368/jhr.50.3.549

Francis, B., & Skelton, C. (2005). *Reassessing gender and achievement: Questioning contemporary key debates*. London: Routledge.

Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus and Giroux.

Fundación Foessa (2016). *Expulsión Social y Recuperación Económica*. Madrid: Cáritas Española Editores.

García-Pérez, J. I., Hidalgo-Hidalgo, M., & Robles-Zurita, J. A. (2014). Does grade retention affect achievement? Some evidence from Spain. *Applied Economics, 46(12),* 1373–1392.

Goldhaber, D. D., Brewer, D. J., & Anderson, D. J. (1999). A three-way error components analysis of educational productivity. *Education Economics, 7*(3), 199–208. doi: 10.1080/09645299900000018

Goldhaber, D., Gross, B., & Player, D. (2010). Teacher Career Paths, Teacher Quality, and Persistence in the Classroom: Are Public Schools Keeping Their Best? *Journal of Policy Analysis and Management, 30*(1), 57–87. doi: 10.1002/pam.20549

Grilli, L., & Rampichini, C. (2009). Multilevel models for the evaluation of educational institutions: a review. In P. Monari, M. Bini, D. Piccolo, & L. Salmaso (Eds.), *Statistical Methods for the Evaluation of Educational Services and Quality of Products* (pp. 61–80). Heidelberg: Physica-Verlag. doi: 10.1007/978-3-7908-2385-1_5

Gromada, A., & Shewbridge, C. (2016). *Student Learning Time. A Literature Review*. OECD Education Working Papers, No. 127. Paris: OECD Publishing. doi: 10.1787/5jm409kqqkjh-en

Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources, 14*(3), 351–388. doi: 10.2307/145575

Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2014). *The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance*. CESifo Working Paper, No. 5120.

Hanushek, E. A., & Woessmann (2007). *The Role of Education Quality in Economic Growth*. World Bank Policy Research Working Paper No. 4122. doi: 10.1596/1813-9450-4122

Hao, L., & Bonstead-Bruns, M. (1998). Parent-Child Differences in Educational Expectations and the Academic Achievement of Immigrant and Native Students. *Sociology of Education, 71*(3), 175–198. doi: 10.2307/2673201

IECA. (2010). Encuesta Social 2010. Educación y familia. http://www.juntadeandalucia.es/institutodeestadisticaycartografia/encsocial/2010/index.htm

IECA. (2017a). Premature dropout rate by sex. Last time accessed September 2017 from http://www.juntadeandalucia.es/institutodeestadisticaycartografia/indsoc/indicadores/1038.htm

IECA. (2017b). Unemployment rate in Andalusia. Last time accessed September 2017 from http://www.juntadeandalucia.es/institutodeestadisticaycartografia/iea/detalleDatosDia.jsp?cod=190&ram=D

INE. (2015). Producto Interior Bruto regional. Año 2014. http://www.ine.es/prensa/np901.pdf

INE. (2017). Cifras de Población a 1 de enero de 2017. Estadística de Migraciones 2016. http://www.ine.es/prensa/cp_2017_p.pdf

Ivie (2013). *Cambio educativo y productivo en España. 1964–2013. Capital Humano, 144*. Valencia: Publicaciones Ivie.

Jeong, H. (2014). A comparative study of scores on computer-based tests and paper-based tests. *Behaviour & Information Technology*, *33*(4), 410–422. doi: 10.1080/0144929X.2012.710647

Jürges, H., & Schneider, K. (2007). Fair ranking of teachers. *Empirical Economics, 32*(2-3), 411–431. doi: 10.1007/s00181-006-0112-3

Kavale, K. A., & Forness, S. R. (2000). What Definitions of Learning Disability Say and Don't Say. A Critical Analysis. *Journal of Learning Disabilities, 33*(3), 239–256. doi: 10.1177/002221940003300303

Keng, L., McClarty, K. L., & Davis, L. L. (2008). Item-level comparative analysis of online and paper administrations of the Texas assessment of knowledge and skill. *Applied Measurement in Education, 21*(3), 207–226. doi: 10.1080/08957340802161774

Kingston, N. M. (2009). Comparability of computer- and paper- administered multiple choice test for K–12 populations: A synthesis. *Applied Measurement in Education, 22*(1), 22–37. doi: 10.1080/08957340802558326

Klassen, S. (2002). Low schooling for Girls, slower Growth for All? Cross country evidence on the Effect of Gender Inequality in Education on Economic Development. *World Bank Economic Review, 16*(3), 345–373. doi: 10.1093/wber/lhf004

Klassen, R. M., & Chiu, M. M. (2010). Effects on Teachers' Self-Efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress. *Journal of Educational Psychology, 102*(3), 741–756. doi: 10.1037/a0019237

Kumar, P. (2014). A study job satisfaction of British Education Trained teacher working at primary schools in Saharanpur district. *Journal of Indian Research, 2*, 141–145.

Kyriakides, L. (2008). Testing the validity of the comprehensive model of educational effectiveness: a step towards the development of a dynamic model of effectiveness. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 19*(4), 429–446. doi: 10.1080/09243450802535208

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review, 99*(5), 1979–2011. doi: 10.1257/aer.99.5.1979

Leckie, G., & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society: Series A (Statistics and Society), 172*(4), 835–851. doi: 10.1111/j.1467-985X.2009.00597.x

Lipman, P. (2003). *High Stakes Education: Inequality, Globalization, and Urban School Reform (Critical Social Thought)*. London: Routledge.

Lockheed, M. E., & Hanushek, E. A. (1994). Concepts of Educational Efficiency and Effectiveness. In T. Husén, & T. N. Postlethwaite (Eds.), *International Encyclopedia of Education* (pp. 1779–1784). Oxford: Pergamon.

Loeb, S., & Page, M. E. (2000). Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecuniary variation. The *Review of Economics and Statistics, 82*(3), 393–408. doi: 10.1162/003465300558894

Luque, M., Miettinen, K., Eskelinen, P., & Ruiz, F. (2009). Incorporating Preference Information in Interactive Reference Point Methods for Multiobjective Optimization. *OMEGA-International Journal of Management Science, 37*(2), 450–462. doi: 10.1016/j.omega.2007.06.001

Luque, M., Ruiz, F., & Steuer, R. E. (2010). Modified Interactive Chebysshev Algorithm (MICA) for Convex Multiobjective Programming. *European Journal of Operational Research, 204*(3), 557–564. doi: 10.1016/j.ejor.2009.11.011

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics, 107*(2), 407–437. doi: 10.2307/2118477

Marcenaro-Gutierrez, O., Luque, M., & Ruiz, F. (2010). An application of multiobjective programming to the study of workers' satisfaction in the Spanish labour market. *European Journal of Operational Research, 203*(2), 430–443. doi: 10.1016/j.ejor.2009.07.017

Marcenaro-Gutierrez, O., & Vignoles, A. (2015). A comparison of teacher and test-based assessment for Spanish primary and secondary students. *Educational Research, 57*(1), 1–21. doi: 10.1080/00131881.2014.983720

Marques, H. (2015). Does the Gender of Top Managers and Owners Matter for Firm Exports? *Feminist Economics, 21*(4), 89–117. doi: 10.1080/13545701.2015.1029958

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Science*. Chesnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Miettinen, K. (1999). *Nonlinear Multiobjective Optimization*. Boston: Kluwer Academic Publishers.

Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art - teacher effectiveness and professional learning. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, *25*(2), 231–256. doi: 10.1080/09243453.2014.885451

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker K., T. (2012). *PIRLS 2011 International Results in Reading*. Chesnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Chesnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

OECD. (2009). *Equally prepared for life? How 15-year-old boys and girls perform in school*. OECD Publishing. doi: 10.1787/9789264064072-en

OECD. (2010a). *PISA 2009 Assessment Framework. Key competencies in reading, mathematics and science*. OECD Publishing. doi: 10.1787/9789264062658-en

OECD. (2010b). *PISA 2009 at a Glance*. OECD Publishing. doi: 10.1787/9789264095298-en

OECD. (2014a). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014).* PISA, OECD Publishing. doi: 10.1787/9789264208780-en

OECD. (2014b). *Are boys and girls equally prepared for life?* OECD Publishing.

OECD. (2014c). *Education at a Glance 2013: OECD Indicators*. Centre for Educational Research and Innovation. doi: 10.1787/199

OECD. (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. PISA, OECD Publishing. doi: 10.1787/9789264229945-en

OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris: PISA, OECD Publishing. doi: 10.1787/9789264266490-en

Podell, D. M., & Soodak, L. C. (1993). Teacher efficacy and bias in special education referrals. *Journal of Educational Research, 86*(4), 247−253. doi: 10.1080/00220671.1993.9941836

Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A comparative evaluation of score results from computerized and paper & pencil mathematics testing in a large scale state assessment program. *The Journal of Technology, Learning, and Assessment, 3*(6), 1–30.

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education Economics, 12*(2), 111–134. doi: 10.1080/0964529042000239140

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458. doi: 10.1111/j.1468-0262.2005.00584.x

Rivkin, S. G., & Schiman, J. C. (2015). Instruction Time, Classroom Quality, and Academic Achievement. *The Economic Journal, 125*, F425–F448. doi: 10.1111/ecoj.12315

Ryan, B. A. (2012). *Computer-Based Versus Paper-Pencil Modes of Administration United States Government End of Course Exams: Student Cumulative Grade Point Averages as Predictors of Success*. Lindenwood, PhD Dissertation. Lindenwood University.

Sim, G., & Horton, M. (2005). Performance and attitude of children in computer based versus paper based testing. In P. Kommmers & G. Richards (Eds.), *Proceedings of ED-MEDIA 2005--World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005* (pp. 3610–3614). Montreal, Canada: Association for the Advancement of Computing in Education (AACE).

Smith, M. C., Mikulecky, L., Kibby, M. W., Dreher, M. J., & Dole, J. A. (2000). What will be the demands of literacy in the workplace in the next millennium? *Reading Research Quarterly, 35*(3), 378–383. doi: 10.1598/RRQ.35.3.3

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics, 39*(3), 312–320. doi: 10.2307/1926047

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. doi: 10.2307/1882010

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current directions in Psychological Science, 18*(5), 259–263. doi: 10.1111/j.1467-8721.2009.01648.x

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive concept. *Teaching and Teacher Education, 17*(7), 783−805. doi: 10.1016/S0742-051X(01)00036-1

Vignoles, A., De Coulon, A., & Marcenaro-Gutierrez, O. (2011). The value of basic skills in the British labour market. *Oxford Economic Papers, 63*(1), 27–48. doi: 10.1093/oep/gpq012

Wallenius, H., Wallenius, J., & Vartia, P. (1978). An approach to solving multiple criteria macroeconomic policy problems and an application. *Management Science, 24*(10), 1021–1030. doi: 10.1287/mnsc.24.10.1021

6. Summary of the thesis in Spanish (resumen de la tesis en español)

Esta tesis pertenece al campo de Economía de la Educación, área de investigación dentro de la Economía que se basa en la aplicación de técnicas econométricas para analizar la realidad educativa. La intención de esta tesis es evaluar diferentes aspectos del proceso de enseñanza-aprendizaje en la educación primaria y secundaria del sistema educativo español, con especial énfasis en la región de Andalucía. Por lo tanto, esta tesis pretende ofrecer propuestas sobre el diseño de políticas para estos sistemas educativos, para lograr los siguientes objetivos: superar las diferencias de género existentes en rendimiento académico, ayudar a los estudiantes a lograr un rendimiento académico equilibrado, proveer a los sistemas educativos con profesorado eficaz y satisfecho – que pueda mejorar el conocimiento de los estudiantes y su interés – y mejorar el conocimiento de los agentes educativos e investigadores sobre las implicaciones de la utilización de los datos obtenidos mediante nuevos métodos de evaluación.

Esta tesis se ha centrado en la educación por su importancia en la sociedad: tal y como han destacado Hanushek y Woessmann (2007)[22], ésta desempeña un papel esencial en el bienestar de la población, lo que sitúa la calidad de la educación en una posición relevante en el debate político. Esta calidad de la educación está tan estrechamente relacionada con el desarrollo de las habilidades cognitivas de la población que autores como Hanushek y Woessmann (2007) las usaron como su *proxy*. En la medida en que estas habilidades afecten a los salarios y, por lo tanto, el bienestar de la población, tendrán una influencia relevante en el crecimiento económico (Klassen, 2002; OCDE, 2015).

Sin embargo, Hanushek y Woessmann (2007) destacaron que, al analizar la calidad de la educación, es esencial tener en cuenta características particulares del país analizado, ya que ignorarlas puede suponer llegar a conclusiones erróneas sobre la relación entre educación y resultados económicos. Este problema puede ser consecuencia de las diferencias en términos de los retornos de la educación entre países. Debido a esto, esta tesis se centra en los contextos económicos y educativos del país – España – y región – Andalucía – en los que se realizan los análisis.

Tal y como se ha indicado, resulta obvio que la educación es esencial en el desarrollo socio-cultural y económico de un país. Sin embargo, no fue hasta la aparición a mediados de los años noventa de los informes[23] PISA[24] y TIMSS[25]/PIRLS[26], realizados por organizaciones como la OCDE[27] o la IEA[28], respectivamente, que los países participantes – y aquellos dispuestos a

[22] Estos autores han hecho una gran contribución a la literatura en Economía de la Educación.

[23] Este tipo de estudios que están orientados a evaluar las competencias de los estudiantes y el rendimiento académico con una orientación internacional son conocidos como *International Large Scale Assessments* (ILSA) tests.

[24] *Programme for International Student Assessment*.

[25] *Trends in International Mathematics and Science Study*.

[26] *Progress in International Reading Literacy Study*.

[27] *Organization for Economic Co-operation and Development*.

participar – desarrollaron un gran espíritu competitivo en términos de rendimiento en estos informes. Cada ciclo de éstas y otras pruebas ILSA ha evaluado principalmente las competencias de los estudiantes en materias como lectura, matemáticas y ciencias – entre otras como, por ejemplo, las competencias de resolución de problemas o la financiera – en diferentes niveles de educación obligatoria.

Debido a la relevancia que estas competencias tienen en el desarrollo de las habilidades cognitivas de la población, esta tesis abordará el estudio de las tres[29].

