IEB
Institut
d'Economia
de Barcelona

Document de treball de l'IEB 2016/7

DOUBLE TOIL AND TROUBLE: GRADE RETENTION AND ACADEMIC PERFORMANCE

Álvaro Choi, María Gil, Mauro Mediavilla, Javier Valbuena

Human Capital

Documents de Treball de l'IEB 2016/7

DOUBLE TOIL AND TROUBLE: GRADE RETENTION AND ACADEMIC PERFORMANCE

Álvaro Choi, María Gil, Mauro Mediavilla, Javier Valbuena

The **IEB** research group in **Human Capital** aims at promoting research in the Economics of Education. The main objective of this group is to foster research related to the education and training of individuals and to promote the analysis of education systems and policies from an economic perspective. Some topics are particularly relevant: Evaluation of education and training policies; monetary and non-monetary effects of education; analysis of the international assessments of the skills of the young (PISA, PIRLS, TIMMS) and adult populations (PIAAC, IALS); education and equality, considering the inclusion of the disabled in the education system; and lifelong learning. This group puts special emphasis on applied research and on work that sheds light on policy-design issues. Moreover, research focused in Spain is given special consideration. Disseminating research findings to a broader audience is also an aim of the group. This research group enjoys the support from the **IEB-Foundation.**

The **Barcelona Institute of Economics (IEB)** is a research centre at the University of Barcelona (UB) which specializes in the field of applied economics. The **IEB** is a foundation funded by the following institutions: Applus, Abertis, Ajuntament de Barcelona, Diputació de Barcelona, Gas Natural, La Caixa and Universitat de Barcelona.

Postal Address:
Institut d'Economia de Barcelona
Facultat d'Economia i Empresa
Universitat de Barcelona
C/ John M. Keynes, 1-11
(08034) Barcelona, Spain
Tel.: + 34 93 403 46 46
ieb@ub.edu
http://www.ieb.ub.edu

Documents de Treball de l'IEB 2016/7

DOUBLE TOIL AND TROUBLE: GRADE RETENTION AND ACADEMIC PERFORMANCE *

Álvaro Choi, María Gil, Mauro Mediavilla, Javier Valbuena

ABSTRACT: Most recent available evidence points to the scarce efficacy of grade retention for levelling the performance of students. Yet, the fact that many countries persist in applying this measure reflects longstanding traditions, cultural factors and social beliefs as well, it would seem, the lack of robust empirical evidence to do otherwise. We contribute to the literature by analysing the impact of grade retention on the reading competencies of lower secondary school students in Spain, a country where almost one out of every three students will repeat at least one grade by age 16. We overcome the absence of longitudinal data by creating a pseudo-panel that combines microdata from two international assessments, PIRLS and PISA. Having controlled for reverse causality, our study confirms the negative and heterogeneous impact of grade retention. This paper provides new evidence of the pressing need to rethink this educational policy, and our results highlight the importance of early intervention as opposed to only employing remedial measures.

JEL Codes: I21, I28, I24

Keywords: Grade retention, academic achievement, PISA, PIRLS

Álvaro Choi
Universitat de Barcelona & Barcelona Institute of Economics.
690, Av. Diagonal
08034 – Barcelona, Spain
E-mail: alvarochoi@ub.edu

María Gil
Universidad Autónoma de Madrid
Campus de Cantoblanco
C/ Francisco Tomás y Valiente, 5
28049, Madrid, Spain
E-mail: maria.gil@uam.es

Mauro Mediavilla
Universitat de València & Barcelona Institute of Economics.
Avda Tarongers s/n
46022 - Valencia, Spain
E-mail: Mauro.Mediavilla@uv.es

Javier Valbuena
Universitat de Girona
Campus Montilivi
17071 – Girona, Spain
E-mail: javier.valbuena@udg.edu

* The authors gratefully acknowledge the financial support from the Areces Foundation through its XIII National Contest for Research in Social Sciences and the Spanish Ministry of Economy and Competitiveness (project EDU2013-42480-R).

1. Introduction

The efficacy of grade retention, the educational practice of holding a student back for an extra year in the same grade, is not only of academic interest, but also of enormous policy relevance given its consequences for students and schools alike (De Witte et al., 2013). Most recent available evidence seems to stress its negative effect on academic performance (Andrew, 2014; Belot et al., 2014; Glick et al., 2010; Jacob et al., 2009); however, many OECD countries, albeit with considerable variation in their grade retention rates, persist in applying this policy as their primary method for enhancing the academic performance of low achievers.

In this paper we assess the impact of grade retention on the reading competencies of lower secondary school students in Spain, a country with a high retention rate and for which no robust evidence is yet available, which can, in part, be explained by the country's lack of longitudinal data. We overcome this issue by introducing prior educational achievement, a key factor in understanding grade retention as the consequence of a cumulative process.

Hence, the aim of this paper is twofold. First, we present robust estimations of the impact of grade retention on academic performance for the Spanish case considering earlier schooling achievement. Second, we propose a novel empirical approach that allows us to overcome methodological issues associated with Spanish data, including reverse causation. Additionally, this approach should be of interest to comparative education researchers in other countries facing similar data constraints.

Using a parametric method we create a pseudo-panel that combines micro-data from the 2006 and 2012 waves of the Progress in International Reading Literacy Study (PIRLS) and the Programme for International Student Assessment (PISA), respectively. Our results highlight the importance of previous achievement and are consistent with a substantial academic performance gap between 'repeaters' (i.e., those asked to repeat the grade) and 'non-repeaters' (i.e., those that advance to the next grade). Furthermore, the results describe negative non-linear grade retention effects along the distribution of primary education grades.

The rest of the paper is organised as follows. Section 2 provides a brief literature review and some background on the Spanish educational system and its application of grade retention. Section 3 describes the data, methodology and their relative merits. Section 4 outlines relevant features of the results. Section 5 concludes.

2. Grade retention as a universal solution

Although available empirical evidence tends to underline the negative effects of grade retention, several OECD countries still apply it as their main policy for tackling low academic performance (Figure 1). While countries such as Japan and Norway prefer automatic promotion, others, including Belgium and Portugal use grade retention intensively. Tradition, cultural factors and social beliefs regarding the benefits and effectiveness of grade retention seem to play an important role in accounting for these differences (Goos et al., 2013).

INSERT FIGURE 1 AROUND HERE

In the Spanish education system, grade retention is widely applied: almost a third of students will have repeated at least one grade by age 16. Moreover, Figure 1 shows that the number of repeaters in Spain increased between 2003 and 2012. According to Dupriez et al. (2008), this means that among those countries with a comprehensive educational system Spain can be classified as one in which grade retention is the main policy for levelling student performance. Figure 2 provides specific information for the Spanish case.

INSERT FIGURE 2 AROUND HERE

Figure 2 shows that grade retention occurs mainly during compulsory secondary education (ages 12 to 16) and is more common among Spanish boys.[1] This seems to be consistent with their higher early school dropout rates (25.6% boys vs. 18.1% girls, in 2013), in line with a number of studies, including Jacob and Lefgren (2009) and Ou and Reynolds (2010), that describe a positive relationship between grade retention and early school dropout. However, although early school dropout figures have fallen since 2008

[1] Education in Spain is compulsory from ages 6 to 16. This comprises six years of primary school and four years of lower secondary education. Although not compulsory, education from ages 3 to 5 is free - in public and private publicly-funded schools – and nearly universal. According to the Spanish Education Act (*Ley Orgánica de Educación* – LOE, art. 20.4), the education act in force during the period analyzed herein, students may only repeat one grade during primary school, and two grades during lower secondary education (ESO). Teachers decide which students should be retained, the main criteria being that of having failed three or more subjects. However, teachers are allowed to promote students that have not passed three subjects. Although the LOE is a national law, there are major differences in retention rates between and within the Spanish regions (*Comunidades Autónomas*).

(mainly due to the economic crisis and high rates of youth unemployment), both Spanish girls and boys were still, in 2014, well above the 10% target established in the EU 2020 Strategy benchmark (15% for the Spanish case).

The main reason for applying the practice of grade retention is to *punish* the students' poor performance, i.e., their inability to pass a certain number of subjects, as established by the syllabus. Several arguments have been used in defence of its use. First, it provides students with time to mature; second, it serves the purpose of establishing minimum academic requirements in order to advance to the next grade; finally, it aims to enhance overall performance by transmitting to students a culture of effort. In this sense, the claim is that it acts as a deterrent to low performance (Manacorda, 2012). Overall, it is assumed that retention can improve the academic performance of low achievers by exposing them to an additional year of teaching so they catch up in terms of curriculum requirements. Nevertheless, grade retention remains a controversial measure.

Those opposed to grade retention emphasise its inefficacy (Jimerson et al., 2002), its high cost (OECD, 2011) and its negative impact on student motivation, given that those required to repeat the grade are separated from their friends, obliged to retake not only the subjects they failed but those they passed, and to suffer the potential stigma of being labelled "slow" students (Martin, 2011). Moreover, the practice may also generate discipline issues in schools (Crothers et al., 2010). All in all, it might negatively affect academic performance and increase the probability of school dropout (Holmes, 1989).

At the international level, there is a vast literature within the economics of education concerned with grade retention policies. Yet, the results are inconclusive ranging from substantial negative effects on academic achievement, to null grade retention and even positive effects (Jacob and Lefgren 2004). In a recent meta-analysis, Allen et al. (2009) highlight the crucial role of selection bias in determining the results of short-term grade retention effects. In the absence of pre-retention measures of academic ability, results show a much stronger negative association between retention and academic performance. Additionally, there is a growing literature on estimating the causal effect of grade retention policies on academic achievement (Dong, 2009; Eide and Showalter, 2001; Glick and Sahn, 2010; Gomes-Neto and Hanushek, 1994; Jacob and Lefgren, 2004, 2009 and Manacorda, 2012). Empirical findings are also mixed showing both negative and positive effects of retention on academic performance and school dropout.

