The determinants of repetition rates in European countries: insights from an empirical analysis using PIRLS 2011 data

Tommaso Agasisti[1] and José Manuel Cordero[2,*]

Abstract
In this paper, we examine the factors that are associated with the probability for a student to be retained in primary school; understanding such phenomenon is very important, because the literature points at claiming that retention perpetrates socioeconomic and educational inequality, instead of reducing it. Empirically, we estimate a multivariate logistic regression with the aim of identifying cross-country determinants of grade retention using data about 16 European countries participating in PIRLS 2011 using data about entry age at school and grade enrollment to identify retained students. Our results highlight that socioeconomic background is not the only factor that plays a role. Early-acquired skills do reduce the probability for a child to become a repeater; thus, the policy-makers should devote their effort in helping families to undertake preschool activities that develop skills also before starting formal schooling.

Keywords: Retention, Educational attainment, Cross-country analysis, Multilevel regression.
JEL Code: I21

* Corresponding author
[2] Universidad de Extremadura (Spain), Department of Economics, Faculty of Economics and Business Studies. T. +34 924 289300, Ext. 86518
jmcordero@unex.es

[1]Politecnico di Milano (Italy), School of Management. T. +39.02.2399.3963.
tommaso.agasisti@polimi.it

1. Introduction

The analysis of grade retention has a long tradition within the field of educational research. In the first stages of the educational system, this strategy is implemented when pupils show a certain level of immaturity or social behavior problems that can affect their learning process (Jimerson & Ferguson, 2007). Likewise, other aspects can also been considered as the existence of family or health problems that derive into school absenteeism. The advocates of retention argue that this intervention grants children more time in the same course to mature and acquire some basic skills to prevent failure and frustration in future grades where they will have to face more advanced learning tasks (Alexander *et al.*, 2003; Hong & Yu, 2007). Moreover, this policy can result in more academically homogeneous classes, which facilitates teacher instruction (Hong & Raudenbush, 2005). These arguments can explain why retention has been a popular practice for decades; also, the interest of academic research about this practice is justified by the possibility to understand more factors associated with higher/lower level of student achievement.

However, the great amount of research conducted on this issue has failed to demonstrate that grade retention provides greater benefits to struggling students than promoting them to the next grade (Jimerson *et al.*, 2002). In fact, in many cases the effects of this practice is just the opposite, since students who experience this situation are more prone to school failure, especially if that strategy is applied in primary education when individual differences in learning rates have long-term consequences (Rumberger, 1995; Ou & Reynolds, 2010). This is particularly worrying if we take into account that retention rates depend significantly on socioeconomic factors (Corman, 2003), thus it can contribute to increase inequality within an educational system.

In this paper we are especially interested in examining the causes behind the decision of retaining a student in the early stages of the education process, which is usually adopted by school authorities following the recommendations made by teachers and, in some cases, considering parents' opinions. Nevertheless, it is noteworthy that the implementation of this strategy depends on the context and traditions of each nation, thus there is a considerable variation in grade retention rates across countries. In this sense, it is possible to find different models to deal with students' early difficulties and academic heterogeneity (Dupriez *et al.*, 2008); these policy-based and managerial

choices have clear effects on the educational experience (and performance) of students as a whole. In some countries (e.g. Belgium, France, Netherlands or Spain) grade retention is a common practice for students who do not reach the minimum levels of attainment established, whereas in other countries this measure is either not allowed or not applied, because they use differentiation and individual teaching (e.g. Norway or Finland) or separate children into distinctive ability groups at an early age (e.g., the United Kingdom). Eisenmon (1997) also reported that repetition rates in developing countries are often high, especially in rural areas.

We use a multivariate approach with the aim of identifying cross-country determinants of grade retention using data about 16 European countries participating in the Progress in International Reading Literacy Study (herafter, PIRLS) in 2011. This project, conducted by the International Association for the Evaluation of Educational Achievement (IEA), comprises data about students' reading achievement after four years of primary schooling. Therefore, our focus is placed on the practice of retention in first stages of the learning process that takes place in schools, which deserve a careful scrutiny, because they can have a greater impact for children's long-term educational attainment (Ikeda & Garcia, 2013).

The empirical strategy used in this paper is similar to the one employed recently by Goos *et al.* (2013a) to analyze international variation in grade retention rates across 34 OECD countries participating in PISA 2009. However, there are two main differences between our approach and the one used in that study. First, those authors are interested in identifying which national educational policy factors explain the divergences in the probability of students repeating a grade, while our goal is exactly the opposite, i.e., we search for common factors that can explain the phenomenon of retention in all the analyzed countries. Second, PISA dataset provides explicit information about whether the student has repeated one or more courses and when (primary or secondary education), since there is a specific question about this issue in the student questionnaire. Unfortunately, this question is not included in the questionnaire completed by students participating in PIRLS, although it is possible to identify retained students by combining the available data about the exact age of the student at the moment of taking the test (years, months and days) and the existing criteria about the

age of admissions in schools in each country[1]. This is not straightforward, thus this imputation procedure constitutes one of the main contributions of the paper, since we develop a method that allows us to distinguish between retained and non-retained students despite this information is not directly available in the original dataset. For that purpose, we use data about the grade retention regulation provided by Eurydice (2011). Given that this information is only available for European countries, our empirical study has been restricted to them.

To anticipate our main results, they show that, despite some structural differences across countries' educational policies that determine the percentage of retained students in each of them, there are some factors that are statistically correlated with the probability for a student to become a repeater. More specifically, these factors are related to the student's socioeconomic background; the higher it is, the lower the probability to be retained, all else equal. Also, attending schools in which students are on average better-off reduces the probability to incur in repetition. Lastly, and mostly important on a policy ground, acquiring skills and competencies through educational activities before schooling (such as learning letters or words, etc.) is beneficial for students – indeed, the probability of becoming repeaters decreases in a statistically significant way.

The remainder of the paper is organised as follows. In Section 2 we present a brief literature review about previous studies related to grade retention policies. In Section 3, we describe our data and variables, the innovative strategy used to identify retained students in each country and the methodology used to perform our analysis. In Section 4, we present the main results. Finally, in Section 5 we discuss those results and their main policy implications.

2. Literature review

The strategy of grade retention has received a lot of attention mainly in the field of education, sociology and psychology, although more recently it has also become the focus of economic research due to the extra costs that it originates for educational

[1] A closely related literature examines the effects of school start age on performance (Bedard & Dhuey, 2006; Deming & Dinarski, 2008; Elder & Lubotsky, 2009). Although it is possible to find some similarities between the effects of both practices, in this paper we focus on grade retention policies.

systems. Those costs include the expenditure of providing the retained student with an additional year of education. For example, the average annual expenditure per student in the United States is around $11,000 (Ikeda, 2011), thus estimates of the cost of that an additional year of schooling becomes staggering, regardless of whether retention rates are placed between 5% or 20% (Aud *et al.*, 2012).