Centrándonos en el país que es el foco de esta tesis – España – los ciclos más recientes de las pruebas ILSA han mostrado una realidad alarmante, lo que ha motivado la elección de este país para su estudio: los estudiantes españoles obtienen un bajo rendimiento en estas tres competencias, en comparación con los estudiantes de los países con mejor rendimiento. Teniendo en cuenta que las puntuaciones de PISA – para alumnado de décimo grado – tienen una media de 500 y una desviación estándar de 100, en el caso de PISA 2015 (OCDE, 2016) España presenta 486 puntos en matemáticas, superada por los 564 puntos (una diferencia de 0,78 desviaciones estándar) de Singapur y 548 (0,62 desviaciones estándar) de Hong Kong (China), siendo el promedio de la OCDE 490 (0,04 desviaciones estándar por encima); España presenta 496 puntos en lectura, 0,39 desviaciones estándar por debajo de Singapur y 0,31 por debajo de Hong Kong (China), siendo la media de la OCDE de 0,03 desviaciones estándar por debajo; y 493 en ciencias, siendo las puntuaciones de Singapur 0,63 desviaciones estándar más altas y 0,45 desviaciones estándar más altas en el caso de Japón, mientras que la media de la OCDE es la misma que en España.

Además, según OCDE (2016), los estudiantes españoles, además de tener un rendimiento inferior al de otros países, tienen una distribución de rendimiento desplazada hacia la izquierda, es decir, los estudiantes españoles tienen bajo rendimiento en promedio.

Por otra parte, el abandono educativo temprano de los estudiantes españoles presentaba una tasa del 20% (24% para niños y 15,8% para niñas) en 2015 (IECA, 2017a), superior a la del resto de la UE-28 en 2015 (11%, un 12,4% para los niños y un 9,2% para las niñas; Eurostat, 2017). Estas cifras han comenzado a disminuir en este contexto económico de crisis que estamos viviendo hoy en día, ya que esta crisis ha reducido drásticamente las oportunidades de empleo (según EPA, 2017, la tasa de desempleo pasó de 13,79% en el cuarto trimestre de 2008 a 20,90% en el cuarto trimestre de 2015). Por lo tanto, estas altas tasas de desempleo han reducido al mismo tiempo el coste de oportunidad de permanecer en el sistema educativo, lo que

[28] *International Association for the Evaluation of Educational Achievement*.

[29] Debido a la gran correlación que la competencia de ciencias tiene con matemáticas (Rivkin & Schiman, 2015) la primera competencia tiene un papel secundario en esta tesis, por lo que será estudiada en la última sección del último capítulo.

significa que la población española está tratando de resolver su falta de cualificación aumentando su nivel educativo, para facilitar su incorporación al mercado de trabajo.

Por lo tanto, todas estas cifras parecen indicar que el sistema educativo español necesita una exploración profunda para determinar qué factores pueden estar conduciendo a estos pobres resultados; una vez identificados estos factores, las autoridades educativas podrán implementar políticas para mejorar la situación española. Sin embargo, dentro de España hay una región – Comunidad Autónoma – en la que esta tesis pone especial interés, ya que su situación es aún peor que la de España en general; ésta es, concretamente, Andalucía. En este sentido, hay muchas razones que hacen necesario el estudio de esta región: ha obtenido puntuaciones por debajo del promedio de España en las tres competencias evaluadas por PISA 2015 – lectura, matemáticas y ciencias – (OCDE, 2016), las tasas de abandono educativo temprano de Andalucía son incluso más altas que las de España, con cifras del 28,9% y del 20,8% para los chicos y las chicas en 2015, respectivamente (una tasa de abandono total del 24,9%, superando las tasas españolas de abandono en 4,9%; IECA, 2017a) y sus tasas de desempleo son también mucho más altas que las españolas – un 26,94% para Andalucía, 24,22% para hombres y 30,25% para mujeres, frente a un total de 18,75% para España, 17,22% para hombres y 20,51% para mujeres, para el primer trimestre de 2017 (IECA, 2017b). Estas cifras son aún más alarmantes cuando se tiene en cuenta que Andalucía es la región más poblada de España, lo que significa que esta pobre situación afectará a muchos estudiantes.

Otra razón para estudiar el caso del sistema educativo andaluz en esta tesis es la novedad que presenta la disponibilidad de una rica base de datos para esta región. Concretamente, se trata de la Encuesta Social 2010: Educación y Vivienda – ESOC10 – (IECA, 2010), que se centró en Andalucía y se llevó a cabo entre estudiantes de 11-12 y 15-16 años y sus familias. Esta ESOC10 se vinculó a los resultados de los registros administrativos (SENECA) de puntuaciones del profesorado, aportados por la Consejería de Educación de la Junta de Andalucía. A esta base de datos combinada se le va a denominar a partir de ahora como ESOC10-SEN, convirtiéndose en un elemento esencial para profundizar en el estudio del proceso de enseñanza-aprendizaje en Andalucía.

Las evidencias y cifras presentadas han señalado la necesidad de mejorar la situación desfavorable de los sistemas educativos español y andaluz. Por lo tanto, esta tesis tiene como objetivo general el análisis de las características educativas de este país y región, para contribuir al diseño óptimo de las políticas educativas, dirigidas a mejorar sus resultados académicos. Para afrontar este ambicioso y exigente objetivo, esta tesis trabaja con una herramienta conocida como "función de producción educativa", por lo que cada metodología y enfoque de esta tesis se centra en el estudio de aspectos específicos de esta función, desde diferentes perspectivas. Pero, ¿qué significa esta función de producción educativa? En general, el concepto de función de producción ha sido ampliamente utilizado en Economía en los últimos años (Solow, 1957) y se

entiende como una forma de relacionar la cantidad de *outputs* producidos por unidad de tiempo con la cantidad de *inputs* utilizados para producir esos *outputs*. En el campo de la Economía de la Educación, la función de producción educativa se utiliza como instrumento para modelar el proceso de enseñanza-aprendizaje y suele definirse como una función que asigna cantidades de *inputs* de las características de la escuela, profesorado y estudiante, a una medida de *output* educativo (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, & York, 1966; Bowles, 1970; Hanushek, 1979). Sin embargo, la extensa literatura sobre la función de producción educativa ha mostrado generalmente resultados contradictorios o débiles, lo que hace más difícil definir intervenciones políticas adecuadas. Debido a estas dificultades, cada uso de la función de producción educativa que se hace en esta tesis se discute y se apoya en la literatura previa, por lo que se pueden obtener resultados fiables.

Hasta este punto se han definido los sistemas educativos en los que se centra esta tesis y el instrumento empleado para analizarlos. Sin embargo, la realidad educativa es tan amplia y está compuesta por una gran cantidad de agentes – que también interactúan entre ellos – que tratar de abordar varios temas educativos al mismo tiempo podría ser problemático. Así, para alcanzar los objetivos indicados para esta tesis, algunos de los *outputs* e *inputs* que forman parte de la estructura de la función de producción educativa, en Andalucía y en España, se han agrupado y estudiado con más detalle de manera separada, centrando la atención en un objetivo particular en cada estudio. Para ello se analizan tres elementos del proceso de enseñanza-aprendizaje, por lo que cada uno de los tres capítulos de esta tesis se dedica a uno de estos elementos, comenzando por el alumnado, pasando por el profesorado, hasta llegar a los métodos de evaluación utilizados por profesorado e instituciones para evaluar a los estudiantes. Cada uno de estos tres capítulos está integrado por dos líneas de investigación – dedicando una sección del capítulo a cada línea, con la excepción del tercer capítulo, compuesto por una línea de investigación – orientada cada una a una casuística específica del elemento correspondiente. En los párrafos siguientes se presenta un breve esbozo sobre el contenido de estos tres capítulos.