Compared to the large body of research at the international level, there is a relative dearth of empirical research into the effect of grade retention in Spain. Indeed, no more than a few recent empirical studies examine this education policy, the results of which can be summarised as follows: grade retention shows a negative association with academic performance and increases the probability of school failure i.e., dropping out of school before completion of compulsory education (Calero et al., 2010; Calero and Escardíbul, 2007; Choi and Calero, 2013; Cordero et al., 2010; Guío and Choi, 2014; Mancebón et al, 2012; Salinas and Santín, 2012). The main limitation presented by these studies is that they are unable to determine whether retention is the direct cause of poor academic achievement/school failure, or the result of the students' prior characteristics that increase their probability of failure. In the case of the latter, retention would simply be a signal of those students at greatest risk of academic failure (i.e., reverse causation). As such, most studies analysing the Spanish case do not estimate precisely the effect of grade retention on academic achievement. This is due, in the main, to the lack of sufficiently rich data (i.e., longitudinal data) for tackling methodological issues such as reverse causation and endogenous selection. To the best of our knowledge, only García-Pérez et al. (2014) sought to control for endogeneity using a switching regression model, but their results crucially depend on endogenous selection. Therefore, there is a lack of robust empirical evidence for Spain.

3. Data and methodology

It is not uncommon for researchers in the social sciences to face various data limitations when dealing with specific issues that might be of great interest to both the scientific community and policymakers, despite the fact that databases are becoming increasingly more exhaustive. The analysis of grade retention in Spain is a good example of this: thus, while the practice is widely applied in the Spanish education system, there are no databases offering sufficient information to assess it properly. In this paper, we attempt to overcome this limitation by creating a pseudo-panel that merges microdata from two international cross-sectional databases. In this section, we describe the two databases used (3.1) and the methodology applied (3.2).

3.1. Data

The use of longitudinal data has a number of advantages over that of either purely cross-sectional or time-series data. The main advantages are the ability to study dynamic relationships and to model the differences or heterogeneity between subjects. The absence of such a database for Spain has led us to create a pseudo-panel that combines microdata from two international cross-sectional databases, namely, PIRLS 2006 and PISA 2012, which follow the same cohort of students.

PISA 2012 assesses, on a triennial basis, the extent to which 15-year-old students have acquired key competencies and skills. The assessment, which focuses on reading, mathematics, science and problem-solving, does not just evaluate whether students can reproduce what they have learned, it also examines how well they can extrapolate from what they have learned and apply that knowledge to unfamiliar settings. This reflects the fact that modern societies reward individuals not for what they know, but for what they can do with that knowledge. A total of 65 countries, 34 belonging to the OECD and 31 partner countries, participated in the PISA 2012 assessment (OECD, 2014). However, PISA does not follow the evolution of students over time and provides no information on their previous achievement. Given the existence of issues of reverse causality, the direct estimation of an educational production function that attempts to measure the impact of grade retention would therefore be biased. Yet, this bias could be reduced by controlling for prior academic performance.

To overcome this we turn to our auxiliary sample, the 2006 data from PIRLS (Progress in International Reading Literacy Study), and the second in the series of studies carried out by the International Association for the Evaluation of Educational Achievement (IEA). Inaugurated in 2001 and conducted every 5 years, PIRLS is the IEA's assessment of students' reading achievement in fourth grade (ages 9/10). PIRLS 2006, implemented in 40 countries, assessed a range of reading comprehension strategies focusing on two main reading purposes – literary and informational.[2] In addition, the database provides information on the students' individual, household and school characteristics (Mullis et al., 2007). Interestingly, for our purposes here, most students participating in PIRLS 2006 were born during 1996 and so belong to the same cohort as PISA 2012 students.

[2] Unfortunately, Spain did not participate in the Trends in International Mathematics and Science Study (TIMSS)-2007 and so we are unable to replicate our results for maths and science.

3.2. Methodology

Our analysis focuses on the reading skills of students aged 15/16 (PISA 2012) taking into account their predicted score at age 9/10, as extracted from the PIRLS 2006 database. This original approach is possible as both international assessments provide relevant information for students belonging to the same cohort. Moreover, both assessments are regarded as being representative at the national level and share similar sampling designs and response rates. Thus, we are able to identify a set of individual and household level variables present in both databases that are relevant for estimating academic performance. This allows us to apply a parametric approach for merging PIRLS 2006 results with those from the PISA 2012 database.

Our approach can be related to the two-sample two-stage least squares (TSTSLS) methodology, which has been widely used in the intergenerational mobility literature[3]. To the best of our knowledge only De Simone (2013) has used a similar approach for assessing educational issues. In our study, we specifically draw on the PISA 2012 data as our ultimate objective is to identify the impact of grade retention during lower secondary school on the academic performance of 15- to 16-year-old students.

3.2.1. Imputing PIRLS 2006 and PISA 2012

The existence of a non-negligible amount of missing values is an issue when dealing with PIRLS and PISA data. Thus, before using these sources as donors for the main database, we need to *clean* the variables we intend using for our analysis. To do so, we use the multiple imputation technique.

Although known since the 1970s (Rubin, 1976), the development and implementation of the technique have been extended in recent years (Acock, 2005; Royston, 2005; Rubin, 1996; Sterne et al., 2009; van Buuren et al., 1999; van Buuren, 2012). This stochastic technique allows full use to be made of the data and unbiased estimators to be obtained, reflecting the uncertainty introduced by the non-response in the parameter estimation while preserving the dispersion of the distribution of the imputed variable (Rubin, 1996). Its implementation is based on replacing the unobserved data for m>1

[3] Jerrim et al. (2014) provide a recent review.

possible simulated values from the maximum incorporation of predictive variables with missing values (Schafer, 1999; White et al., 2011).

The applicability of the method, in its general outline, has been enhanced with Monte Carlo methods based on Markov chains, known as MICE (Multiple Imputation by Chained Equations) algorithms. Besides, multiple imputation is considered a flexible methodology for working with multivariate data and monotonous or arbitrary distributions of missing values. Finally, its suitability requires that the distribution pattern of missing values is random (MCAR – missing completely at random, or MAR – missing at random).

The PIRLS 2006 database for Spain consists of 4,094 observations. For the purpose of the analysis, we consider only students born in 1996, which gives us a final sample of 3,771 observations. Missing value issues were detected. For some observations there is no information on child schooling and education/occupation of the parents. Since this information is of relevance to our investigation, these observations were eliminated from the sample (listwise deletion), but we subsequently checked to ensure the resultant database presented similar average values for all the variables.

The PISA 2012 database for Spain consists of 25,313 observations. Missing value issues were also detected (Table A.1.1 in Appendix 1). In this case we found 131 observations for which there was no information regarding the majority of the variables of interest. As in PIRLS, they were eliminated from the sample (listwise deletion). The descriptive analysis of the resulting databases is available upon request.

As a preliminary step for imputation, we evaluate the randomness of missing values using the dichotomized test of correlations and checking for the absence of outliers. We apply the multiple imputation technique using the MICE algorithm (Royston and White, 2011; StataCorp, 2013). The MICE method is implemented in the chained method and uses a Gibbs-like algorithm to impute multiple variables sequentially using univariate fully conditional specifications. Finally, the estimation imputes the variables in a specific order – from the most to the least observed. In this case, following the recommendation of Rubin (1996) and Acock (2005), we use all the available variables in the model to estimate unobserved data from three different empirical approaches (logit, ordered logit and multinomial logit) according to the particular characteristics of each variable. After the imputation process, we have a final database with 40 complete simulated databases of PIRLS and PISA.

3.2.2. Merging PIRLS and PISA data: creation of a pseudo-panel

We consider a parametric approach for linking PIRLS 2006 scores to those of PISA 2012: we first estimate an educational production function drawing on the PIRLS (auxiliary) database, and using as regressors those individual and household level variables also available in the PISA (main) sample.[4] We then apply the parameters obtained in the former regression to the PISA sample and obtain the predicted value that a student in the PISA database would have obtained in PIRLS. We therefore add an additional column to the PISA 2012 database: the student's predicted score in PIRLS 2006. This procedure is repeated five times for each plausible value in PIRLS. As can be seen, this methodology follows a similar approach to that of the TSTSLS technique suggested by Arellano and Meghir (1992).

3.2.3. Hierarchical linear model

PISA designs its sample using a two-stage method. In the first stage, a sample of schools is randomly selected from the whole list of schools that provide schooling for 15-year-old students. In the second stage, a random sample of 35 students is chosen from each of the schools selected in the first stage. The probability of a school being selected in PISA is proportional to its size. As a consequence, larger schools have a higher probability of being chosen; however, students at larger schools have a lower probability of being selected than students enrolled at the smaller schools that have been chosen in the PISA sample. Therefore, the principle of independence of variables among the students of each school does not hold, given that students enrolled at the same school typically share similar socio-economic circumstances, which means the average correlation between the student variables within a school are higher than that found between students at different schools (Hox, 1995). Consequently, we rely on a hierarchical linear model (HLM, hereafter) to take into account the nested structure of the database. In our analysis, our data are structured at two levels: students (level 1) and schools (level 2). We estimate the following equations:

[4] OLS estimates from this auxiliary regression are available upon request.