There is a vast literature related to the existing duality between the practice of promoting students and the alternative of retaining them in order to determine what strategy is more favorable (Bali *et al.*, 2005). The revision of this literature is beyond the scope of this work, although the main conclusion that can be drawn from different meta-analyses (Holmes, 1989; Jimerson, 2001) and literature reviews (Jimerson *et al.*, 2002, Xia & Kirby, 2009; Huddleston, 2014) about this topic of research is that retention does not appear to benefit students academically and, in most cases, has a negative effect on achievement. Indeed, retention can be considered as one of the most powerful predictors of dropping out of school later on in the academic career (Roderick, 1994; Eide & Showalter, 2001)[2], and this is the culminating event of a gradual process of disconnection from school, including lack of motivation, low effort and absence in classes (Fredericks *et al.*, 2004). Moreover, some studies suggest that this practice can also have negative effect on parent educational expectations (Hughes *et al.*, 2013) and cause social, emotional, attitudinal and behavioral problems (McCoy & Reynolds, 1999; Pagani *et al.*, 2001).

However, some recent works have pointed out the limitations of those earlier works because they suffer from important methodological shortcomings (Lorence, 2006; Wu *et al.*, 2008; Allen *et al.*, 2009; Reschly & Christenson, 2013). The main drawback comes from the existence of an endogeneity problem, since the characteristics of the pupils, such as ability or motivation, as well as their family and school environment are likely to affect simultaneously the possibility of grade failure and also academic attainment. The common structure of more recent studies focused on analyzing the impact of retention policies on students outcomes are based on the use of longitudinal data, which allows for monitoring them through different stages of the educational, and rigorous statistical methods to deal with selection bias (See Goos *et al.*, 2013b for details). The conclusions of some of these studies suggest that there may be some initial

[2] The empirical results of different studies show that retained students have between two and eleven more probabilities to drop out later in school than those who are not retained.

positive effects on achievement for retained students in the short term, although these effects fade fairly quickly (Jacob & Lefgren, 2004, 2009; Dong, 2010; Im *et al.*, 2013).

In all the former analysis, grade retention is considered as a predictor for different outcomes such as test scores, dropping out rates or behavioral skills. However, there is also a different body of research focused on studying the determinants of grade retention (Ferguson *et al.*, 2001; Willson & Hughes, 2009). In all these studies there is a common factor related to the probability of repeating a grade, the low academic achievement, although there are also other influential variables such as male gender, living in poverty or coming from Hispanic and Black racial/ethnic backgrounds (Frey, 2005), as well as little parental involvement or the lack of discipline at home (Greene & Winters, 2007). In addition to these individual factors, some school variables can also being identified as predictors of the possibility of being retained, among which large classes, frequent teacher absence or an unfavorable environment in the class are the most significant (Creemers & Kyriakides, 2008).

Despite the prevalence of grade retention practices in many countries, the majority of the aforementioned research studies on this topic has been conducted in the United States, where grade retention is only used as the last remedy in primary education. Therefore, the extension of those results to other countries is not enough clear, since the educational context can be very different in other frameworks (especially in Europe) where grade retention can be considered as a more common practice, such as in The Netherlands, Belgium, France or Spain, or it is practically never applied (e.g. Scandinavian countries).

Among the studies investigating the main predictors of the probability of being retained in other nations, most empirical studies are based on data collected in a particular setting, including both developed and developing countries[3]. Nevertheless, the examinations of the causes behind the decision about grade retention using a cross-country approach are still scarce in the literature. This type of analysis is complex because each country has a singular student population with its own characteristics as

[3] Some examples are the empirical analysis performed for Brazil (Gomes-Neto & Hanushek, 1994), Bolivia and Guatemala (Patrinos & Psacharopoulos, 1996), Lebanon (El-Hassan, 1998); Colombia (Angrist *et al.*, 2002), Canada (Guèvremont *et al*., 2007), Switzerland (Bonvin *et al*., 2008), China (Chen *et al*., 2010), Germany (Ehmke *et al*., 2010), Netherlands (Kloosterman & de Graaf), Australia (Martin, 2011), Uruguay (Manacorda, 2012), Belgium (Goos *et al*., 2013b; Belot & Vandenberghe, 2014) or Spain (Cordero *et al.*, 2014; García-Pérez *et al*., 2014).

well as different educational contexts, thus the interaction between them produce diverse and stratified educational outcomes within countries (Dalton, 2012).

Recently, Goos *et al.* (2013a) provide a good analytical review about some previous studies adopting this approach. All of them present a similar structure based on the utilization of regression linear models (multilevel in some cases) to examine the effects of family characteristics as well as school resources or policies on grade retention rates using data from different countries. However, from our point of view, all of them present a significant limitation that makes us question the validity of the results obtained. For example, Lee & Barro (2001) and Le Donné (2014) use data from international achievement tests aggregated at country level, which it is often interpreted as a potential bias for the estimates. Willms & Somers (2001) used a definition for grade retention (average time a student needs in order to complete a grade during the first three years of primary education) that might be open to discussion. Dupriez *et al.* (2008) classifies countries into different typologies, according to the way they deal with heterogeneity among students, before conducting their empirical analysis, thus the results obtained might be conditioned by this previous distinction.

The empirical work carried out by Goos *et al.* (2013a) using PISA 2009 data shares the same structure (it is based on a hierarchical linear regression model), but it overcomes all the aforementioned drawbacks. First, it uses data at different levels (student, school and country). Second, it identifies retained students according to their own responses about this specific issue. Third, the analysis is conducted for the whole sample (34 countries) without making any previous classification. The main finding of this work is that national educational policy factors account for a significant portion (around 20-25%) of the variation in grade retention rates across countries.

This study in inserted in the second stream of contributions mentioned in this section, namely those that investigate the factors associated with the probability for a student to be retained; more specifically, the paper aims to extending this literature based on cross-country comparison by using data retrieved from a different dataset, PIRLS 2011, which incorporates some additional information about students' abilities before starting the school, that might have a great impact on early school performance and, therefore, on the probability of students repeating a grade (Elder & Lubotsky, 2009). Likewise, we are also interested in exploring whether the divergences detected across countries in

grade retention rates in previous studies, in which data refer to fifteen-year old students, still remain when the evaluated students are fourth graders.

3. Methods

3.1. Dataset

In this study we use data from European countries participating in PIRLS 2011. This dataset provides international comparative data about students' reading achievement in the fourth year of primary schooling (see Mullis *et al.*, 2012 for details). Besides this academic content, PIRLS also collects a rich array of background information about the students' background, the school environment and instructional practices. This information comes from the responses given to different questionnaires completed by students, parents, teachers and school principals.

One of the main advantages of using this dataset is the information provided by parents and teachers, which provides us with additional analytic power for gaining an in-depth understanding of student s' performance. In particular, PIRLS includes more detailed data about students' familiar background as well as habits in early childhood that might have influence on their future attainment. This information, which is not available in other international databases like PISA, may be particularly interesting in the context of our study, since we are interested in identifying common features in the profile of retained students across different countries.