Empezando por el primer capítulo, éste se encuentra dedicado al análisis de la función de producción educativa desde la perspectiva de los estudiantes. Los estudiantes son analizados en primer lugar debido a que son los recursos más importantes que posee todo sistema educativo y su aprendizaje es el objetivo último de estos sistemas. La primera cuestión que se analiza es la diferente manera en que los niños y las niñas aprenden y cómo estas diferencias pueden afectar a sus logros académicos. Uno de los resultados más robustos de los ILSA tests (PISA, TIMSS, PIRLS, etc.), para la mayoría de países y ciclos, es el rendimiento superior de las niñas en lectura, en comparación con los niños. Por el contrario, se observa con frecuencia que los chicos superan a las chicas en matemáticas, aunque este resultado parece ser más específico de cada país. Estas diferencias pueden apreciarse, por ejemplo, en el caso de PISA 2009 (OCDE, 2010), PISA 2012 (OCDE, 2014) y PISA 2015 (OCDE, 2016).

En este contexto, esta perpetuación de las diferencias de género en rendimiento académico parece requerir un análisis más profundo, a fin de determinar qué factores pueden estar condicionando esta situación. Debido a esto, la primera línea de investigación en este capítulo de los estudiantes se basa en el análisis de las diferencias de género en el rendimiento académico. El enfoque utilizado se centra en teorías sociales que tratan de explicar estas diferencias de género, en contraste con otras teorías biológicas[30]. En particular, esta investigación se centra en una variable particular que puede estar causando estas diferencias de género: las expectativas de los estudiantes, basándose en el supuesto de que las expectativas de los estudiantes se encuentran estrechamente relacionadas con su rendimiento académico final. En la medida en que se ha comprobado que estas expectativas son diferentes para los niños y las niñas (Fortin, Oreopoulos, & Phipps, 2015, para el caso de los Estados Unidos) esto podría significar que las niñas esperan asistir a la universidad en mayor medida que los niños, lo que podría explicar por qué las niñas estén dejando atrás a los niños en el colegio. Por ello, el objetivo de esta sección es comprobar la siguiente hipótesis: los niños y niñas andaluces presentan diferencias en rendimiento académico y las diferencias en sus expectativas sobre su nivel educativo futuro – y otras variables de estudiantes y familiares como, por ejemplo, estatus socio-económico o mala conducta – están condicionando de manera significativa estas diferencias en rendimiento académico.

Sin embargo, a pesar de la relevancia de este tema, la escasa información sobre las expectativas de los estudiantes sobre su futuro nivel educativo ha impedido a más investigadores analizar el efecto de esta variable en los resultados académicos de los estudiantes. Afortunadamente, la base de datos de ESOC10-SEN contiene información rica sobre el número de años de escolaridad que los niños y niñas esperan completar en el futuro. Esta información, junto con la metodología empleada, es decir, una mejora de la técnica de descomposición de Oaxaca desarrollada por Fortin, Oreopoulos, y Phipps (2015)[31], permite medir el efecto sobre la brecha de género en rendimiento académico de las diferencias de género en las expectativas de los estudiantes.

Como se ha destacado, el hecho de que existan diferencias de género en rendimiento académico indica que es necesario considerar las diferencias que niños y niñas presentan en su forma de aprender, por lo que esto se tendrá en cuenta en las siguientes líneas de investigación. La próxima sección se centra en analizar qué factores pueden contribuir – o hacer más difícil – que los estudiantes alcancen el mejor rendimiento académico que puedan alcanzar. Sin embargo, lograr que todos los estudiantes obtengan el rendimiento más alto en todas las asignaturas es muy poco realista, ya que los estudiantes tienen diferentes capacidades y

[30] Una revisión de la literatura de ambas teorías puede encontrarse en Francis y Skelton (2005).

[31] Esta metodología presenta la novedad de que permite obtener los efectos de respuesta educativa y de composición para las diferencias de género en rendimiento académico.

habilidades de aprendizaje. Por lo tanto, la hipótesis a contrastar en esta sección es la existencia de un conjunto de *inputs* educativos sobre los que es posible actuar para lograr un rendimiento equilibrado en tres asignaturas específicas: lectura (español), matemáticas e inglés (para estudiantes andaluces de primaria y secundaria), siendo el efecto de estos *inputs* diferente por sexo.

Esta hipótesis es contrastada mediante la optimización de la función de producción educativa, construida mediante técnicas econométricas. Esto supone la combinación de técnicas econométricas con metodologías de programación matemática – concretamente, programación multiobjetivo. Esta última metodología es empleada por su capacidad para generar soluciones eficientes mediante la optimización simultánea de varias funciones objetivo (una para cada una de las asignaturas consideradas) mediante un enfoque de puntos de referencia (Miettinen, 1999; Luque, Miettinen, Eskelinen, & Ruiz, 2009; Luque, Ruiz, & Steuer, 2010). Su aplicación al sistema educativo andaluz sería una novedad que podría suponer una ayuda importante para la toma de decisiones políticas orientadas a la obtención de un rendimiento académico equilibrado por parte de los estudiantes de primaria y secundaria en Andalucía. Se ha empleado la base de datos ESOC10-SEN para determinar este perfil del estudiante andaluz de primaria y secundaria más exitoso y "equilibrado" en términos del rendimiento académico en estas tres asignaturas, debido a la rica información que posee sobre los estudiantes y las características de sus familias.

Aunque los estudiantes presenten las características que les ayudasen a obtener un rendimiento académico equilibrado, no sería hasta que llegasen a la escuela y empezaran su proceso de aprendizaje que podrían desarrollar sus conocimientos y competencias. En este proceso, el profesorado es responsable de transmitir el conocimiento a los estudiantes y de hacerles aprender, lo que denota el importante papel que desempeñan los docentes en los sistemas educativos; es por esto que el profesorado es estudiado en el segundo capítulo. Sin embargo, aunque muchos trabajos de investigación se han centrado en la calidad del profesorado, no logran delimitar los diferentes aspectos que la componen. Esto es necesario en la medida en que un mejor conocimiento del proceso de enseñanza-aprendizaje mejoraría la aplicación de las políticas educativas, ya que estarían más enfocadas.

Con el fin de arrojar luz sobre la delimitación de los determinantes de la calidad del profesorado, la primera línea de investigación en este capítulo define los conceptos de eficacia y eficiencia del profesorado. A pesar de la relevancia de estos conceptos, suelen confundirse en la literatura de Economía de la Educación (Lockheed & Hanushek, 1994). La presente investigación pretende reconocer empíricamente los factores y las metodologías de enseñanza que podrían condicionar la eficacia del profesorado, en las que el uso eficiente de los recursos de clase por parte del profesorado podría ocupar un lugar relevante. Esta relación entre la eficacia del profesorado y la eficiencia no ha sido claramente establecida – a mi leal saber y entender – en la literatura anterior, por lo que esta investigación supondría una nueva visión en

la delimitación de estos conceptos. Esta delimitación se discute a continuación, junto con la relevancia del profesorado eficaz en los sistemas educativos.