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^{n} \beta_{1j} X_{kij} + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0, \sigma^2) \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \sum_{1}^{j} \gamma_{01} Z_{lj} + \mu_{0j} \quad \mu_{0j} \sim N(0, \tau_0) \quad (2)$$

$$\beta_{1j} = \gamma_{10} \quad (3)$$

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{kij} + \gamma_{01} Z_{1j} + \mu_{0j} + \varepsilon_{ij} \quad (4)$$

where Y_{ij} is the change in the reading skills score of student "i" enrolled at school "j" between ages 10 and 15/16. X_{kij} is a vector of "k" independent variables at the individual level and Z_j is a vector of "l" variables at the school level. Equation 4 is obtained by substituting equations 2 and 3 (level 2) for the β in equation 1 (level 1). In our model specification, we estimate fixed effects (eq. 3).

The dependent variable is the difference between the scores in the reading skills of students between ages 10 and 15/16. This difference is calculated using the sets of plausible values (random values calculated from the distribution of the results in the assessments) provided by PIRLS and PISA. This gives a total of 25 combinations. Results in PIRLS and PISA are originally scaled to a mean of 500 and a standard deviation of 100 *within* each of the respective surveys. A score of 500 points in PIRLS, however, is not equivalent to a score of 500 points in the PISA scale, given the different number of countries participating in the assessments. Therefore, following Brown et al. (2007), we tackle this issue using international z-scores for the countries participating in both assessments, standardising scores for each survey at the student level, with a mean of 0 and a standard deviation of 1, across 25 jurisdictions.

A set of individual, household and school level variables have been included in the model (Table A.1.2 in Appendix 1). Additionally, interactions between the predicted score in PIRLS 2006 (quartiles) and grade retention have been introduced. This allows us to take into account previous performance and to assess different impacts of grade retention on students with different profiles (from lower to higher achievement). By so doing we are able to overcome the reverse causality issue that, as far as we know, affects all the previous results in the literature reporting the impact of grade retention in Spain. Nevertheless, we acknowledge that, while efforts have been made to introduce a wide range of controls into the analysis, we cannot discard the possibility that unobservable variables may affect the results. Individual- and school-level weights have been applied throughout.

4. Results and discussion

This section is structured in two subsections. Subsection 4.1 provides a brief overview of the main characteristics of repeaters and non-repeaters. The results of the hierarchical linear model outlined in the section above are presented in subsection 4.2.

4.1. Descriptive analysis

Table 1 provides information for a set of features in which repeaters and non-repeaters differ. The former seem to perform worse than non-repeaters at age 9/10; however, the gap between the two groups increases greatly between ages 9/10 and 15/16. This may indicate the negative impact of grade retention, although we acknowledge, as Table A.1.2 shows, that there is a reduction in the dispersion of scores at age 10, due to the merging method applied that depends on a limited set of variables.

INSERT TABLE 1 AROUND HERE

A non-surprising finding for the Spanish case (see Section 2) is the fact that boys seem to repeat a grade more frequently than girls. Spanish girls, as in most countries (OECD, 2014), outperformed boys in the PISA 2012 reading competence too. Small differences were also found in the distribution of repeaters and non-repeaters by date of birth, as they were when analysing the percentage of students that did not attend pre-primary education. However, the fact that pre-primary education in Spain is, while not compulsory, nearly universal reduces the potential weight of this variable for explaining differences between the two groups.

Household characteristics seem to be relevant for understanding grade retention. First, the percentage of non-repeaters that live with both their parents is considerably higher than for repeaters. This finding is in line with other studies including Fernández Enguita et al. (2010). Second, being a first or second generation immigrant seems to increase the likelihood of repeating a grade during lower secondary school. Cordero et al. (2014) find differences in the probability of grade retention between first and second generation immigrants. And third, repeaters tend to live in households with a lower socioeconomic

status. This finding may well be related to the fact that the percentage of repeaters enrolled at public (state) schools is higher than that at private schools (i.e., in 2012, 61.4% of non-repeaters and 76.8% of repeaters were enrolled in public schools). A number of studies, including Escardíbul and Villarroya (2009) and Mancebón et al. (2010), have described the existence of socioeconomic-status-based student selection processes in Spain not only by private independent schools, but also by publicly-subsidised schools.

There also seem to be marked differences in retention rates between the Spanish regions (Table A.1.2), which might point to the existence of different "cultures of grade retention" within Spain. This clearly runs counter to one of the main arguments in favour of applying grade retention, namely, equality of treatment. Although analysing this question falls beyond the scope of this paper, it is clearly one that merits further attention in the future.

4.2. HLM Results

Our empirical approach seeks to quantify the impact of grade retention during secondary education on students' reading competencies. This means that we are not solely interested in the score the student obtained at age 15/16, but also on the gain or loss that can be attributed to the policy. Thus, our outcome of interest is the difference in attainment between primary (PIRLS) and secondary school (PISA). Although the educational production function in our analysis included a set of individual, household and school level variables, we focus our attention specifically on variables aimed at identifying the impact of grade retention. Coefficient estimates for all the variables are presented in Table 2.

INSERT TABLE 2 AROUND HERE

Our results show a stark negative impact of grade retention on the reading competencies of students. The difference in reading competencies, *ceteris paribus*, between primary and secondary education of the students retained one grade during secondary school is, on average, 0.41 standard deviation points in favour of the non-repeaters. The magnitude of this negative effect increases by more than 60% when a student is retained for two or more grades years at this educational stage, indicating that the negative effect

is also cumulative. Furthermore, we have included in our estimation equation interaction terms measuring the impact of retention across the distribution of scores in primary education (i.e., quartiles). Our results indicate that the magnitude of the negative effect is decreasing in prior academic performance, affecting more severely the best students among the low achievers. This has policy implications, as it suggests that schools should not apply this policy homogeneously, if at all, among their low performing students. We interpret this finding to be an indication of a strong negative impact on the motivation and self-esteem of the individuals that are retained (Holmes, 1989).

To seek confirmation of our results, we proceeded to replicate the analysis by matching PISA 2012 scores with data from the 2007 *Evaluación General de Diagnóstico* (EGD), a national scale assessment tool measuring the performance of Spanish students in mathematics, science and reading comprehension at age 12. The results obtained in this auxiliary analysis clearly scaffold our earlier findings. A brief description of the 2007 EGD and the main results of the auxiliary analysis are reported in Appendix 2.

As for the remaining individual and household controls, positive coefficients were associated with students living in a household with a non-nuclear family structure (i.e., single parent household) and first generation immigrant households. Here, the definition of the outcome variable needs to be borne in mind. A positive coefficient indicates that the difference in performance for the two respective groups (between ages 9/10 and 15/16) increased slightly. Auxiliary analyses, in which the dependent variable is the performance of secondary school students, show that these continue to be factors that have a negative impact on academic performance at age 15/16. The socioeconomic status variable (ESCS index) is also statistically significant in our model. This means the socioeconomic-based gap in reading competence increases between primary and secondary school. This result is in line with Choi and Jerrim (2016) and stresses the need for the early identification of low socioeconomic status students as students at risk and, hence, the need to bolster targeted support mechanisms.

In the case of school variables, grade retention shows a greater effect on students that live in small cities. The results of students that attend schools with greater autonomy in the allocation of their resources also seem to improve between ages 9/10 and 15/16.

5. Conclusions

Grade retention is widely used in Spain despite the fact that a) educational legislation considers it a policy of last resort and b) there is a lack of consistent studies determining its efficacy for improving academic performance and, subsequently, reducing school failure and early school dropout. While social beliefs and teacher attitudes may play a role in accounting for the former (Arregi et al., 2009), the lack of adequate data for addressing methodological issues, such as reverse causation, explain the latter.

Our results confirm that, once previous performance is taken into account, the sign of the effect of grade retention on the reading competencies of Spanish students remains negative. These results are in line with most of the previous literature on the impact of grade retention described in Section 2. Equally noteworthy is the heterogeneous impact of grade retention during lower secondary school. The reading competencies of secondary school repeaters that performed relatively better during primary school declined relatively more. This might be related to a greater fall in student motivation. Unfortunately, our data do not allow us to go any further than the formulation of this hypothesis. Yet, independent of the mechanism that accounts for this result, there is a clear policy implication: while teachers should be careful when deciding which students need to repeat a grade, they need to be especially cautious when retaining students whose previous academic performance was relatively better.

The negative impact of grade retention and the importance of previous achievement have further policy implications. The most obvious is perhaps the need for the early detection of students at risk of grade retention during the initial stages of their education. The results presented in Section 4 provide a clear indication of the characteristics of students at risk of grade retention. Among other possible measures, ensuring targeted supports and services are available for poorly performing students that are automatically promoted might serve as an alternative to grade retention (Darling-Hammond, 1998).

Finally, we should acknowledge certain limitations in the study reported here. First, we focus on the short-term impact of retention at the secondary school level. Empirical evidence available from other countries seems to describe different effects of grade retention at earlier stages in the education system. Grade retention may also have long-term effects, such as an impact on the probability of accessing higher education

(Andrew, 2014). Second, we focus only on reading competencies and cannot discard the possibility that the effect of grade retention may be heterogeneous by competencies. Finally, we recognise our study does not control for unobservable variables such as motivation. Nevertheless, while we await better databases, this paper has reported an innovative methodology – which should be of interest for other countries facing the same data constraints as Spain – to offer strong evidence of the ineffectiveness of grade retention in a country where it is widely applied, and to make a telling case for the reconsideration of this policy.

References

Acock, A. (2005). Working with Missing Values. *Journal of Marriage and Family*, 67: 1012-1028.

Allen, C. S., Chen, Q., and Willson, Hughes V. L. (2009). Quality of Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multi-Level Analysis. *Educational Evaluation and Policy Analysis*, 31: 480–99.

Andrew, M. (2014). The Scarring Effects of Primary-Grade Retention? A Study of Cumulative Advantage in Educational Career. *Social Forces*, 93: 653-685.