The main problem that arises when using PIRLS to address this issue is that the dataset does not include a specific question about whether students have ever repeated a grade during primary education. However, this limitation has not diminished our intention of conducting our research, but rather the identification of retained students has become one of the main challenges of this research. This identification has been possible because PIRLS is a grade-based sample, i.e., the survey is administered to students enrolled in the same course (fourth year of primary education)[4], so students who are older than their classmates will be, in principle, retained students. A similar argument was also adopted by Corman (2003) to examine statewide educational policies' effects on grade repetition in the United States. In our case, this is a more difficult task, since the admission criteria to enter the schools varies across countries, so we need to develop

[4] This is different in PISA, because all the students are born in the same year.

an approach to recognize adequately whether a student has been previously retained or not. This approach will be based on combining the information about students' exact birth date, the date of testing and the age of admission in the school[5].

For that purpose, we rely on data collected by the EACEA in a recent report about the existing regulation concerning the admission criteria for compulsory primary education and grade promotion policies (Eurydice 2011)[6]. As we assume that the reliability of our results depends on the accuracy of this imputation process, in the next section 3.2 we provide a detailed explanation about the precise method employed to distinguish between retained and non-retained students in different countries.

3.2. The identification of retained students

The report on "*Grade retention during Compulsory Education in Europe: Regulation and Statistics*" prepared by Eurydice (2011) includes an extensive volume of data about the regulation of retention policies in European countries[7]. That is the main reason why we only consider countries from this continent in our empirical analysis, despite PIRLS 2011 collects data about 45 countries around the world.

Our main focus of attention is the official age for starting compulsory primary education in different countries and, even most importantly, the regulation about the specific date or period in the year by which the child must have reached the required age to start in the school. In most of countries, pupils start the school when they are six, although in some countries they can start earlier (at the age of four in Ireland or five in England and Netherlands) or later (at the age of seven in Bulgaria, Denmark, Finland and Sweden). With regard to the former group of countries, the variation across them is actually greater. In some cases, children enrolled in primary education do not need to fulfill the requirement of age when they arrive to the school for the first time, but they must have reached it before the end of the calendar year. This group of countries comprises Belgium, Bulgaria, Denmark, Finland, France, Hungary, Italy, Netherlands, Norway, Poland, Slovenia, Spain and Sweden. In other countries, the regulation establishes that

[5] McEwan & Shapiro (2008) also used the exact birth dates to analyze the effect of delaying school enrollment on educational outcomes.

[6] The Education, Audiovisual and Culture Executive Agency (EACEA) is responsible for the management of programs funded by the European Commision in the field of education.

[7] Those countries are Austria, Belgium, Bulgaria, Czech Republic, Denmark, England, Finland, France, Germany, Hungary, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Romania, Slovak Republic, Slovenia, Spain and Sweden.

students need to reach the age of admission before a specific date, which usually corresponds to the start of the school year in September. This is the case in Austria, Czech Republic, England, Portugal, Romania and Slovak Republic. In Germany, the cut-off date established by the law is September 30, while in Ireland the students are admitted in primary schools if they reach the age of six before July 1.

In the first group the identification of retained students is quite simple, since we only need to check the information about the year of birth. Therefore, in most of them the retained students will be those who were born before 2001. Bulgaria and the three Nordic countries, where the students started primary education one year later, represent the exceptions, thus the category of retained students is represented by those who were born before the year 2000. In the second group, we consider as retained students those who were born before September 1 in 2000, although this date was modified for German and Irish students according to the criteria mentioned above.

However, the classification of retained students derived from the application of this basic criterion is not totally reliable in some countries. For instance, in Germany the reference period may be extended until the end of the year by some regional authorities (länder), while in other countries such as Austria, Czech Republic, Denmark or the Netherlands, the period of admission can be modified for some children if their parents make the appropriate request and there is enough evidence that they are sufficiently mature to attend school. To avoid potential misidentification of retained students in those four countries, we decided to exclude them from our analysis. As a result, our sample includes data about more than 76,000 students from 16 European countries. Table 1 reports the percentage of retained and non-retained students in each country considered in the present study, according to the criterion established.

(*Table 1 around here*)

In order to test the reliability of our classification for retained students, we compare the average values recorded using our imputation method with statistical data provided by PISA 2009, where there is a specific question about whether the student has repeated a year at primary education (ISCED level 1), and the estimation calculated by Eurydice (2011) using data from the Eurostat database[8]. According to the values reported in Table 2, our classification seen to be consistent with those data. In fact, the correlation

[8] See Eurydice (2011) pp. 34-35 for details.

coefficients (Pearson and Spearman) between them are higher than 0.8 for all the combinations.

(*Table 2 around here*)

In Table 3 we also show the existing differences in reading achievement between students labeled as repeaters and non-repeaters in each country, which are substantial in most cases. The largest gap between the two groups can be found in Eastern countries such as the Slovak Republic, Poland, Bulgaria, Romania and Hungary, where the difference is higher than 70 points. In contrast, the differences are relatively slight in Ireland, Netherlands and Denmark, although the existence of an average difference of 33 points between the two groups for the whole sample provides us with an additional argument to support the validity of our classification.

(*Table 3 around here*)

3.3. Explanatory variables

Once we have defined the dependent variable, which classifies students into two different categories (repeaters and non-repeaters), our aim is to identify common factors related to the students´ and school background that might be associated with the possibility of being retained during primary education. Likewise, we are also interested in exploring whether the existing variation in grade retention rates across countries is significant or not once we have taken into account individual and school variables.

With regard to variables at the student level, it is important to highlight that, apart from testing the influence of the usual variables considered in the literature (e.g. gender, mother's level of education or possessions at home)[9], we also pay attention to some variables related to the acquisition of certain abilities in the earliest stages of life. This interest in based on recent evidence about the importance of these skills on explaining differences in educational attainment, and takes into account the major role that families play in shaping those skills (Cunha *et al*, 2010). In particular, we are concerned about whether the acquisition (or not) of some specific skills at home before starting the

[9] Todd & Wolpin (2003) survey the educational production function literature. Some recent references about these particular issues can be found in Else-Quest et al. (2010) for gender effects, Carneiro *et al.* (2013) for mothers´ education or Evans et al. (2010) for books at home.

school or the enrollment (or not) in preschool courses are associated (or not) with the probability of repeating a grade during primary education.

In our empirical exercise, we include two different sets of covariates that should, in principle, have an opposite effect on the probability of being retained. Firstly, we define a set of variables that reflect positive characteristics, so it can be expected that they reduce the probability of being a repeater. Subsequently, we use a different criterion to define another group of variables representing disadvantaging features, so they are likely to increase this probability.

In addition to individual variables, we also examine the potential influence of the presence of some educational resources together with some indicators related to the school environment such as socioeconomic background of the students attending the school or the frequency of disturbance problems during classes. The detailed definition of all the variables included in the analysis is reported in Table 4.

(*Table 4 around here*)

Finally, we also investigate the issue raised by Eurydice (2011) and Goos *et al*. (2012) about whether there exist significant differences among countries in the application of grade retention policies by including country fixed-effects (dummies).