La hipótesis a contrastar en esta sección para el profesorado español es doble: en primer lugar, el profesorado eficaz también es eficiente, siendo esta eficacia una forma relevante de medir el rendimiento del profesorado; en segundo lugar, hay factores que pueden clasificarse, tanto teórica como empíricamente, en factores relacionados con la eficacia o eficiencia del profesorado.

La identificación del profesorado eficaz es un tema relevante en la medida en que las "tablas de clasificación", creadas mediante el uso de la puntuación media del profesorado, no tienen en cuenta la variabilidad del muestreo y otras fuentes de error (Leckie & Goldstein, 2009). El uso de una clasificación adecuada del profesorado en términos de eficacia es particularmente relevante para que el profesorado eficaz pueda ser reconocido y recompensado en base a su desempeño en el fomento de altos niveles de participación y aprendizaje de los estudiantes. Por lo tanto, se propondrá una clasificación del profesorado basada en su eficacia como una mejor manera de clasificarlos de acuerdo a su desempeño, en comparación con el uso del promedio de las calificaciones de sus estudiantes.

La metodología utilizada para este análisis es un procedimiento de dos etapas, cuyo primer paso mide la eficacia del profesorado mediante el uso de fronteras estocásticas (*stochastic frontier analysis*, SFA) – una metodología aplicada, por ejemplo, por Battese y Coelli (1988) o Jürges y Schneider (2007) – empleando como *output* el rendimiento académico en lectura y matemáticas de los estudiantes y, como *inputs*, los recursos escolares disponibles para el profesorado. En el segundo paso, se han empleado las puntuaciones de eficiencia obtenidas en el primer paso, junto con las prácticas de aprendizaje en el aula del profesorado, para explicar el interés de los estudiantes en las lecciones de lectura y matemáticas, utilizando modelos de regresión multinivel – como Afshartous y Wolf (2007) o Grilli y Rampichini (2009) para el caso de las escuelas. La base de datos utilizada para este estudio es la combinación de PIRLS y TIMSS para el ciclo de 2011 de estudiantes españoles de cuarto grado (9/10 años).

Además, aunque se requiere de profesorado eficaz en cualquier sistema educativo, no sólo es importante encontrarlo y contratarlo, sino también atraer y retener en esta profesión al potencial profesorado eficaz. Sin embargo, el estatus que tiene el profesorado en España no es muy alto (Dolton & Marcenaro-Gutiérrez, 2011), lo que hace que la enseñanza sea una carrera poco atractiva. En este sentido, la satisfacción laboral de los docentes juega un papel relevante, al motivarlos para que realicen su trabajo de la mejor manera posible (Klassen & Chiu, 2010). Esta satisfacción ha recibido poca atención en la literatura y ha sido considerada como parte del rendimiento de los sistemas educativos, siguiendo los argumentos de muchos autores como Kumar (2014), quienes encontraron una relación positiva entre la eficacia del profesorado en educación primaria y la satisfacción en el trabajo, lo que significa que el profesorado más

satisfecho con su trabajo es el más efectivo en la enseñanza. Por lo tanto, el rendimiento académico de los estudiantes no es el único resultado relevante de los sistemas educativos, sino también su capacidad para hacer que el profesorado se sienta satisfecho con su trabajo. Además, como se ha destacado anteriormente, la distribución del rendimiento académico de los estudiantes españoles está desplazada a la izquierda, lo que denota que los estudiantes españoles tienen un rendimiento pobre en promedio (OCDE, 2016). Por lo tanto, en la segunda línea de investigación de este capítulo del profesorado, se definen y estudian simultáneamente varias formas alternativas de medir los resultados de la educación, distintos del rendimiento académico, esto es, el porcentaje de estudiantes que no alcanzaron los estándares de aprendizaje básico y la satisfacción del profesorado con el trabajo. Por lo tanto, estos dos *outputs*, junto con el rendimiento académico de los estudiantes, se van a analizar para el sistema educativo español, para conseguir un equilibrio entre ellos, pues exigir un máximo para todos ellos – mínimo en el caso de los estudiantes que no alcanzan un nivel básico de rendimiento académico – sería poco realista.

Por lo tanto, la hipótesis a contrastar en esta sección es doble: hay factores que contribuyen de manera simultánea a maximizar los "*outputs* buenos" (rendimiento promedio de los estudiantes y la satisfacción del profesorado con su trabajo) y a minimizar los "*outputs* malos" (porcentaje de los estudiantes que no alcanza los niveles básicos de aprendizaje) a lo largo del proceso de enseñanza-aprendizaje en España. La segunda es que la satisfacción del profesorado español es un *output* relevante a considerar, ya que permite mejorar el rendimiento académico en el sistema educativo.

Para realizar este estudio, en una primera etapa se estimaron las relaciones de las características del profesorado con su satisfacción y los resultados educativos del alumnado mediante métodos econométricos y, como segunda etapa, se aplicaron técnicas de programación matemática, concretamente, de programación multiobjetivo. Se utilizaron microdatos de TIMSS 2011 para España en cuarto grado. Como se van a considerar simultáneamente tres *outputs*, este estudio se va a centrar en la asignatura de matemáticas (rendimiento de los estudiantes en matemáticas, el porcentaje de estudiantes con un nivel básico de rendimiento en matemáticas y la satisfacción del profesorado).

Una vez que los niños y las niñas presentan un desempeño equilibrado y el profesorado es eficaz y está satisfecho con su trabajo, otra cosa importante para un sistema educativo es la forma de medición del desempeño de los estudiantes, de manera que cualquier desviación negativa en el rendimiento académico pueda ser identificada y resuelta de manera rápida y precisa. Debido a esto, el tercer capítulo de esta tesis se dedica a los métodos de evaluación. Concretamente, la línea de investigación de este capítulo pretende analizar los procedimientos de evaluación académica recientemente desarrollados, ya que estos deben medir de manera objetiva el rendimiento en los sistemas educativos (Battauz, Bellio, & Gori, 2011; Marcenaro-

Gutiérrez & Vignoles, 2015); de lo contrario, estos procedimientos de evaluación podrían potencialmente condicionar la relación entre los *inputs* y los *outputs*. El papel y el lápiz han sido las herramientas básicas para realizar los exámenes de evaluación en diferentes asignaturas (*Paper and Pencil Assessment*, PPA). Sin embargo, con el desarrollo de nuevas tecnologías, las pruebas de evaluación administradas mediante el uso de ordenadores están adquiriendo una importancia creciente hoy en día. Estas pruebas por ordenador (*Computer-Based Assessment*, CBA) pueden tener muchos factores positivos; sin embargo, este método de evaluación también tiene el riesgo de distorsionar la medición objetiva de las competencias de los estudiantes. Esto ocurriría si los resultados obtenidos dependieran de la habilidad del estudiante en el uso de ordenadores y no de sus verdaderas competencias en términos de la asignatura evaluada.