Arellano, M., and C. Meghir. (1992). Female labour supply and on-the-job search: an empirical model estimated using complementary data sets. *The Review of Economic Studies*, 59: 537–559.

Arregi, A., Martínez, P., Sainz, A. and Ugarriza, J. R. (2009). *Efecto de las repeticiones de curso en el proceso de enseñanza-aprendizaje del alumnado*. Bilbao: ISEI-IVEI.

Belot, M. and Vandenberghe, V. (2014). Evaluating the 'threat' of grade repetition: exploiting the 2001 reform by the French-Speaking Community of Belgium. *Education Economics*, 22: 73-89.

Brown, Giorgina, Micklewright, John, Schnepf, Sylke V., and Robert Waldmann. (2007). International Surveys of Educational Achievement: How Robust are the Findings? *Journal of the Royal Statistical Society Series A,* 170: 623–646.

Calero J., Choi A., and Waisgrais S. (2010). Determinantes del Riesgo de Fracaso Escolar en España: Una Aproximación a través de un Análisis Logístico Multinivel aplicado a PISA 2006. *Revista de Educación,* 2010: 225-256.

Calero J., and Escardibul J. (2007). Evaluación de Servicios Educativos: el rendimiento en los centros públicos y privados medido en PISA 2003. *Hacienda Pública Española*, 183: 33-66.

Choi, Á., and Calero J. (2013). Determinantes del riesgo de fracaso escolar en España en PISA 2009 y propuestas de reforma. *Revista de Educación*, 362: 562-593.

Choi, Á., and Jerrim, J. (2016). The use and (misuse) of PISA in guiding policy reform: evidence from Spain. *Comparative Education.* DOI: 10.1080/03050068.2016.1142739.

Cordero, J.M., Crespo, J.M., and Santín, D. (2010). Factors Affecting Educational Attainment: Evidence From Spanish PISA 2006 Results. *Regional and Sectoral Economic Studies*, 10: 55-76.

Cordero, J.M., Manchón, C., and Simancas, R. (2014). La repetición de curso y sus factores condicionantes en España. *Revista de Educación*, 365: 12-37.

Crothers, L.M., Schreiber, J.B., Schmitt, A.J., Bell, G.R., Blasik, J., Comstock, L.A. et al. (2010). A Preliminary Study of Bully and Victim Behavior in Old-for-Grade Students: Another Potential Hidden Cost of Grade Retention or Delayed School Entry. *Journal of Applied School Psychology*, 26: 327-338.

Darling-Hammond, L. (1998). Alternatives to grade retention. *School Administrator*, 55: 18-21.

De Simone, G. (2013). Render unto primary the thing which are primary's: Inherited and fresh learning divides in Italian lower secondary education. *Economics of Education Review*, 35: 12-23.

De Witte, K., Nicaise, I., Lavrijsen, J., Van Landeghem, G., Lamote, C. and Van Damme, J. (2013). The Impact of Institutional Context, Education and Labour Market Policies on Early School Leaving: a comparative analysis of EU countries. *European Journal of Education*, 48: 331–345.

Dong, Y. (2009). Kept back to get ahead? Kindergarten retention and academic performance. *European Economic Review*, 54: 219–236.

Dupriez, V., Dumay, X., and Vause, A. (2008). How Do School Systems Manage Pupils' Heterogeneity? *Comparative Education Review*, 52: 245-273.

Eide, E., and Showalter, M. (2001). The effect of grade retention on educational and labor market outcomes. *Economics of Education Review*, 20: 563–576.

Escardíbul, J-O, and Villarroya, A. (2009). The inequalities in school choice in Spain in accordance to PISA data. *Journal of Education Policy*, 6: 673-696.

Fernández Enguita, M; Mena, L., and Riviere, J. (2010). *Fracaso y abandono escolar en España*. Barcelona: Fundació La Caixa.

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

Guio, J.M., Choi, A. (2014). Evolución del riesgo de fracaso escolar en España durante la década del 2000: Análisis de los resultados de PISA con un modelo logístico de dos niveles. *Estudios sobre Educación*, 26: 33-62.

Glick, P., Sahn, D. (2010). Early academic performance, grade repetition, and school attainment in Senegal: a panel data analysis. *World Bank Economic Review*, 24: 93–120.

Gomes-Neto, J.B., Hanushek, E.A. (1994). Causes and consequences of grade retention: evidence from Brazil. *Economic Development and Cultural Change*, 43: 117–148.

Goos, M., Schreier, B., Knipprath, H., De Fraine, B., Van Damme, J., and Trautwein, U. (2013). Can cross-country differences in the practice of grade retention be explained? A closer look at national educational policy factors. *Comparative Education Review*, 57: 54-84.

Holmes, T.C. (1989). Grade level retention effects: a meta-analysis of research studies. In: Shepard, Lorrie A., Smith, Mary L. (Eds.), *Flunking Grades: Research and Policies on Retention*, pp. 16–33, Falmer Press: London.

Hox, J. (1995). *Applied Multilevel Analysis.* Amsterdam: TT-Publikaties.

Jacob, B.A., and Lefgren, L. (2004). Remedial education and student achievement: a regression–discontinuity analysis. *The Review of Economics and Statistics*, 86, 226–244.

Jacob, B.A., and Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1: 33-58.

Jerrim, J., Choi, Á, and Simancas, R. (2014). Two-Sample Two-Stage Least Squares (TSTSLS) Estimates of Earnings Mobility: How consistent are they? *DoQSS Working Papers*, 17.

Jimerson, S.R., Anderson, G.E., and Whipple, A.D. (2002). Winning the battle and losing the war: examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, 39: 441-457.

Manacorda, M. (2012). The cost of grade retention. *The Review of Economics and Statistics*, 94: 596-606.

Mancebón, M.J., Calero, J., Choi, Á., and Ximénez-de-Embún, D. P. (2012). The Efficiency of Public and Publicly Subsidized High Schools in Spain: Evidence from PISA 2006. *Journal of the Operational Research Society*, 63: 1516-1533.

Martin, A.J. (2011). Holding back and holding behind: grade retention and students' non-academic and academic outcomes. *Educational Research Journal*, 37: 739-763.

MECD. (2014). *Las cifras de la educación en España 2014*. Madrid: MECD.

Mullis, I., Martin, M., Kennedy, A., and Pierre Foy (2007). *PIRLS 2006 International Report*. Chestnut Hill: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

OECD (2011). When students repeat grades or are transferred out of school: What does it mean to education systems? *PISA in Focus*, 2011/6.

OECD (2014). *PISA 2012 Results in Focus: What 15-year-olds know and what they can do with what they know*. Paris: OECD.

Ou, S.R., and Reynolds, A.J. (2010). Grade Retention, Postsecondary Education, and Public Aid Receipt. *Educational Evaluation and Policy Analysis*, 32:118-139.

Royston, P. (2005). Multiple imputation of missing values: update. *Stata Journal*, 5: 188-201.

Royston, P. and White, I. (2011). Multiple Imputation by Chained Equations (MICE): Implementation in STATA. *Journal of Statistical Software*, 45: 1-20.

Rubin, D. (1976). Inference and missing data, *Biometrika*, 63, 581-592.

Rubin, D. (1996). Multiple Imputation After 18+ Years. *Journal of the American Statistials Association,* 91: 473-489.

Salinas, J., and Santín D. (2012). Selección escolar y efectos de la inmigración sobre los resultados académicos españoles en PISA 2006. *Revista de Educación*, 358: 382-405.

StataCorp. (2013). *Stata Multiple Imputation: Reference Manual. Release 13*. Texas: Stata Press.

Schafer, J. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research,* 8, 3-15.

Sterne, J., White, I., Carlin, J., Spratt, M., Royston, P., Kenward, M., Wood, A. & Carpenter, J. (2009). Multiple imputation for missing data in epidemiological

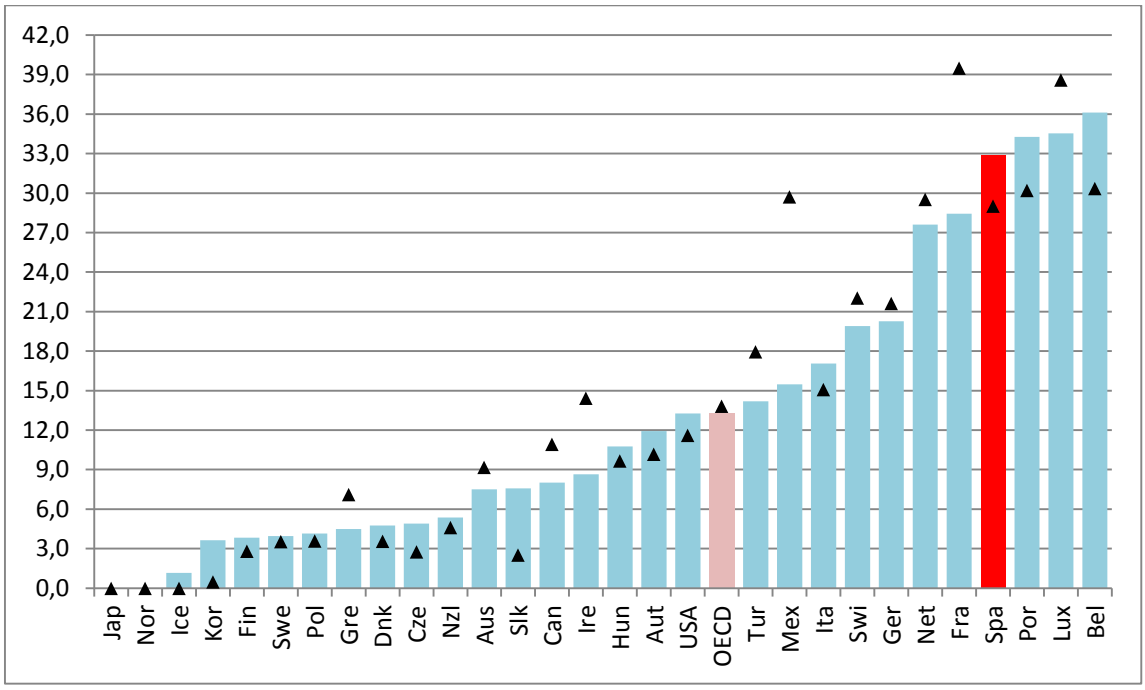
and clinical research: Potential and Pitfalls. *British Medical Journal*, 338(b2): 393.