Table 5 reports the descriptive statistics for individual and school variables[10]. Given that all the variables are binary, mean values can be interpreted as proportions. According to this information it is possible to observe that only 8.2% of students have repeated a grade in primary school. Our sample of individuals is almost evenly distributed by gender, the language spoken at home not always coincides with the one used in the test (25% of cases) and the percentage of students´ mothers who only finished lower secondary education is higher than the proportion holding an university degree. Most students has a computer, an own room and access to internet, although the proportion of students with a shortage of books at home is higher than those who has a great library. The assistance to preschool is widespread, although there is a great heterogeneity with regard to the abilities acquired before starting the school.

[10] The information provided by descriptive statistics displayed by countries can also be helpful to interpret the results of our analysis. However, we decided not to include them because of space limitations. Nevertheless, they are available upon request.

The mean values of the school variables allow us to detect that there are few schools operating in a harsh environment, although the proportion of schools placed in a high-income area is relatively low. The problems derived from absenteeism seem to be less important than those related to class disturbance. Finally, there is an almost insignificant percentage of schools where there are not computers available for instruction, but this proportion is relatively important for the case of libraries.

(Table 5 around here)

3.4. Data analysis

The model used in the empirical analysis is a hierarchical linear regression (Bryk and Raudenbush, 1992; Goldstein, 1995), which considers that the students are grouped (nested) at a higher level, represented by the schools. The use of this multilevel approach allows us to avoid potential problems of bias in the estimations derived from classic methods, such as OLS regression, due to the existence of correlation between the values of the school variables of pupils from the same school (Hox, 2002). Since the dependent variable is categorical, these regressions assume a binomial logistic model structure.

In this paper, we adopt a two-level approach in which the binary dependent variable to be estimated is the probability that the student "*i*" from school "*j*" is included within the corresponding group of retakers: $P(Y_{ij}=1|\beta)=P_{ij}$. This probability can be modeled using the following logistic function:

$$\log\left[\frac{P_{ij}}{(1-P_{ij})}\right]=\beta_{0j}+\beta_{ij}X_{ij}+r_{ij} \qquad (1)$$

$$\beta_{0j}=\gamma_{00}+\gamma_{01}Z_j+u_{0j}$$

$$\beta_{1j}=\gamma_{10}+u_1$$

In this equation, the probability that the student meets the established requirement depends on a vector of independent variables at the individual level (X_{ij}) and a vector of school variables (Z_j), but also takes into account the deviation of the school *j* (u_j) with respect to the results of all the schools (γ_0) and the deviation of the student *i* with respect to the average of the results obtained by students who belong to the same school *j*.

The values of the estimated coefficients in the model cannot be interpreted directly as in a linear regression, thus we need to estimate the odds ratios of the independent variable, which represent the relationship between the probability of an event happening in one group and the probability of it happening in the other group. Therefore, the odds ratios associated with an explanatory variable will take a value greater than one if that variable increases the probability that a student has been retained and less than one if that variable decreases the probability of such an event. The former is associated with positive coefficients while the latter presents negative coefficients.

4. Results

In this section we present the results obtained by applying the multilevel logistic regression model explained above to our dataset. All statistical analyses were performed considering sample weightings in the estimations to ensure that sampled students adequately represent the analyzed total population (Rutkowski *et al.*, 2010).

We first perform various versions of equation (1) for each set of variables (positive and negative), by including or excluding various variables – even though maintaining variables belonging to different categories, and especially student-level and school-level variables. At the end, we report the results that maximize the explanatory power of the overall models, although the complete sets of results are available from the authors. The results of the empirical analysis are reported in Table 6.

(*Table 6 around here*)

While the two general models provide results that are qualitatively and quantitatively similar, we start by discussing those from the model with variables that measures factors and characteristics, which are expected to be associated with a lower probability of being a repeater.

Females are less likely to be repeaters (odds ratio = 0.75), as are those students whose language spoken at home is the same of the test (odds ratio = 0.73), this latter variables capturing the status of immigrant student or belonging to a minority within the country. The most important factor reducing the probability to be a repeater is the mother's education: the students whose mother is highly educated (i.e. university degree or more) are 0.37 times less likely to be repeaters.

Also, experiences before schooling play an important role: this is measured both by the variable about having attended a preschool of at least three years (odds ratio: 0.64) and by the information about the skills before schooling, acquired in the family and at the kindergarten (ability of reading letters, odds ratio = 0.49; ability of writing words, odds ratio = 0.74).

As evidenced by a huge literature, the effect of her/his family's socioeconomic background is a major determinant of a student's achievement (Haveman & Wolfe, 1995). In this paper, we use as proxies for this measure some variables reflecting various dimensions as the cultural status (number of books at home > 200, odds ratio = 0.82) and the home possessions, such as "having a PC at home" (odds ratio = 0.69) and "having a single room" (odds ratio = 0.85). Overall, these variables indicate that positive conditions at home for studying (an educated mother that can eventually help studying, a room for being concentrated and a PC for working, a high number of books at disposal) act in reducing the probability to become a repeater. As a consequence, policy-makers should then pay attention also to stimulate and help families in creating these favourable conditions – a more profound discussion about the implications of these results is in the final section of the paper.

Turning our attention to the school-level characteristics, the only factor, which is correlated with a reduced probability of being repeater in a statistically significant way, is a measure for the proportion of students from disadvantaged background being lower than 10% (odds ratio = 0.76). In other words, if a student is attending a school populated by more affluent students, the probability of being a repeater is lower. This evidence can be the result of (i) students/families self-selection (more affluent students attend schools where there is a higher proportion of advantaged students) (ii) peer-effects, where students with a better background (and who have better results at school) exert a positive influence on the overall educational climate and ability to improve achievement and reduce the phenomenon of repetition, and (iii) the positive impact of a higher average level of academic self-concept of advantaged students (for a discussion of the relationships between academic self-concept and achievement, see Marsh & Martin, 2011; an estimation of the effect of school's average socioeconomic status on students' achievement is in Perry & McConney, 2010; a theoretical discussion of peer effects and its effects on students at elementary schools is in Ammermueller & Pischke, 2009). The data we have at-hand do not allow exploring which of the two (complementary) factors

is more important. The other two variables that measure the characteristics of the schooling experience (eventual problems of class disturbance during lessons and the wealth of the area in which the school operates) have coefficients with the expected sign – i.e. odds ratios <1 – but do not gain statistical significance.

Some additional insights come from the analysis of the results from the competing model in which the variables measure factors that obstacle achievement, thus increasing the probability of being a repeater. More specifically, two personal characteristics are particularly influent in determining repeating status: (i) being immigrant (or, better, speaking a language at home different from that of the test) odds ratio = 1.98 and (ii) problems in reading letters before going to the school, odds ratio = 2.10 (the latter is also complemented by the problem of not having attended a preschool of at least three years, odds ratio = 1.96). Conversely to what described above, the absence of adequate resources at home harms achievement and increase the probability of being a repeater; the odds ratios associated with having less than 25 books, not having a PC, room or internet connection at home are 1.56, 1.49, 1.22 and 1.19 respectively – and all are statistically significant. A further factor of interest is that, while having parents who help with homework does not reduce the probability of being a repeater, students whose parents "never help homework by asking what learned" are 1.49 times more likely to be repeaters.