Si nos centramos en la comparabilidad de las puntuaciones en PPA y CBA, no existen resultados claros en este tema. En consecuencia, debido a estos resultados no confluentes, esta sección tiene la intención de determinar los efectos del uso de CBA en lugar de PPA como método de evaluación. En el contexto de una era digital, PISA 2012 aplicó una evaluación informática para el 40% de la muestra española (como en otros países) en las competencias tradicionales de lectura y matemáticas[32]. Gracias a esto, esta línea de investigación hace uso del análisis de las diferencias entre los resultados obtenidos en PPA y CBA por los estudiantes españoles para las competencias de lectura y matemáticas, empleando la base de datos PISA 2012. La hipótesis a contrastar es que las puntuaciones en CBA son diferentes de las obtenidas en PPA, debido a la influencia que tienen las habilidades de los estudiantes usando medios digitales sobre las puntuaciones en CBA. Para lograr este objetivo, se examinaron las características de los estudiantes y su entorno de aprendizaje, tanto en el hogar como en la escuela, mediante una metodología de modelos multinivel. En concreto, se estudiaron las características que podrían causar las posibles diferencias entre ambos tipos de pruebas en la medición del rendimiento académico de los estudiantes y el grado en que algunas de estas características podría aumentar o reducir la brecha entre ellas. Los resultados de esta sección tienen una gran relevancia, especialmente en el contexto de PISA, ya que la OCDE trata de consolidar el CBA como principal metodología de sus evaluaciones, a partir del ciclo de PISA 2015.

Tras delimitar todos los trabajos de investigación de esta tesis, a continuación se recogen y discuten los resultados obtenidos. Comenzando por los estudiantes, se obtuvo que los estudiantes andaluces de secundaria muestran diferencias de género en su probabilidad de suspender una asignatura, siendo ésta mayor para los niños. Además, las diferencias en las expectativas de los niños y las niñas pueden explicar casi toda la ventaja de las niñas en rendimiento académico, ya que ésta es causada en gran medida por la respuesta educativa de las

[32] Este 40% también realizó los tests en PPA.

niñas – diferencias de género no observadas en las características – a las expectativas. Así, cuando las niñas esperan completar un mayor número de años de escolaridad, presentan una mayor productividad en su trabajo y obtienen mejores resultados que los niños en lectura y matemáticas, lo que puede explicar gran parte de la brecha de género en el rendimiento académico; sin embargo, la contribución del efecto composición – diferencias de género observadas en las características – para los estudiantes andaluces es menor que la obtenida en otras investigaciones para los Estados Unidos (por ejemplo, Fortin, Oreopoulos, & Phipps, 2015). Otro resultado interesante es que los niños usan sus habilidades innatas en la escuela más que las niñas, mientras que las niñas tienen una mejor conducta, lo que permite a las niñas obtener mejores calificaciones[33]. Para concluir, se ha observado que los niños son más sensibles que las niñas a los cambios en el estatus socio-económico de sus familias.

Estos resultados sugieren la necesidad de realizar intervenciones políticas con el objetivo de ayudar a los niños que no tienen buenos resultados en comparación con las niñas, por lo que los programas de atención temprana pueden resultar útiles. Los programas de orientación vocacional y universitaria también son relevantes para proporcionar a los estudiantes la información adecuada para moldear sus expectativas, ya que se ha encontrado que éstas resultan importantes para explicar el rendimiento académico de los estudiantes.

Sin embargo, los estudiantes no presentan solamente diferencias de rendimiento académico debido a su género, sino también debido a diferencias en sus capacidades de aprendizaje y características personales. En este sentido, conseguir que todos los estudiantes obtengan el rendimiento más alto mediante la aplicación de ciertas políticas parece un objetivo poco realista para un gobierno. Por esta razón, conseguir el perfil del estudiante que obtiene el rendimiento académico más equilibrado en Andalucía parece ser un objetivo más interesante y alcanzable. En este sentido, se encontró que el aumento del tiempo que los estudiantes dedican a realizar las tareas por sí mismos en casa ayuda a lograr un rendimiento equilibrado, siendo este tiempo de trabajo aún más productivo que el de la enseñanza privada. Además, fomentar los hábitos de lectura convenciendo a los progenitores sobre su importancia a través de, por ejemplo, campañas publicitarias, podría ser útil para obtener un mayor rendimiento académico del estudiante; además, la intervención del profesorado a través de la transmisión de información a los progenitores sobre estas cuestiones también podría ser positiva. Uno de los principales resultados es el relacionado con las becas, ya que se ha encontrado que son útiles sólo para los estudiantes que obtienen, en promedio, un rendimiento inferior, por lo que el fomento de las becas podría motivar a este grupo particular de estudiantes. Los estudiantes repetidores también deben recibir suficiente atención, ya que obtienen un menor rendimiento académico, por lo que los esfuerzos de política deberían estar orientados a hacer los itinerarios

[33] Se encontró que este resultado no estaba basado en una mejor evaluación de las niñas por parte del profesorado.

vocacionales más atractivos para la juventud española, con el objetivo de ayudar a los estudiantes que no están obteniendo buenos resultados por la vía académica, de manera que puedan hacer uso de otras de sus competencias en estas vías vocacionales.

Se ha encontrado que la combinación de un enfoque multiobjetivo con técnicas econométricas es útil para identificar las características deseables que definen el perfil de los estudiantes que obtienen puntuaciones equilibradas en lectura (español), matemáticas e inglés y ha ayudado a detectar algunas cuestiones que pueden ser difíciles de identificar sólo con un análisis econométrico – como las diferencias de género en este perfil de logro equilibrado. Además, un análisis de sensibilidad ha mostrado la robustez de las soluciones obtenidas.

No obstante, aunque los estudiantes presenten todas las características necesarias para obtener resultados equilibrados en sus asignaturas, la forma en que son orientados para obtenerlos y aprender es muy importante para alcanzar estos resultados. Por ello, el trabajo del profesorado en la gestión de los recursos que le da la escuela (en los que está incluido el alumnado), junto con sus prácticas docentes orientadas a que el alumnado aprenda, son esenciales para la mejora de cualquier sistema educativo. Básicamente, los factores que componen estos dos tipos de actividades de los docentes han sido delimitados, tanto teórica como empíricamente, en prácticas eficientes y efectivas del profesorado – respectivamente. Una conclusión relevante obtenida de su análisis es que la eficiencia está integrada en el concepto de eficacia, por lo que pedir al profesorado que maneje sus recursos de manera adecuada también lograría que sus estudiantes estuvieran más interesados en el aprendizaje.

En cuanto al análisis de las variables de eficiencia del profesorado de cuarto grado, se ha comprobado que el uso adecuado de los recursos disponibles es esencial para interesar al alumnado en las lecciones de lectura y, por lo tanto, aumentar la eficacia del profesorado. El número de estudiantes por clase debería ser controlado e incluso podría ser aconsejable la división de clases en grupos, ya que las aulas superpobladas parecen ser un problema para las lecciones de matemáticas. También se encontró que una mayor cantidad de horas de enseñanza puede ser beneficiosa para los estudiantes, hasta el umbral de 5 horas a la semana[34].