Van Buuren, S. (2012). *Flexible Imputation of Missing Data*. Boca Raton: CRC Press.

Van Buuren, S., Boshuizen, H., and Knook, D. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18: 681-694.

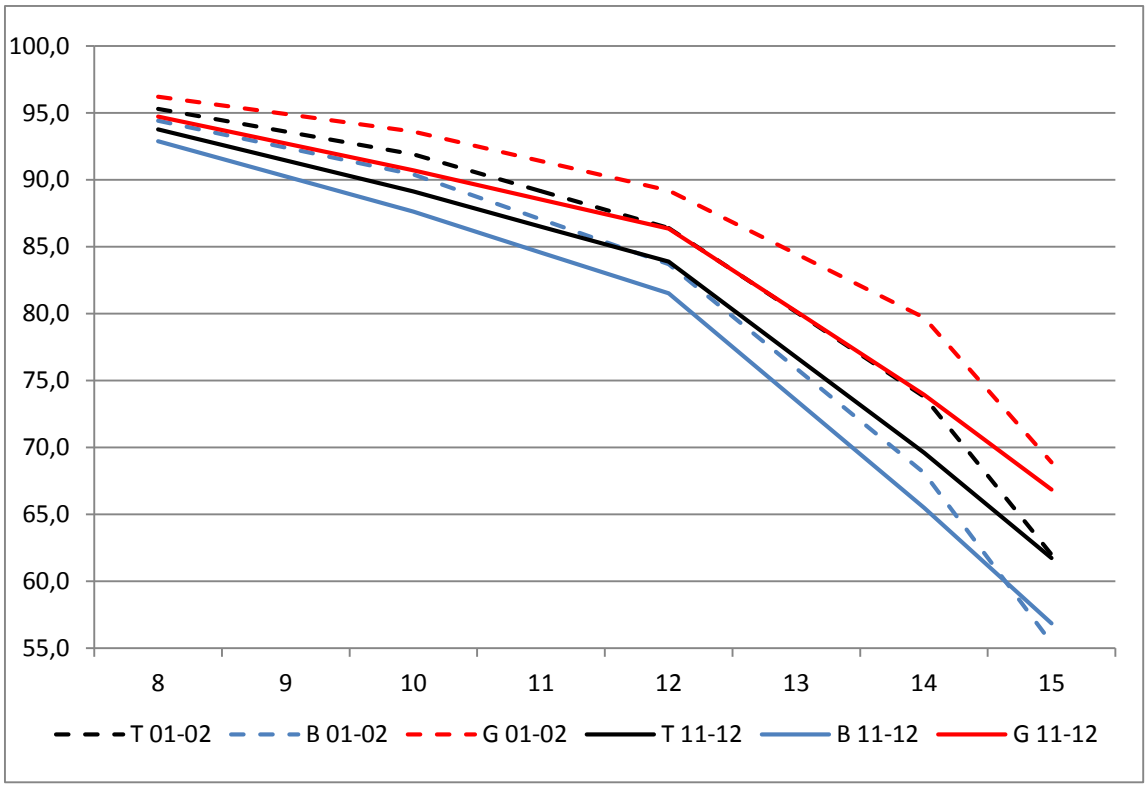
White, I., Royston, P., and Wood, A. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30: 377-399.

Figure 1. Grade retention at age 15 (%). OECD countries, Programme for International Student Assessment (PISA) 2012.



Source: Based on OECD data. Note: ▲, repetition rate in year 2003.

Figure 2. Students attending reference grade (%): evolution between primary and compulsory secondary education; academic years 2001/02 and 2011/12.



Source: Based on data from MECD (2014 –table C2). T: total; B: boys; G: girls.

Table 1. Main characteristics of repeaters and non-repeaters

	Non-repeater	Repeater		Non-repeater	Repeater
Mean PIRLS z_score*	0.42	0.25	% did not attend ISCED0*	4.38	7.16
Mean PISA z_score*	0.26	-0.52	% non-nuclear household*	8.62	14.51
% girls*	52.1	46.7	% immigrant background*	7.3	15.13
% born January to March	25.46	26.73	Mean ESCS (index)*	0.05	-0.56
% born October to Dec.	24.6	24.36	% public school*	61.42	76.78

Source: Based on PISA 2012 data. Note: * difference in means between repeaters and non-repeaters is statistically significant at the 95% level.

Table 2. Results from the hierarchical linear model; reading competency, PISA 2012.

	Coefficient	Standard E.
Intercept	-0.688***	0.181
A. Individual and household characteristics		
Repeated one year during lower secondary education	-0.405***	0.041
Repeated two or more years during lower secondary education	-0.650***	0.080
Interaction term: Repeated x first quartile in PIRLS score	-0.403***	0.066
Interaction term: Repeated x second quartile in PIRLS score	-0.257***	0.061
Interaction term: Repeated x third quartile in PIRLS score	-0.182***	0.056
Sex (girl=1)	0.202***	0.025
Born between April and September	0.125***	0.025
Born between October and December	-0.010	0.029
Attended ISCED0 one year	0.066	0.079
Attended ISCED0 more than one year	0.041	0.070
Single parent or other situation	0.099***	0.031
Ref. Two-parents in the household		
Immigrant household: first generation	0.181***	0.052
Immigrant household: second generation	-0.058	0.081
Ref. Non-immigrant household		
International language at home (ref. language of the test)	0.053	0.046
ESCS (Socio-economic status index)	0.214***	0.020
HEDRES (Home educational resources index)	0.011	0.011

Table 2. Results from the hierarchical linear model; reading competency, PISA 2012 (continued).

B. School characteristics		
School ownership: private	-0.039	0.031
(Ref. public school)		
School size (number of students)	-0.000	0.000
City size (100,000 to 1,000,000 inhabitants)	0.073**	0.036
City size (more than 1,000,000 inhabitants)	0.219***	0.036
(Ref. less than 100,000 inhabitants)		
Student-teacher ratio	0.004	0.005
Mean years of education of parents	0.018	0.013
Percentage of immigrant students >20%	0.045	0.054
Index of ICT school availability (ICTSCH)	-0.038***	0.012
Class size (number of students per class)	0.004	0.002
Responsibility of school in curriculum and assessments (index RESPCUR)	-0.039	0.029
Responsibility of school in allocating resources (index RESPRES)	0.103**	0.044
Variances	Null model	Complete model
Schools (μ_{0j})	0.128	0.060
Individuals (ε_{ij})	0.531	0.144
Total ($\mu_{0j} + \varepsilon_{ij}$)	0.659	0.204
% of the total variance explained by the variables		69,04%
% of the level 1 (students) variance explained by the variables		71,14%
% of the level 2 (schools) variance explained by the variables		53.13%

Source: Based on PISA 2012 and PIRLS 2006 data.

Note: ***, statistically significant at 99%; **, 95%; *, 90%. Regional dummies included.

Appendix 1. Summary statistics

Table A.1.1. PISA 2012: Missing Values Analysis

	Valid values	*Missing values*	*% Missing*
Repeated grade at ISCED 2	24,243	939	3.73
Repeated grade at ISCED 1	23,184	1,998	7.93
Immigrant background (index IMMIG)	24,696	486	1.93
Language spoken at home	24,558	624	2.48
Early childhood education (ISCED 0)	24,804	378	1.50
Family structure (index FAMSTRUC)	23,674	1,508	5.99
School type (index SCHLTYPE)	24,445	737	2.93
City Size	24,958	224	0.89
Average class size	22,161	3,021	12.00
Economic, social and cultural status (index ESCS)	24,991	191	0.76
Home educational resources (index HEDRES)	24,938	244	0.97
Student-teacher ratio	23,523	1,659	6.59
School size (index SCHSIZE)	23,848	1,334	5.30
Responsibility of school staff in issues relating to curriculum and assessment (index RESPCUR)	25,056	126	0.50
Responsibility of school staff in allocating resources (index RESPRES)	25,056	126	0.50
ICT availability at school (index ICTSCH)	24,286	896	3.56
PC at home	25,028	154	0.61
Highest educational level of parents	24,809	373	1.48
Parents highest occupation	24,882	300	1.19
Books at home	24,920	262	1.04
Father Current Job Status	24,031	1,151	4.57
Mother Current Job Status	24,600	582	2.31
Own room at home	25,026	156	0.62

Source: Based on PISA 2012 database. Sample size n= 25,182.