Turning to the school-level variables, students attending schools where there is a high proportion of students from disadvantaged background (>50%) are more likely to be repeaters (odds ratio: 1.49); again, it is difficult to disentangle if the effect is due to disadvantaged students' self-selection in schools with similar peers, or to ex-post (negative) peer effects. In this case, there is also a slight negative effect of attending a class in which "disturbance during lesson" is a major problem (odds ratio = 1.22); a possible interpretation is that "at margin" disadvantaged students are negatively affected by attending classes where it is difficult to stay concentrated on academic tasks, and this influence their probability to be forced repeating one grade.

The models also includes country fixed-effects (dummies) that must be interpreted as reflecting country-specific structural differences in the repeating rates; in this perspective, three groups of countries can be easily detected[11]:

- Those in which the proportion of repeaters is very high, as odd ratio is >1 and statistically significantly different from the base case (Bulgaria, where the repeating rate is around 3.5%), after having considered the explanatory factors included in the empirical analysis. This is the case of Finland, France, Belgium, etc.).

- The converse case is that of countries with repeating rates lower than the base case, net of the role of other explanatory variables (for instance, Italy, Poland and England).

- Lastly, there is a group of countries for which the probability of being resilient students is not statistically different from that estimated (using also covariates) for the baseline case (see Slovak Republic, Slovenia, etc.).

These estimates must be read and interpreted together with the description of the different educational settings and mechanisms which are in place, in the various countries, for deciding how frequently and which circumstances the students must repeat a grade. In this light, the estimated country-level dummies identify specific choices made by the policy-makers about the overall features of the educational systems, as well as the consolidated "cultures" about the use of repetition as a pedagogic tool for favouring student achievement and regulating progress across grades.

5. Concluding remarks

In this paper, we analyze a (large) set of factors potentially associated with the probability for European students to be retained in early stages of his/her educational career. For this purpose, we use PIRLS 2011 data for 16 European countries. A novel procedure to identify if a student has been retained is proposed; such method is based on combining information about entry age at school and the grade the student is enrolled when the survey was carried out. The main findings suggest that, in addition to

[11] These results are consistent in both models, as can be noted by comparing the two columns of results. Nevertheless, the magnitude of the effects seems a little bigger when considering the "negative" characteristics as explanatory variables (those increasing the probability of repetition, right column).

structural differences in policy-making across countries (in our empirical approach, captured by country dummies), there are common factors that make a student more/less at-risk to be retained. Among these factors, four groups can be broadly identified: (i) family's socioeconomic background, (ii) early childhood activities and skills, (iii) available (economic and cultural) resources, and (iv) the student composition of the school attended. Generally speaking, the higher the socioeconomic condition of the family (and its resources made available for the student) the lower the probability that a student is retained; moreover, attending a school where the proportion of disadvantaged students is lower reduce the probability of being retained, all else equal. All these findings corroborate previous literature about the determinants of student achievement, which highlights the positive effect exerted by the family and peers' socioeconomic background (Haveman & Wolfe, 1995; Sirin, 2005; Perry & McConney, 2010; Willms, 2010).

Therefore, the evidence about the positive influence of the average socioeconomic condition of the schools should raise some concerns. Indeed, if having better peers help in attenuating at-risk factors, it is likely that attending disadvantaged schools (i.e. those in which there is a high proportion of disadvantaged students) does increase the probability of becoming repeaters. This channel can reinforce vicious circles of inequality: disadvantaged students – who are likely to become repeaters – attend more probably disadvantaged schools, and this increases their negative attitude towards obtaining good results. The policy-makers, when aware of these statistical relationships, should operate to facilitate a more diverse socioeconomic composition of student population within the schools.

Nevertheless, in our opinion, the most important result is the one related to the early childhood skills; the indicators that measure the activities undertaken and the competences acquired before schooling (i.e. in the family context and/or in formal pre-schooling experiences) are all statistically significant and positively correlated with a lower probability of being retained during primary schooling. It is interesting to note that this suggests that the probability of repeating one or more years is connected with academic skills acquired before starting primary school level. Moreover, this finding seems coherent with the most modern theories of human capital and the related studies about educational production functions (EPFs), which indicates (i) that education is a cumulative process, and (ii) that it starts very early in life, well before starting formal

educational processes (see Cuhna *et al.*, 2006). These findings call for policy interventions in the direction of improving preschool activities and opportunities, and of helping families in supporting the educational and cultural experiences of their children.

More research will be needed in future to explore more profoundly the results discussed here. First, it would be interesting to evaluate if the positive/negative effect of better/worse-off peers is of a similar magnitude for disadvantaged students and their advantaged counterparts. Understanding more about the effect exerted by peers on the probability to repeat can help the policy-makers and school managers in targeting specific interventions towards particular subgroups of students. Second, more detailed information about the skills acquired before schooling can help in designing proper educational activities and in creating extracurricular opportunities, also by considering different and heterogeneous student profiles. Third, it could be helpful to check whether the factors associated with higher probability of being a repeater do accompany the students later in their educational life, or if are there factors that – in the "cumulative" process of education – can act for reducing the impact of the characteristics that make some students more at-risk.

A final reflection is worth of consideration about the role of national different cultures and policies. Obviously, it is not possible to coordinate different attitudes towards the use of repetition as an educational tool; nevertheless, the research in the field is today unanimous in highlighting that this practice harms, and does not help, to reduce inequalities. In this sense, it would be clever to rethinking the role of this instrument in the framework of the overall educational policies for helping disadvantaged students – acting both directly and indirectly through those factors that reduce at-risk situations.

References

Alexander, K.L., Entwisle, D.R., Dauber, S.L. (2003). *On the success of failure: A reassessment of the effects of retention in the primary school grades,* Cambridge University Press, New York.

Ammermueller, A., Pischke, J. (2009). Peer effects in European primary schools: Evidence from PIRLS. *Journal of Labor Economics*, 27, 315- 348.

Allen, C. S., Chen, Q., Wilson, V. L., Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: a meta-analytic multilevel analysis. *Educational Evaluation and Policy Analysis*, 31, 480-499.

Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E. (2012). *The condition of education 2012*. NCES 2012-045. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Bali, V., Anagnostopoulos, D., Roberts, R. (2005). Toward a Political Explanation of Grade Retention. *Educational Evaluation and Policy Analysis,* 27 (2): 133-55.

Belot, M., Vandenberghe, V. (2014). Evaluating the 'threat'effects of grade repetition: exploiting the 2001 reform by the French-Speaking Community of Belgium. *Education Economics*, *22*(1), 73-89.

Bonvin, P., Bless, G., Schüpbach, M. (2008). Grade Retention: Decision-Making and Effects on Learning As Well As Social and Emotional Development. *School Effectiveness and School Improvement*, 19 (1): 1-19.