En el caso de las metodologías empleadas por los docentes eficaces, se encontró que el profesorado debería organizar sus lecciones con otros docentes, pero que tendría que impartirlas por su cuenta. Además, se ha encontrado que el profesorado eficaz en lectura monitoriza y discute los deberes con sus estudiantes, mientras que el profesorado eficaz en matemáticas hace que los estudiantes realicen un examen escrito en matemáticas cada (o casi cada) lección y pone énfasis en las pruebas nacionales de matemáticas; por lo tanto, estas prácticas deberían ser incluidas en el plan de estudios del profesorado. Sin embargo, hay otras prácticas, como hacer

[34] Esto ha sido aplicado en España con la nueva legislación LOMCE (BOE, 2013) – Ley Orgánica de Mejora de la Calidad Educativa.

que los estudiantes trabajen en los problemas por su cuenta sin guiarlos durante el proceso o sin prestarles suficiente atención, que pueden hacer que los estudiantes pierdan interés en la lección. Esto ocurriría también con la metodología de lectura en voz alta en clase, debido a las dificultades que los estudiantes pueden tener para leer a su propio ritmo y entender el contenido y sentido del texto.

En general, estas metodologías de enseñanza eficaz deben incorporarse a las prácticas y al plan de estudios del profesorado. Además, el uso de la clasificación propuesta del profesorado, de acuerdo con su eficacia, puede ayudar a desarrollar políticas educativas y a fijar los salarios o compensaciones de los docentes de una manera más justa que mediante el uso de clasificaciones según rendimiento medio de su alumnado.

Por lo tanto, tener un profesorado eficaz es deseable en todo sistema educativo, pero encontrar, atraer, contratar y mantener a estos docentes podría ser una tarea difícil en países como España, donde el estatus del profesorado no es muy alto (Dolton & Marcenaro-Gutiérrez, 2011). Por ello, se ha propuesto que la satisfacción del profesorado podría que ser un mecanismo para atraer y retener en el medio y largo plazo a los mejores graduados del sistema educativo. Esto ha llevado a sugerir la satisfacción del profesorado como un *output* e indicador de lo bueno que es un sistema educativo, junto con el rendimiento académico y el porcentaje de estudiantes que no alcanzan los estándares básicos de aprendizaje.

Un análisis econométrico, combinado con un enfoque multiobjetivo, ha contribuido a identificar el perfil de aquellos docentes que alcanzan un resultado equilibrado en los tres *outputs* analizados. El análisis indica que el profesorado debería asistir a cursos de formación especializada para seguir mejorando la forma en que comunica sus conocimientos a los estudiantes y, también, felicitarlos – como una forma de retroalimentación – cuando están obteniendo buenos resultados. En este sentido, las autoridades educativas tienen un papel relevante en el fomento de la inclusión de estas prácticas en el plan de estudios ofrecido a los futuros docentes por los Institutos de Educación. También es necesario resolver, antes del comienzo del curso, el problema potencial de la falta de profesorado para las clases programadas.

La experiencia del profesorado también es un factor importante, ya que se obtuvo que tiene un impacto significativo y positivo en el rendimiento académico de los estudiantes hasta cierto punto; a partir de ahí, este efecto se reduce (teniendo una forma de U invertida). Este efecto marginal decreciente de la experiencia del profesorado podría reflejar pocos incentivos para hacer esfuerzos en la enseñanza, como resultado de una progresión casi plana de la carrera docente. Esta situación podría mejorarse mediante la creación de rutas de progresión profesional, basadas en objetivos equilibrados para cada uno de los diferentes *outputs* educativos. Además, los responsables de la formulación de políticas deberían prestar más atención a la promoción de los hábitos de lectura y a la sensibilización de los progenitores sobre

la importancia de la lectura a través de, por ejemplo, campañas publicitarias. Adicionalmente, el profesorado desempeña un papel importante en la promoción de los hábitos de lectura entre los estudiantes. En el caso de los estudiantes con diglosia, los responsables de la formulación de políticas educativas deberían esforzarse en la integración del alumnado de diferentes orígenes culturales, lo que contribuirá además a la cohesión social. Además, el sentimiento de seguridad del profesorado, en relación su integridad personal y sus condiciones de trabajo, tiene una influencia muy positiva en el logro de los objetivos indicados, por lo que la creación de un ambiente de trabajo adecuado debería ser un principio clave en todas las escuelas.

Por otra parte, es importante destacar que las soluciones del enfoque multiobjetivo muestran que la satisfacción del profesorado es el único indicador cuyo valor óptimo supera el máximo de la muestra, lo que indica que habría que hacer más esfuerzo para mejorarlo y que, al mismo tiempo, existe una gran posibilidad de intervención en éste. A la vista de estos resultados, puede afirmarse que la mejora de la satisfacción de los docentes puede ser un área prometedora para futuras investigaciones.

Con el fin de evaluar la forma en que el profesorado está realizando su trabajo y el nivel de rendimiento académico que los estudiantes están obteniendo gracias a sus lecciones, es necesario disponer de métodos de evaluación objetivos y precisos. Gracias a los nuevos avances tecnológicos, los métodos de evaluación han evolucionado hacia el formato digital, por lo que es necesario averiguar si estas nuevas metodologías digitales están midiendo las habilidades que se pretenden. Para ello, se han estudiado las diferencias de rendimiento académico presentadas por el alumnado en la evaluación de papel y lápiz (PPA) y las pruebas de evaluación por ordenador (CBA) en PISA 2012, en el caso de los estudiantes españoles.

Los resultados han mostrado que la realización de las pruebas en formato CBA "favorece" a las niñas más que a los niños, en el sentido de que ayuda a las niñas a mejorar su posición relativa en la competencia matemática, en la que las niñas obtienen una puntuación media más baja que los niños cuando son evaluadas por PPA. Por otra parte, la ventaja que las niñas ya tenían en la competencia de lectura se mantuvo sin cambios o incluso aumentó ligeramente en CBA.

Los resultados reportados en esta investigación han destacado muchos hechos interesantes: el uso de CBA podría hacer que los estudiantes "menos capacitados" reduzcan "artificialmente" la desventaja relativa que presentan al realizar pruebas mediante PPA, aunque esto no significa un mejor desempeño real en el proceso de enseñanza-aprendizaje. Por esta razón, las iniciativas de política educativa que tomen el marco de evaluación de PISA como un modelo de mejores prácticas deben ser cuidadosamente consideradas, ya que el efecto de las intervenciones en algunos grupos de estudiantes puede estar condicionado por el método de evaluación empleado. Esta es una cuestión relevante en la medida en que el CBA se utilizará para evaluar el desempeño de los sistemas educativos a partir de PISA 2015, por lo que resulta

necesario tener en cuenta que la influencia de ciertos factores puede potencialmente causar diferencias entre los resultados de PPA y CBA. Por lo tanto, los responsables políticos deben ser informados sobre estas diferencias, ya que la política educativa y las decisiones presupuestarias que afectan a estos factores podrían tener un efecto inesperado en el desempeño de los estudiantes, en comparación con sus resultados en PPA.

Para concluir, esta tesis ha tratado de analizar la realidad educativa de los sistemas educativos españoles y andaluces, centrándose en algunos aspectos de los tres elementos analizados del proceso de enseñanza-aprendizaje: estudiantes, profesorado y métodos de evaluación. Para ello, se han empleado varias metodologías econométricas y multiobjetivo – y una combinación de ellas. Para lograr estos objetivos, esta tesis ha estudiado la función de producción educativa de Andalucía y España. Aunque se han derivado muchas sugerencias de políticas en esta tesis, su aplicación requiere una implicación activa e inversión por parte del gobierno y otros participantes en las decisiones educativas – familias, docentes, etc. – para hacer que tengan éxito. Por esta razón, estos agentes deben gestionar sus recursos y confiar en los investigadores que realizan este tipo de análisis para mejorar sus sistemas educativos; de lo contrario, cualquier esfuerzo realizado en este campo de Economía de la Educación será en vano. Por la parte de los investigadores, el uso de datos empíricos representativos para la población estudiada refuerza la confianza que estos agentes podrían tener en los resultados obtenidos; además, la comprensión adecuada por parte de los investigadores de los datos estudiados es esencial para obtener conclusiones válidas.