Table A.1.2. Summary statistics

	Mean	S.d.	Min.	Max	N
PIRLS 2006 reading z_scores: PV1	0.374	0.005	-0.991	1.076	20,437
PIRLS 2006 reading z_scores: PV2	0.289	0.005	-1.016	0.941	20,437
PIRLS 2006 reading z_scores: PV3	0.312	0.005	-1.055	0.989	20,437
PIRLS 2006 reading z_scores: PV4	0.323	0.005	-0.908	0.972	20,437
PIRLS 2006 reading z_scores: PV5	0.401	0.005	-0.870	1.058	20,437
PISA reading z_scores plausible value 1	0.144	0.020	-3.856	3.220	20,437
PISA reading z_scores plausible value 2	0.142	0.021	-3.733	3.038	20,437
PISA reading z_scores plausible value 3	0.143	0.020	-3.800	3.267	20,437
PISA reading z_scores plausible value 4	0.144	0.021	-3.972	3.121	20,437
PISA reading z_scores plausible value 5	0.140	0.020	-4.233	2.969	20,437
A. Individual and household characteristics					
Did not repeat during lower secondary education	0.847	0.360	0	1	20,306
Repeated one year during lower secondary education	0.136	0.343	0	1	20,306
Repeated two years during lower secondary education	0.172	0.130	0	1	20,306
Sex (girl)	0.512	0.006	0	1	20,437
Born between January and March	0.241	0.005	0	1	20,437
Born between April and September	0.495	0.006	0	1	20,437
Born between October and December	0.246	0.005	0	1	20,437
Did not attend ISCED0	0,048	0,003	0	1	20,285
Attended ISCED0 one year	0,069	0,003	0	1	20,285
Attended ISCED0 more than one year	0,883	0,005	0	1	20,285
Single parent or other situation	0.093	0.003	0	1	19,570
Non-immigrant household	0.914	0.280	0	1	20,234
Immigrant household: first generation	0.074	0.261	0	1	20,234
Immigrant household: second generation	0.012	0.108	0	1	20,234
Language at home: language of the test	0.813	0.016	0	1	20,102
Language at home: international language	0.187	0.016	0	1	20,102
ESCS (Socio-economic status index)	-0.054	0.030	-3.92	2.73	20,437
HEDRES (Home educational resources index)	0.122	0.014	-3.93	1.12	20,437

Table A.1.2. Summary statistics (continued)

B. School characteristics					
	Mean	**S.d.**	**Min.**	**Max**	**N**
Publicly-subsidised private school	0.652	0.027	0	1	19,789
Independent private school	0.266	0.025	0	1	19,789
Public school	0.082	0.015	0	1	19,789
School size (number of students)	722,795	23,47	45	4,128	19,350
City size (less than 100,000 inhabitants)	0.261	0.024	0	1	20,243
City size (100,000 to 1,000,000 inhabitants)	0.653	0.026	0	1	20,243
City size (more than 1,000,000 inhabitants)	0.086	0.016	0	1	20,243
Student-teacher ratio	12,623	0.342	1.111	139	19,077
Mean years of education of parents	12,570	1.965	5	16.5	20,437
Percentage of immigrant students >20%	0.144	0.019	0	1	20,437
Index of ICT school availability (ICTSCH)	-0.162	0.024	-2.804	2.826	19,895
Class size (number of students per class)	25,572	0.282	13	48	17,911
Responsibility of school in curriculum and assessments (index RESPCUR)	-0.436	0.044	-1.26	1.44	20,311
Responsibility of school in allocating resources (index RESPRES)	-0.394	0.039	-0.80	2.71	20,311
Regional dummy: Rest of the country	0.203	0.030	0	1	20,437
Regional dummy: Andalucía	0.194	0.024	0	1	20,437
Regional dummy: Aragón	0.026	0.004	0	1	20,437
Regional dummy: Asturias	0.020	0.003	0	1	20,437
Regional dummy: Baleares	0.021	0.003	0	1	20,437
Regional dummy: Cantabria	0.012	0.002	0	1	20,437
Regional dummy: Castilla y León	0.049	0.007	0	1	20,437
Regional dummy: Catalunya	0.169	0.022	0	1	20,437
Regional dummy: Extremadura	0.025	0.004	0	1	20,437
Regional dummy: Galicia	0.050	0.007	0	1	20,437
Regional dummy: La Rioja	0.007	0.001	0	1	20,437
Regional dummy: Madrid	0.132	0.018	0	1	20,437
Regional dummy: Murcia	0.031	0.005	0	1	20,437
Regional dummy: Navarra	0.015	0.002	0	1	20,437
Regional dummy: País Vasco	0.047	0.004	0	1	20,437

Source: Based on PISA 2012 and PIRLS 2006 data.

Appendix 2. Results of the analysis combining EGD-2007 and PISA 2012

We tested the robustness of our results using an alternative database to PIRLS 2006. For this purpose, we selected a national student assessment programme (*Evaluación General de Diagnóstico*, in Spanish, EGD hereafter), belonging to the multi-year General Assessment Plans of the Spanish educational system (since 1994/95). By analogy with our main study, we chose the data corresponding to 2007, thus ensuring that we follow a similar cohort to that in our main analysis.

The EGD is conducted with students in the last grade of primary school (12 years old, and so we selected students born in 1995) and aims to identify student competencies and knowledge at this educational level in four areas: the Natural, Social and Cultural Environment; Literature and Spanish Language; English Language; and Mathematics. The procedures used in collecting the data and treating the results are similar to those applied in other international assessments.

The sample is obtained by applying a stratified two-stage sampling: in the first stage (private and public) schools are randomly selected within a stratum (in this case the region and the school ownership model); in the second stage, one class is randomly chosen. Then, all students enrolled in this class make up the sample. Once the sample is selected, the students perform the standardised tests in each of the four educational areas. They also complete a questionnaire about their attitudes, and personal, social and school background, as do their parents, teachers and school principals. The response rate is very high, above 95% for the target populations.

The three databases differ in a number of aspects. For example, PISA focuses on competencies, PIRLS measures curricular content at the international level, while EDG assesses country specific curricular content. Furthermore, the PISA and PIRLS scores are standardised to z-international scores, taking into account the countries participating in each test, whereas the EDG is also standardised but, logically, it does not take into consideration any other countries. Finally, the EDG does not involve the calculation of plausible values in order to measuring its outcomes.

The procedures adopted in cleaning and merging the EGD database with the PISA database are the same as those described for PIRLS.

Table A.2.1. Results from the hierarchical linear model; reading competency, PISA 2012.

	Coefficient	Standard E.
Intercept	0.641***	0.174
A. Individual and household characteristics		
Repeated one year during lower secondary education	-0.220***	0.046
Repeated two or more years during lower secondary education	-0.510***	0.083
Interaction term: Repeated x first quartile in EGD-2007 score	-0.670***	0.070
Interaction term: Repeated x second quartile in EGD-2007 score	-0.521***	0.055
Interaction term: Repeated x third quartile in EGD-2007 score	-0.366***	0.055
Variances	Null model	Complete model
Schools (μ_{0j})	0.187	0.063
Individuals (ε_{ij})	0.697	0.456
Total ($\mu_{0j} + \varepsilon_{ij}$)	0.884	0.519
% of the total variance explained by the variables		41,30%
% of the level 1 (students) variance explained by the variables		34,58%
% of the level 2 (schools) variance explained by the variables		66.30%

Source: Based on PISA 2012 and EGD-2007 data.

Note: ***, statistically significant at 99%; **, 95%; *, 90%. Regional dummies and usual control variables included.

2012

2012/1, Montolio, D.; Trujillo, E.: "What drives investment in telecommunications? The role of regulation, firms' internationalization and market knowledge"
2012/2, Giesen, K.; Suedekum, J.: "The size distribution across all "cities": a unifying approach"
2012/3, Foremny, D.; Riedel, N.: "Business taxes and the electoral cycle"
2012/4, García-Estévez, J.; Duch-Brown, N.: "Student graduation: to what extent does university expenditure matter?"
2012/5, Durán-Cabré, J.M.; Esteller-Moré, A.; Salvadori, L.: "Empirical evidence on horizontal competition in tax enforcement"
2012/6, Pickering, A.C.; Rockey, J.: "Ideology and the growth of US state government"
2012/7, Vergolini, L.; Zanini, N.: "How does aid matter? The effect of financial aid on university enrolment decisions"
2012/8, Backus, P.: "Gibrat's law and legacy for non-profit organisations: a non-parametric analysis"
2012/9, Jofre-Monseny, J.; Marín-López, R.; Viladecans-Marsal, E.: "What underlies localization and urbanization economies? Evidence from the location of new firms"
2012/10, Mantovani, A.; Vandekerckhove, J.: "The strategic interplay between bundling and merging in complementary markets"
2012/11, Garcia-López, M.A.: "Urban spatial structure, suburbanization and transportation in Barcelona"
2012/12, Revelli, F.: "Business taxation and economic performance in hierarchical government structures"
2012/13, Arqué-Castells, P.; Mohnen, P.: "Sunk costs, extensive R&D subsidies and permanent inducement effects"
2012/14, Boffa, F.; Piolatto, A.; Ponzetto, G.: "Centralization and accountability: theory and evidence from the Clean Air Act"
2012/15, Cheshire, P.C.; Hilber, C.A.L.; Kaplanis, I.: "Land use regulation and productivity – land matters: evidence from a UK supermarket chain"
2012/16, Choi, A.; Calero, J.: "The contribution of the disabled to the attainment of the Europe 2020 strategy headline targets"
2012/17, Silva, J.I.; Vázquez-Grenno, J.: "The ins and outs of unemployment in a two-tier labor market"
2012/18, González-Val, R.; Lanaspa, L.; Sanz, F.: "New evidence on Gibrat's law for cities"
2012/19, Vázquez-Grenno, J.: "Job search methods in times of crisis: native and immigrant strategies in Spain"
2012/20, Lessmann, C.: "Regional inequality and decentralization – an empirical analysis"
2012/21, Nuevo-Chiquero, A.: "Trends in shotgun marriages: the pill, the will or the cost?"
2012/22, Piil Damm, A.: "Neighborhood quality and labor market outcomes: evidence from quasi-random neighborhood assignment of immigrants"
2012/23, Ploeckl, F.: "Space, settlements, towns: the influence of geography and market access on settlement distribution and urbanization"
2012/24, Algan, Y.; Hémet, C.; Laitin, D.: "Diversity and local public goods: a natural experiment with exogenous residential allocation"
2012/25, Martinez, D.; Sjögren, T.: "Vertical externalities with lump-sum taxes: how much difference does unemployment make?"
2012/26, Cubel, M.; Sanchez-Pages, S.: "The effect of within-group inequality in a conflict against a unitary threat"
2012/27, Andini, M.; De Blasio, G.; Duranton, G.; Strange, W.C.: "Marshallian labor market pooling: evidence from Italy"
2012/28, Solé-Ollé, A.; Viladecans-Marsal, E.: "Do political parties matter for local land use policies?"
2012/29, Buonanno, P.; Durante, R.; Prarolo, G.; Vanin, P.: "Poor institutions, rich mines: resource curse and the origins of the Sicilian mafia"
2012/30, Anghel, B.; Cabrales, A.; Carro, J.M.: "Evaluating a bilingual education program in Spain: the impact beyond foreign language learning"
2012/31, Curto-Grau, M.; Solé-Ollé, A.; Sorribas-Navarro, P.: "Partisan targeting of inter-governmental transfers & state interference in local elections: evidence from Spain"
2012/32, Kappeler, A.; Solé-Ollé, A.; Stephan, A.; Välilä, T.: "Does fiscal decentralization foster regional investment in productive infrastructure?"
2012/33, Rizzo, L.; Zanardi, A.: "Single vs double ballot and party coalitions: the impact on fiscal policy. Evidence from Italy"
2012/34, Ramachandran, R.: "Language use in education and primary schooling attainment: evidence from a natural experiment in Ethiopia"
2012/35, Rothstein, J.: "Teacher quality policy when supply matters"
2012/36, Ahlfeldt, G.M.: "The hidden dimensions of urbanity"
2012/37, Mora, T.; Gil, J.; Sicras-Mainar, A.: "The influence of BMI, obesity and overweight on medical costs: a panel data approach"
2012/38, Pelegrín, A.; García-Quevedo, J.: "Which firms are involved in foreign vertical integration?"