Bryk, A.S., Raudenbush, S.W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage Publications, Newbury Park, Thousand Oaks, CA.

Carneiro, P., Meghir, C., Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, *11*(s1), 123-160.

Chen, X., Liu, C., Zhang, L., Shi, Y., Rozelle, S. (2010). Does taking one step back get you two steps forward? Grade retention and school performance in poor areas in rural China. *International Journal of Educational Development*, 30, 544-559.

Cordero, J.M., Manchón, C., Simancas, R. (2014): La repetición de curso y sus factores condicionantes en España, *Revista de Educación, in press, doi:*10.4438/1988-592X-RE-2014-365-263

Corman, H. (2003). The Effects of State Policies, Individual Characteristics, Family Characteristics and Neighborhood Characteristics on Grade Repetition in the United States. *Economics of Education Review*, 22 (4): 409-20.

Creemers, Bert P. M., and Leonidas Kyriakides. 2008. *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. London: Routledge.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1, 697-812.

Cunha, F., Heckman, J. J., Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica, 78*(3), 883-931.

Dalton, B. (2012): Grade Level and Science Achievement: US Performance in Cross-National Perspective. *Comparative Education Review,* 56(1), 125-154.

Deming, D., Dynarski, S. (2008). The lengthening of childhood. *Journal of Economic Perspectives*, 22, 71–92.

Dong, Y. (2010). Kept back to get ahead? Kindergarten retention and academic performance, *European Economic Review*, 54, 219-236.

Dupriez, V., Dumay, X., Vause, A. (2008). How do school systems manage pupils´ heterogeneity?. *Comparative Education Review*, 52 (2), 245-273.

Ehmke, T., Drechsel, B., Carstensen, C.H. (2010). Effects of Grade Retention on Achievement and Self-Concept in Science and Mathematics. *Studies in Educational Evaluation*, 36 (1-2): 27-35.

Eide, E. R., Showalter, M. H. (2001). The effect of grade retention on educational and labor market outcomes. *Economics of Education Review*, 20(6): 563-576.

Eisenmon, T. (1997). *Reducing repetition: Issues and strategies*. Paris: IIEP-UNESCO.

Elder, T. E., Lubotsky, D. H. (2009). Kindergarten entrance age and children's achievement impacts of state policies, family background, and peers. *Journal of Human Resources, 44*(3), 641-683.

El-Hassan, K. (1998). Relation of academic history and demographic variables to grade retention in Lebanon. *Journal of Educational Research*, 91, 279-288.

Else-Quest, N. M., Hyde, J. S., Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological bulletin,136*(1), 103.

Eurydice (2011): *Grade retention during compulsory* education in Europe: Regulation and Statistics, Education, Audiovisual and Culture Executive Agency. Available at: http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/126EN.pdf

Evans, M. D., Kelley, J., Sikora, J., Treiman, D. J. (2010). Family scholarly culture and educational success: Books and schooling in 27 nations. *Research in Social Stratification and Mobility*, *28*(2), 171-197.

Ferguson, P., Jimerson, S.R., Dalton, M.J. (2001). Sorting Out Successful Failures: Exploratory Analyses of Factors Associated With Academic and Behavioral Outcomes of Retained Students. *Psychology in the Schools*, 38 (4): 327-41.

Fredericks, J. A., Blumenfeld, P. C., Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59–109.

Frey, N. (2005). Retention, social promotion and academic redshirting: What do we know and need to know? *Remedial and Special Education*, 26 (6), 332-346.

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*(12), 1373-1392.

Goldstein, H. (1995). *Multilevel statistical models (2nd edition)*, New York: Wiley Publishers.

Gomes-Neto, J.B., Hanushek, E.A. (1994): Causes and Consequences of Grade Repetition: Evidence from Brazil. *Economic Development and Cultural Change*, 43:1 117–148.

Goos, M., Schreier, B.M. Knipprath, H.M., De Fraine, B. Van Damme, J., Trauwein, U. (2013a). How Can Cross-Country Differences in the Practice of Grade Retention Be Explained? A Closer Look at National Educational Policy Factors. *Comparative Education Review*, 57 (1), 54-84.

Goos, M., Van Damme, J., Onghena, P., Petry, K., de Bilde, J. (2013b). First-grade retention in the Flemish educational context: Effects on children's academic growth, psychosocial growth, and school career throughout primary education. *Journal of school psychology*, *51*(3), 323-347.

Greene, J.P., Winters, M.A. (2007). Revisiting Grade Retention: An Evaluation of Florida's Test-Based Promotion Policy. *Education Finance and Policy*, 2 (4): 319-40.

Guèvremont, A., Roos, N.P., Brownell, M. (2007). Predictors and Consequences of Grade Retention: Examining Data From Manitoba, Canada. *Canadian Journal of School Psychology*, 22 (1): 50-67.

Hanushek, E. A., Rivkin, S. G., Taylor, L. L. (1996). *Aggregation and the estimated effects of school resources.* National Bureau of Economic Research, No. w5548.

Haveman, R., Wolfe, B. (1995). The determinants of children's attainments: a review of methods and findings. *Journal of Economic Literature*, XXXIII, 1829-1878.

Holmes, C. T. (1989). Grade-level retention effects. A meta-analysis of research studies" in Shepard, L. y Smith, M.L. (eds.): *Flunking grade: Research and policies on retention*, London: The Falmer Press, 16-33.

Hong, G., Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27, 205–224.

Hong, G., Yu, B. (2007). Early-grade retention on children´s reading and math learning in elementary years. *Educational Evaluation and Policy Analysis*, 29, 239-261.

Hox, J. (2002). *Multilevel Analysis. Techniques and Applications*, Mahwah: Lawrence Erlbaum Associates.

Huddleston, A. P. (2014). Achievement at whose expense? A literature review of test-based grade retention policies in U.S. school. *Education Policy Analysis Archives, 22* (18).

Hughes, J. N., Kwok, O. M., Im, M. H. (2013). Effect of Retention in First Grade on Parents' Educational Expectations and Children's Academic Outcomes. *American Educational Research Journal, 50*(6), 1336-1359.

Ikeda, M. (2011). PISA in focus: When students repeat grades or are transferred out of school: What does it mean for education systems. Available at the OECD: http://www.oecd.org/dataoecd/35/58/48363440.pdf.

Ikeda, M., García, E. (2013). Grade repetition: A comparative study of academic and non-academic consequences. *OECD Journal: Economic Studies, 2013*(1), 1-47.

Im, M. H., Hughes, J. N., Kwok, O. M., Puckett, S., Cerda, C. A. (2013). Effect of retention in elementary grades on transition to middle school. *Journal of School Psychology, 51*(3), 349-365.

Jacob, B. A. (2005): Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools. *Journal of Public Economics,* 89 (5-6), 761- 96.

Jacob, B.A., Lefgren, L. (2004). Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. *Review of Economics and Statistics*, 84 (1), 226-244.

Jacob, B.A., Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1, 3, 33-58.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, 30 (3), 420-437.

Jimerson, S. R., Anderson, G. E., 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(4): 441-457.