Otra implicación relevante de esta tesis es que las conclusiones obtenidas para el país y la región bajo análisis no pueden extenderse directamente a otros países, ya que cada país presenta sus propias casuísticas educativas y culturales. Sin embargo, las metodologías y procedimientos empleados en esta tesis podrían ser utilizados como un primer enfoque para abordar las mismas temáticas de esta tesis, pero para otros países. Para concluir, la realidad educativa mundial es tan grande hoy en día que es imposible cubrir todos los temas que la engloban en solo una tesis. Es por esto que mis líneas de investigación futura tratarán otros importantes temas educativos que no han sido estudiados en esta tesis.

References of the Spanish summary (Referencias del resumen en español)

Afshartous, D., & Wolf, M. (2007). Avoiding 'data snooping' in multilevel and mixed effects models. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 170*(4), 1035–1059. doi: 10.1111/j.1467-985X.2007.00494.x

Battauz, M., Bellio, R., & Gori, E. (2011). Covariate error adjustment for multilevel models with application to educational data. *Journal of Educational and Behavioral Statistics*, *36*(3), 283–306. doi: 10.3102/1076998610366262

Battese, G. E., & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics, 38*(3), 387–399. doi: 10.1016/0304-4076(88)90053-X

BOE (2013). *Organic Law 8/2013, 9th December, for the improvement of the education quality (LOMCE)*. Nº 295, 10th December 2013, 97858–97921. Spain.

Bowles, S. (1970). Towards an educational production function. In W. Lee Hansen (Ed.), *Education, income, and human capital* (pp. 11–61). New York: Columbia University Press for the National Bureau of Economic Research.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Dolton, P., & Marcenaro-Gutierrez, O. (2011). If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance. *Economic Policy, 26*(65), 5–55. doi: 10.1111/j.1468-0327.2010.00257.x

EPA (2017). Encuesta de Población Activa. Último acceso septiembre 2017 desde http://www.ine.es/dynt3/inebase/es/index.htm?padre=982&capsel=986 ; Último acceso septiembre 2017 desde http://www.juntadeandalucia.es/institutodeestadisticaycartografia/iea/resultadosConsulta.jsp?CodOper=25&codConsulta=72916#consulta_72916

Eurostat (2017). Early leavers from education and training. Último acceso septiembre 2017 desde http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=t2020_40&plugin=1

Fortin, N. M., Oreopoulos, P., & Phipps, S. (2015). Leaving boys behind. Gender disparities in high academic achievement. *Journal of Human Resources, 50*(3), 549–579. doi: 10.3368/jhr.50.3.549

Francis, B., & Skelton, C. (2005). *Reassessing gender and achievement: Questioning contemporary key debates*. London: Routledge.

Grilli, L., & Rampichini, C. (2009). Multilevel models for the evaluation of educational institutions: a review. In P. Monari, M. Bini, D. Piccolo, & L. Salmaso (Eds.), *Statistical*

Methods for the Evaluation of Educational Services and Quality of Products (pp. 61–80). Heidelberg: Physica-Verlag. doi: 10.1007/978-3-7908-2385-1_5

Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources, 14*(3), 351–388. doi: 10.2307/145575

Hanushek, E. A., & Woessmann (2007). *The Role of Education Quality in Economic Growth*. World Bank Policy Research Working Paper No. 4122. doi: 10.1596/1813-9450-4122

IECA. (2010). Encuesta Social 2010. Educación y familia. http://www.juntadeandalucia.es/institutodeestadisticaycartografia/encsocial/2010/index.htm

IECA. (2017a). Premature dropout rate by sex. Último acceso septiembre 2017 desde http://www.juntadeandalucia.es/institutodeestadisticaycartografia/indsoc/indicadores/1038.htm

IECA. (2017b). Unemployment rate in Andalusia. Último acceso septiembre 2017 desde http://www.juntadeandalucia.es/institutodeestadisticaycartografia/iea/detalleDatosDia.jsp?cod=190&ram=D

Jürges, H., & Schneider, K. (2007). Fair ranking of teachers. *Empirical Economics, 32*(2-3), 411–431. doi: 10.1007/s00181-006-0112-3

Klassen, S. (2002). Low schooling for Girls, slower Growth for All? Cross country evidence on the Effect of Gender Inequality in Education on Economic Development. *World Bank Economic Review, 16*(3), 345–373. doi: 10.1093/wber/lhf004

Klassen, R. M., & Chiu, M. M. (2010). Effects on Teachers' Self-Efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress. *Journal of Educational Psychology, 102*(3), 741–756. doi: 10.1037/a0019237

Kumar, P. (2014). A study job satisfaction of British Education Trained teacher working at primary schools in Saharanpur district. *Journal of Indian Research, 2*, 141–145.

Leckie, G., & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society: Series A (Statistics and Society)*, *172*(4), 835–851. doi: 10.1111/j.1467-985X.2009.00597.x

Lockheed, M. E., & Hanushek, E. A. (1994). Concepts of Educational Efficiency and Effectiveness. In T. Husén, & T. N. Postlethwaite (Eds.), *International Encyclopedia of Education* (pp. 1779–1784). Oxford: Pergamon.

Luque, M., Miettinen, K., Eskelinen, P., & Ruiz, F. (2009). Incorporating Preference Information in Interactive Reference Point Methods for Multiobjective Optimization. *OMEGA-International Journal of Management Science*, *37*(2), 450–462. doi: 10.1016/j.omega.2007.06.001

Luque, M., Ruiz, F., & Steuer, R. E. (2010). Modified Interactive Chebysshev Algorithm (MICA) for Convex Multiobjective Programming. *European Journal of Operational Research, 204*(3), 557–564. doi: 10.1016/j.ejor.2009.11.011

Marcenaro-Gutierrez, O., & Vignoles, A. (2015). A comparison of teacher and test-based assessment for Spanish primary and secondary students. *Educational Research, 57*(1), 1–21. doi: 10.1080/00131881.2014.983720

Miettinen, K. (1999). *Nonlinear Multiobjective Optimization*. Boston: Kluwer Academic Publishers.

OECD. (2010). *PISA 2009 at a Glance*. OECD Publishing. doi: 10.1787/9789264095298-en

OECD. (2014). *Are boys and girls equally prepared for life?* OECD Publishing.

OECD. (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. PISA, OECD Publishing. doi: 10.1787/9789264229945-en

OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris: PISA, OECD Publishing. doi: 10.1787/9789264266490-en

Rivkin, S. G., & Schiman, J. C. (2015). Instruction Time, Classroom Quality, and Academic Achievement. *The Economic Journal, 125*, F425–F448. doi: 10.1111/ecoj.12315

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics, 39*(3), 312–320. doi: 10.2307/1926047