2012/39, Agasisti, T.; Longobardi, S.: "Inequality in education: can Italian disadvantaged students close the gap? A focus on resilience in the Italian school system"

2013

2013/1, Sánchez-Vidal, M.; González-Val, R.; Viladecans-Marsal, E.: "Sequential city growth in the US: does age matter?"
2013/2, Hortas Rico, M.: "Sprawl, blight and the role of urban containment policies. Evidence from US cities"
2013/3, Lampón, J.F.; Cabanelas-Lorenzo, P-; Lago-Peñas, S.: "Why firms relocate their production overseas? The answer lies inside: corporate, logistic and technological determinants"
2013/4, Montolio, D.; Planells, S.: "Does tourism boost criminal activity? Evidence from a top touristic country"
2013/5, Garcia-López, M.A.; Holl, A.; Viladecans-Marsal, E.: "Suburbanization and highways: when the Romans, the Bourbons and the first cars still shape Spanish cities"
2013/6, Bosch, N.; Espasa, M.; Montolio, D.: "Should large Spanish municipalities be financially compensated? Costs and benefits of being a capital/central municipality"
2013/7, Escardíbul, J.O.; Mora, T.: "Teacher gender and student performance in mathematics. Evidence from Catalonia"
2013/8, Arqué-Castells, P.; Viladecans-Marsal, E.: "Banking towards development: evidence from the Spanish banking expansion plan"
2013/9, Asensio, J.; Gómez-Lobo, A.; Matas, A.: "How effective are policies to reduce gasoline consumption? Evaluating a quasi-natural experiment in Spain"
2013/10, Jofre-Monseny, J.: "The effects of unemployment benefits on migration in lagging regions"
2013/11, Segarra, A.; García-Quevedo, J.; Teruel, M.: "Financial constraints and the failure of innovation projects"
2013/12, Jerrim, J.; Choi, A.: "The mathematics skills of school children: How does England compare to the high performing East Asian jurisdictions?"
2013/13, González-Val, R.; Tirado-Fabregat, D.A.; Viladecans-Marsal, E.: "Market potential and city growth: Spain 1860-1960"
2013/14, Lundqvist, H.: "Is it worth it? On the returns to holding political office"
2013/15, Ahlfeldt, G.M.; Maennig, W.: "Homevoters vs. leasevoters: a spatial analysis of airport effects"
2013/16, Lampón, J.F.; Lago-Peñas, S.: "Factors behind international relocation and changes in production geography in the European automobile components industry"
2013/17, Guío, J.M.; Choi, A.: "Evolution of the school failure risk during the 2000 decade in Spain: analysis of Pisa results with a two-level logistic mode"
2013/18, Dahlby, B.; Rodden, J.: "A political economy model of the vertical fiscal gap and vertical fiscal imbalances in a federation"
2013/19, Acacia, F.; Cubel, M.: "Strategic voting and happiness"
2013/20, Hellerstein, J.K.; Kutzbach, M.J.; Neumark, D.: "Do labor market networks have an important spatial dimension?"
2013/21, Pellegrino, G.; Savona, M.: "Is money all? Financing versus knowledge and demand constraints to innovation"
2013/22, Lin, J.: "Regional resilience"
2013/23, Costa-Campi, M.T.; Duch-Brown, N.; García-Quevedo, J.: "R&D drivers and obstacles to innovation in the energy industry"
2013/24, Huisman, R.; Stradnic, V.; Westgaard, S.: "Renewable energy and electricity prices: indirect empirical evidence from hydro power"
2013/25, Dargaud, E.; Mantovani, A.; Reggiani, C.: "The fight against cartels: a transatlantic perspective"
2013/26, Lambertini, L.; Mantovani, A.: "Feedback equilibria in a dynamic renewable resource oligopoly: pre-emption, voracity and exhaustion"
2013/27, Feld, L.P.; Kalb, A.; Moessinger, M.D.; Osterloh, S.: "Sovereign bond market reactions to fiscal rules and no-bailout clauses – the Swiss experience"
2013/28, Hilber, C.A.L.; Vermeulen, W.: "The impact of supply constraints on house prices in England"
2013/29, Revelli, F.: "Tax limits and local democracy"
2013/30, Wang, R.; Wang, W.: "Dress-up contest: a dark side of fiscal decentralization"
2013/31, Dargaud, E.; Mantovani, A.; Reggiani, C.: "The fight against cartels: a transatlantic perspective"
2013/32, Saarimaa, T.; Tukiainen, J.: "Local representation and strategic voting: evidence from electoral boundary reforms"
2013/33, Agasisti, T.; Murtinu, S.: "Are we wasting public money? No! The effects of grants on Italian university students' performances"
2013/34, Flacher, D.; Harari-Kermadec, H.; Moulin, L.: "Financing higher education: a contributory scheme"
2013/35, Carozzi, F.; Repetto, L.: "Sending the pork home: birth town bias in transfers to Italian municipalities"

2013/36, Coad, A.; Frankish, J.S.; Roberts, R.G.; Storey, D.J.: "New venture survival and growth: Does the fog lift?"
2013/37, Giulietti, M.; Grossi, L.; Waterson, M.: "Revenues from storage in a competitive electricity market: Empirical evidence from Great Britain"