Jimerson, S. R., Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School Psychology Quarterly*, 22 (3), 314-339.

Kloosterman, R., de Graaf, P.M. (2010). Non-Promotion or Enrolment in a Lower Track? The Influence of Social Background on Choices in Secondary Education for Three Cohorts of Dutch Pupils. *Oxford Review of Education,* 36 (3): 363-84.

Le Donné, N. (2014). European Variations in Socioeconomic Inequalities in Students' Cognitive Achievement: The Role of Educational Policies. *European Sociological Review, 30*(3), 329-343.

Lee, J.W., Barro, R.J. (2001). Schooling Quality in a Cross-Section of Countries. *Economica,* 68 (272), 465-88.

Lorence, J. (2006). Retention and academic achievement research revisited from a United States perspective. *International Education Journal*, 7, 731–777.

Manacorda, M. (2012). The cost of grade retention. *The Review of Economics and Statistics*, 94 (2), 596-606.

Marsh, H. W., Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81(1), 59-77.

Martin, A. (2011). Holding back and holding behind: grade retention and students´ non academic and academic outcomes. *British Educational Research Journal*, 37 (5), 739-763.

McCoy, A. R., Reynolds, A. J. (1999). Grade retention and school performance: an extended investigation. *Journal of School Psychology*, 37, 273-298.

McEwan, P. J., Shapiro, J. S. (2008). The Benefits of Delayed Primary School Enrollment Discontinuity Estimates Using Exact Birth Dates. *Journal of Human Resources, 43*(1), 1-29.

Mullis, I.V.S., Martin, M.O., Foy, P., Drucker, K.T. (2012). *PIRLS 2011 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Ou, S., Reynolds, A. J. (2010). Grade retention, post-secondary education and public aid receipt. *Educational Evaluation and Policy Analysis*, 32, 118-139.

Pagani, L., Tremblay, R., Vitaro, F., Boulerice, B., McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology*, 13, 297-315.

Patrinos, H.A., Psacharopoulos, G. (1996): Socioeconomic and Ethnic Determinants of Age-Grade Distortion in Bolivian and Guatemalan Primary Schools. *International Journal of Educational Development*, 16, 3–14.

Perry, L., McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *The Teachers College Record*, 112(4), 7-8.

Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31 (4), 729-759.

Reschly, A. L., Christenson, S. L. (2013). Grade retention: Historical perspectives and new research. *Journal of School Psychology*, *51*(3), 319-322.

Rumberger, R. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32, 583-625.

Rutkowski, L., González, E., Joncas, M., von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher*, 39 (2), 142-151.

Shepard, L., Smith, M.L. (1990). Synthesis of research on grade retention", *Educational Leadership*, 47.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417-453.

Todd, P. E., Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485), 3-33.

Willms, J. D., Somers, M.A. (2001). Family, Classroom, and School Effects on Children's Educational Outcomes in Latin America. *School Effectiveness and School Improvement* 12 (4), 409-45.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *The Teachers College Record*, 112(4), 3-4.

Willson, V., Hughes, J. (2009). Who Is Retained in First Grade? A Psychosocial Perspective. *Elementary School Journal*, 109 (3): 251-66.

Wu, W., West, S. G., Hughes, J. N. (2008). Short-term effects of grade retention on the growth rate of Woodcock–Johnson III broad math and reading scores. *Journal of School Psychology*, 46, 85–105.

Wu, W., West, S. G., Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes and school relationships. *Journal of Educational Psychology*, 102, 135–152.

Xia, N., Kirby, S. N. (2009). *Retaining students in grade: A literature review of the effects of retention on students' academic and nonacademic outcomes*. (Technical Report No. 678). Available at RAND Education website: http://www.rand.org/pubs/technical_reports/TR678/

Table 1. Proportion of students considered as repeaters across countries

Country	Observations	Not-repeaters	Repeaters	% Repeaters
Bulgaria	5,261	5,075	186	3.5%
Finland	4,640	4,389	251	5.4%
France	4,438	3,975	463	10.4%
Hungary	5,204	4,802	402	7.7%
Ireland	4,524	3,739	785	17.4%
Italy	4,189	4,087	102	2.4%
Norway	3,190	3,178	12	0.4%
Poland	5,005	4,925	80	1.6%
Portugal	4,085	3,794	291	7.1%
Romania	4,665	4,478	187	4.0%
Slovak Republic	5,630	5,438	192	3.4%
Slovenia	4,512	4,302	210	4.7%
Spain	8,580	7,813	767	8.9%
Sweden	4,622	4,486	136	2.9%
England	3,927	3,853	74	1.9%
Belgium (French)	3,727	3,051	676	18.1%
Total	76,199	71,385	4,814	6.3%

Source: authors' elaborations on PIRLS 2011 data

Table 2. Comparison of grade retention rates in primary education in European countries according to different criteria

Countries	Our approach	PISA 2009	Eurostat 2008
Belgium	17.17%	22.0%	21.4%
Bulgaria	3.54%	2.7%	6.0%
England	1.88%	2.0%	5.3%
Finland	5.41%	2.4%	4.1%
France	10.43%	17.8%	19.4%
Hungary	7.72%	6.2%	1.2%
Ireland	17.35%	11.0%	-
Italy	2.43%	1.0%	2.7%
Norway	0.38%	0.0%	0.15%
Poland	1.60%	0.4%	0.4%
Portugal	7,12%	22.4%	28.0%
Romania	4.01%	2.3%	-
Slovak Republic	3.41%	1.9%	5.3%
Slovenia	4.65%	-	-
Spain	8.94%	12.2%	16.3%
Sweden	2.94%	3.8%	-

Source: authors' elaborations on PIRLS 2011, PISA 2009 and Eurostat 2008 data

Table 3. Differences in reading achievement between repeaters and non-repeaters by country

Country	Label	Mean	Std. Err.	Gap
Belgium	Non-repeater	517.70	1.11	51.46
	Repeater	466.25	2.19	
Bulgaria	Non-repeater	541.34	1.10	87.89
	Repeater	453.45	5.97	
England	Non-repeater	553.25	1.31	60.40
	Repeater	492.85	10.54	
Finland	Non-repeater	571.35	0.92	58.48
	Repeater	512.87	4.14	
France	Non-repeater	528.97	1.01	67.32
	Repeater	461.65	2.77	
Hungary	Non-repeater	551.28	1.05	73.49
	Repeater	477.79	4.27	
Ireland	Non-repeater	551.71	1.23	3.05
	Repeater	548.66	2.91	
Italy	Non-repeater	542.32	1.05	42.60
	Repeater	499.72	7.64	
Norway	Non-repeater	509.94	1.09	51.48
	Repeater	458.46	22.77	
Poland	Non-repeater	532.07	1.01	93.51
	Repeater	438.57	8.21	
Portugal	Non-repeater	547.34	1.04	57.91
	Repeater	489.43	3.75	
Romania	Non-repeater	521.04	1.26	81.66
	Repeater	439.38	7.42	
Slovak Rep	Non-repeater	539.14	0.91	94.06
	Repeater	445.08	5.15	
Slovenia	Non-repeater	531.53	1.04	54.79
	Repeater	476.74	6.15	
Spain	Non-repeater	524.34	0.73	56.77
	Repeater	467.57	2.36	
Sweden	Non-repeater	540.93	0.98	65.75
	Repeater	475.18	6.36	
TOTAL	**Non-repeater**	**540.17**	**0.24**	**33.41**
	Repeater	**506.76**	**0.81**	