2014

2014/1, Montolio, D.; Planells-Struse, S.: "When police patrols matter. The effect of police proximity on citizens' crime risk perception"
2014/2, Garcia-López, M.A.; Solé-Ollé, A.; Viladecans-Marsal, E.: "Do land use policies follow road construction?"
2014/3, Piolatto, A.; Rablen, M.D.: "Prospect theory and tax evasion: a reconsideration of the Yitzhaki puzzle"
2014/4, Cuberes, D.; González-Val, R.: "The effect of the Spanish Reconquest on Iberian Cities"
2014/5, Durán-Cabré, J.M.; Esteller-Moré, E.: "Tax professionals' view of the Spanish tax system: efficiency, equity and tax planning"
2014/6, Cubel, M.; Sanchez-Pages, S.: "Difference-form group contests"
2014/7, Del Rey, E.; Racionero, M.: "Choosing the type of income-contingent loan: risk-sharing versus risk-pooling"
2014/8, Torregrosa Hetland, S.: "A fiscal revolution? Progressivity in the Spanish tax system, 1960-1990"
2014/9, Piolatto, A.: "Itemised deductions: a device to reduce tax evasion"
2014/10, Costa, M.T.; García-Quevedo, J.; Segarra, A.: "Energy efficiency determinants: an empirical analysis of Spanish innovative firms"
2014/11, García-Quevedo, J.; Pellegrino, G.; Savona, M.: "Reviving demand-pull perspectives: the effect of demand uncertainty and stagnancy on R&D strategy"
2014/12, Calero, J.; Escardíbul, J.O.: "Barriers to non-formal professional training in Spain in periods of economic growth and crisis. An analysis with special attention to the effect of the previous human capital of workers"
2014/13, Cubel, M.; Sanchez-Pages, S.: "Gender differences and stereotypes in the beauty"
2014/14, Piolatto, A.; Schuett, F.: "Media competition and electoral politics"
2014/15, Montolio, D.; Trillas, F.; Trujillo-Baute, E.: "Regulatory environment and firm performance in EU telecommunications services"
2014/16, Lopez-Rodriguez, J.; Martinez, D.: "Beyond the R&D effects on innovation: the contribution of non-R&D activities to TFP growth in the EU"
2014/17, González-Val, R.: "Cross-sectional growth in US cities from 1990 to 2000"
2014/18, Vona, F.; Nicolli, F.: "Energy market liberalization and renewable energy policies in OECD countries"
2014/19, Curto-Grau, M.: "Voters' responsiveness to public employment policies"
2014/20, Duro, J.A.; Teixidó-Figueras, J.; Padilla, E.: "The causal factors of international inequality in co2 emissions per capita: a regression-based inequality decomposition analysis"
2014/21, Fleten, S.E.; Huisman, R.; Kilic, M.; Pennings, E.; Westgaard, S.: "Electricity futures prices: time varying sensitivity to fundamentals"
2014/22, Afcha, S.; García-Quevedo, J,: "The impact of R&D subsidies on R&D employment composition"
2014/23, Mir-Artigues, P.; del Río, P.: "Combining tariffs, investment subsidies and soft loans in a renewable electricity deployment policy"
2014/24, Romero-Jordán, D.; del Río, P.; Peñasco, C.: "Household electricity demand in Spanish regions. Public policy implications"
2014/25, Salinas, P.: "The effect of decentralization on educational outcomes: real autonomy matters!"
2014/26, Solé-Ollé, A.; Sorribas-Navarro, P.: "Does corruption erode trust in government? Evidence from a recent surge of local scandals in Spain"
2014/27, Costas-Pérez, E.: "Political corruption and voter turnout: mobilization or disaffection?"
2014/28, Cubel, M.; Nuevo-Chiquero, A.; Sanchez-Pages, S.; Vidal-Fernandez, M.: "Do personality traits affect productivity? Evidence from the LAB"
2014/29, Teresa Costa, M.T.; Trujillo-Baute, E.: "Retail price effects of feed-in tariff regulation"
2014/30, Kilic, M.; Trujillo-Baute, E.: "The stabilizing effect of hydro reservoir levels on intraday power prices under wind forecast errors"
2014/31, Costa-Campi, M.T.; Duch-Brown, N.: "The diffusion of patented oil and gas technology with environmental uses: a forward patent citation analysis"
2014/32, Ramos, R.; Sanromá, E.; Simón, H.: "Public-private sector wage differentials by type of contract: evidence from Spain"
2014/33, Backus, P.; Esteller-Moré, A.: "Is income redistribution a form of insurance, a public good or both?"
2014/34, Huisman, R.; Trujillo-Baute, E.: "Costs of power supply flexibility: the indirect impact of a Spanish policy change"

2014/35, Jerrim, J.; Choi, A.; Simancas Rodríguez, R.: "Two-sample two-stage least squares (TSTSLS) estimates of earnings mobility: how consistent are they?"
2014/36, Mantovani, A.; Tarola, O.; Vergari, C.: "Hedonic quality, social norms, and environmental campaigns"
2014/37, Ferraresi, M.; Galmarini, U.; Rizzo, L.: "Local infrastructures and externalities: Does the size matter?"
2014/38, Ferraresi, M.; Rizzo, L.; Zanardi, A.: "Policy outcomes of single and double-ballot elections"

2015

2015/1, Foremny, D.; Freier, R.; Moessinger, M-D.; Yeter, M.: "Overlapping political budget cycles in the legislative and the executive"
2015/2, Colombo, L.; Galmarini, U.: "Optimality and distortionary lobbying: regulating tobacco consumption"
2015/3, Pellegrino, G.: "Barriers to innovation: Can firm age help lower them?"
2015/4, Hémet, C.: "Diversity and employment prospects: neighbors matter!"
2015/5, Cubel, M.; Sanchez-Pages, S.: "An axiomatization of difference-form contest success functions"
2015/6, Choi, A.; Jerrim, J.: "The use (and misuse) of Pisa in guiding policy reform: the case of Spain"
2015/7, Durán-Cabré, J.M.; Esteller-Moré, A.; Salvadori, L.: "Empirical evidence on tax cooperation between sub-central administrations"
2015/8, Batalla-Bejerano, J.; Trujillo-Baute, E.: "Analysing the sensitivity of electricity system operational costs to deviations in supply and demand"
2015/9, Salvadori, L.: "Does tax enforcement counteract the negative effects of terrorism? A case study of the Basque Country"
2015/10, Montolio, D.; Planells-Struse, S.: "How time shapes crime: the temporal impacts of football matches on crime"
2015/11, Piolatto, A.: "Online booking and information: competition and welfare consequences of review aggregators"
2015/12, Boffa, F.; Pingali, V.; Sala, F.: "Strategic investment in merchant transmission: the impact of capacity utilization rules"
2015/13, Slemrod, J.: "Tax administration and tax systems"
2015/14, Arqué-Castells, P.; Cartaxo, R.M.; García-Quevedo, J.; Mira Godinho, M.: "How inventor royalty shares affect patenting and income in Portugal and Spain"
2015/15, Montolio, D.; Planells-Struse, S.: "Measuring the negative externalities of a private leisure activity: hooligans and pickpockets around the stadium"
2015/16, Batalla-Bejerano, J.; Costa-Campi, M.T.; Trujillo-Baute, E.: "Unexpected consequences of liberalisation: metering, losses, load profiles and cost settlement in Spain's electricity system"
2015/17, Batalla-Bejerano, J.; Trujillo-Baute, E.: "Impacts of intermittent renewable generation on electricity system costs"
2015/18, Costa-Campi, M.T.; Paniagua, J.; Trujillo-Baute, E.: "Are energy market integrations a green light for FDI?"
2015/19, Jofre-Monseny, J.; Sánchez-Vidal, M.; Viladecans-Marsal, E.: "Big plant closures and agglomeration economies"
2015/20, Garcia-López, M.A.; Hémet, C.; Viladecans-Marsal, E.: "How does transportation shape intrametropolitan growth? An answer from the regional express rail"
2015/21, Esteller-Moré, A.; Galmarini, U.; Rizzo, L.: "Fiscal equalization under political pressures"
2015/22, Escardíbul, J.O.; Afcha, S.: "Determinants of doctorate holders' job satisfaction. An analysis by employment sector and type of satisfaction in Spain"
2015/23, Aidt, T.; Asatryan, Z.; Badalyan, L.; Heinemann, F.: "Vote buying or (political) business (cycles) as usual?"
2015/24, Albæk, K.: "A test of the 'lose it or use it' hypothesis in labour markets around the world"
2015/25, Angelucci, C.; Russo, A.: "Petty corruption and citizen feedback"
2015/26, Moriconi, S.; Picard, P.M.; Zanaj, S.: "Commodity taxation and regulatory competition"
2015/27, Brekke, K.R.; Garcia Pires, A.J.; Schindler, D.; Schjelderup, G.: "Capital taxation and imperfect competition: ACE vs. CBIT"
2015/28, Redonda, A.: "Market structure, the functional form of demand and the sensitivity of the vertical reaction function"
2015/29, Ramos, R.; Sanromá, E.; Simón, H.: "An analysis of wage differentials between full-and part-time workers in Spain"
2015/30, Garcia-López, M.A.; Pasidis, I.; Viladecans-Marsal, E.: "Express delivery to the suburbs the effects of transportation in Europe's heterogeneous cities"
2015/31, Torregrosa, S.: "Bypassing progressive taxation: fraud and base erosion in the Spanish income tax (1970-2001)"

2015/32, Choi, H.; Choi, A.: "When one door closes: the impact of the hagwon curfew on the consumption of private tutoring in the republic of Korea"
2015/33, Escardíbul, J.O.; Helmy, N.: "Decentralisation and school autonomy impact on the quality of education: the case of two MENA countries"
2015/34, González-Val, R.; Marcén, M.: "Divorce and the business cycle: a cross-country analysis"
2015/35, Calero, J.; Choi, A.: "The distribution of skills among the European adult population and unemployment: a comparative approach"
2015/36, Mediavilla, M.; Zancajo, A.: "Is there real freedom of school choice? An analysis from Chile"
2015/37, Daniele, G.: "Strike one to educate one hundred: organized crime, political selection and politicians' ability"
2015/38, González-Val, R.; Marcén, M.: "Regional unemployment, marriage, and divorce"
2015/39, Foremny, D.; Jofre-Monseny, J.; Solé-Ollé, A.: "'Hold that ghost': using notches to identify manipulation of population-based grants"
2015/40, Mancebón, M.J.; Ximénez-de-Embún, D.P.; Mediavilla, M.; Gómez-Sancho, J.M.: "Does educational management model matter? New evidence for Spain by a quasiexperimental approach"
2015/41, Daniele, G.; Geys, B.: "Exposing politicians' ties to criminal organizations: the effects of local government dissolutions on electoral outcomes in Southern Italian municipalities"
2015/42, Ooghe, E.: "Wage policies, employment, and redistributive efficiency"

2016

2016/1, Galletta, S.: "Law enforcement, municipal budgets and spillover effects: evidence from a quasi-experiment in Italy"
2016/2, Flatley, L.; Giulietti, M.; Grossi, L.; Trujillo-Baute, E.; Waterson, M.: "Analysing the potential economic value of energy storage"
2016/3, Calero, J.; Murillo Huertas, I.P.; Raymond Bara, J.L.: "Education, age and skills: an analysis using the PIAAC survey"
2016/4, Costa-Campi, M.T.; Daví-Arderius, D.; Trujillo-Baute, E.: "The economic impact of electricity losses"
2016/5, Falck, O.; Heimisch, A.; Wiederhold, S.: "Returns to ICT skills"
2016/6, Halmenschlager, C.; Mantovani, A.: "On the private and social desirability of mixed bundling in complementary markets with cost savings"

UNIVERSITAT DE BARCELONA

ieb@ub.edu
www.ieb.ub.edu

Human Capital