Table 4. Definition of variables included in the analysis

Level	Positive	Negative
Student	GIRL: Dummy variable	BOY: Dummy variable
	LANG: Language of the test is always spoken at home	LANG: Language of the test is not always spoken at home
	MOTHEDU: Mother has an university degree	MOTHEDU: Mother has a lower secondary education level
	PRESCHOOL: Student attend preschool at least 3 years	PRESCHOOL: Student did not attend preschool
	PC: There is a computer at home	PC: There is not a computer at home
	ROOM: Student has an own room	ROOM: Student does not have an own room
	INTERNET: Student has internet connection at home	INTERNET: There is no access to internet at home
	BOOKS200: There are more than 200 books at home	BOOKS25: There are less than 25 books at home
	SKILLSletters: Student could recognize most of the letters of the alphabet before starting primary school	SKILLSletters: Student could not recognize the letters of the alphabet before starting primary school
	SKILLSwords: Student could read some words before starting primary school	SKILLSwords: Student could not read any word before starting primary school
	SKILLSwrite: Student could write some words before starting primary school	SKILLSwrite: Student could not write any word before starting primary school
	ASKLEARNED: Parents ask student what he/she has learned almost every day	ASKLEARNED: Parents never (or almost never) ask student what he/she has learned
	HELPHOMEWORK: Parents help student with homework almost every day	HELPHOMEWORK: Parents never (or almost never) help student with homework
School	BACKGROUND: Proportion of students from disadvantaged background < 10%	BACKGROUND: Proportion of students from disadvantaged background > 50%
	HIGHINCOME: The average income level of the area is high	LOWINCOME: The average income level of the area is low
	DISTURBANCE: The disturbance during classes is not a problem	DISTURBANCE: The disturbance during classes is a moderate or serious problem
	ABSENTEEISM: The absenteeism during classes is not a problem	ABSENTEEISM: The absenteeism during classes is a moderate or serious problem
	COMPINST: There are computers available for instruction	COMPINST: There are not computers available for instruction
	LIBRARY: There is a library with more than 5,000 books at the school	LIBRARY: There is no at the school

Table 5. Descriptive statistics

Variable	Mean	Std. Dev.
Dep. variable		
REPEAT	0.082	0.275
Student variables		
GIRL	0.493	0.500
BOY	0.507	0.500
LANG (good)	0.761	0.427
LANG (bad)	0.023	0.149
MOTHEDU (good)	0.315	0.464
MOTHEDU (bad)	0.439	0.496
PRESCHOOL (good)	0.534	0.499
PRESCHOOL (bad)	0.044	0.205
PC (good)	0.934	0.248
PC (bad)	0.055	0.227
ROOM (good)	0.730	0.444
ROOM (bad)	0.257	0.437
INTERNET (good)	0.854	0.354
INTERNET (bad)	0.132	0.338
BOOKS200	0.198	0.398
BOOKS25	0.239	0.427
SKILLSletters (good)	0.392	0.488
SKILLSletters (bad)	0.175	0.380
SKILLSwords (good)	0.237	0.425
SKILLSwords (bad)	0.335	0.472
SKILLSwrite (good)	0.297	0.457
SKILLSwrite (bad)	0.234	0.424
ASKLEARNED (good)	0.637	0.481
ASKLEARNED (bad)	0.013	0.113
HELPHOME (good)	0.482	0.500
HELPHOME (bad)	0.040	0.197
School variables		
BACKGROUND (good)	0.358	0.479
BACKGROUND (bad)	0.142	0.349
HIGHINCOME	0.083	0.276
LOWINCOME	0.276	0.447
DISTURBANCE (good)	0.321	0.467
DISTURBANCE (bad)	0.201	0.400
ABSENTEEISM (good)	0.622	0.485
ABSENTEEISM (bad)	0.076	0.265
COMPINST (good)	0.876	0.329
COMPINST (bad)	0.037	0.190
LIBRARY (good)	0.278	0.448
LIBRARY (bad)	0.118	0.323

Table 6. Results of the empirical analysis

Dependent variable: REPEAT	Covariates: Good values Odd ratio/ z-value	Covariates: Bad values Odd ratio/ z-value
Individual-level variables		
Girl	0.753*** -5.38	
Boy		1.373*** 6.03
Language	0.731*** -4.66	1.981*** 5.27
Mother's education	0.370*** -11.81	1.209*** 3.09
Skills before school (letters)	0.493*** -8.04	2.126*** 9.23
Skills before school (write)		1.346*** 3.49
Skills before school (words)	0.743*** -2.95	
Preschool=NO		1.964*** 5.68
Preschool= YES (3 years)	0.637*** -7.81	
Parents never help homework by asking what learned		1.486* 2.26
Books at home: 200 or more	0.817* -2.07	
Books at home: 25 or less		1.559*** 6.79
Resource at home: PC (yes/no)	0.689*** -3.69	1.492*** 3.86
Resource at home: ROOM (yes/no)	0.853* -2.38	1.223** 3.18
Resource at home: INTERNET (yes/no)	0.889 -1.51	1.188* 2.32
School-level variables		
Proportion of students from disadvantaged background <10%	0.763** -3.18	
Proportion of students from disadvantaged background >10%		1.487*** 3.73
In the class the problem of disturbance during lessons is relevant? (no/yes)	0.914 -1.08	1.216* 2.63
Surround area of the school: High Income	0.909 -0.67	
Surround area of the school: Low Income		1.120 1.20

Dependent variable: REPEAT	Covariates: Good values Odd ratio/ z-value	Covariates: Bad values Odd ratio/ z-value
Country effects		
Bulgaria	1.00	1.00
	.	.
Finland	2.484***	4.073***
	5.37	7.97
France	4.647***	7.749***
	9.94	12.17
Hungary	2.542***	3.722***
	6.26	7.84
Ireland	6.556***	11.830***
	11.24	14.50
Italy	0.608**	1.091
	-2.56	0.41
Norway	0.165***	0.274**
	-3.69	-2.61
Poland	0.421***	0.614**
	-4.81	-2.66
Portugal	2.020***	3.391***
	4.14	6.80
Romania	1.022	1.223
	0.11	1.02
Slovak Republic	0.864	1.127
	-0.65	0.52
Slovenia	1.288	2.438***
	1.45	5.26
Spain	3.759***	6.294***
	9.07	11.22
Sweden	1.045	2.230***
	0.22	3.81
England	0.275***	1.888**
	-6.39	2.91
Belgium	8.085***	14.480***
	12.98	15.86
R2	0.16	0.13
N	76,199	76,199