The version of record of this article has been published and is available in International Social Science Journal (date of publication 13 October 2023)

https://doi.org/10.1111/issj.12468

Education policy reform and the impact of free preschool education on boys' and girls' reading competence

María Ladrón de Guevara Rodríguez[1]

Luis Alejandro Lopez-Agudo[2]

Oscar David Marcenaro-Gutierrez[3]

Abstract

Numerous authors have reported a positive relationship between preschool enrolment and academic performance in later years, even helping to reduce the academic gap existing between students from different socioeconomic backgrounds. In this context, this paper goes further by analysing the impact that early childhood education (from 3 to 6 years) has on Spanish boys' and girls' reading achievement when they are in 4th grade (10 years). For this purpose, we take advantage of the 2006 Spanish education reform that promoted free second stage preschool education by using data from PIRLS 2011 and 2016 and an instrumental variable approach, in order to get closer to causality than previous literature for Spain. We find that attending preschool has a positive influence on reading achievement, well above the impact found when using simple Ordinary Least Squares estimates, being higher for girls than for boys.

Keywords: preschool; instrumental variables; PIRLS; Spain.

JEL Codes: I20, I21, I28.

Acknowledgements: This work has been partly supported by *Ministerio de Ciencia e Innovación* (under Research Project PID2020-119471RB-I00), by the Andalusian Regional Government (SEJ-645), by the *Universidad de Málaga* (under Research Project B1-2022_23) and by the *Fundación Ramón Areces*. We also acknowledge the scholarship FPU20/01509 of the *Ministerio de Universidades* and the training received from the *Programa de Doctorado en Economía y Empresa* of the *Universidad de Malaga.*

[1] Programa de Doctorado en Economía y Empresa, Universidad de Málaga, Málaga, Spain. Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: marialadron@uma.es. Tel.: +34 952131206. ORCID: 0000-0002-5087-422X

[2] Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: lopezagudo@uma.es. Tel.: +34 952131207. ORCID: 0000-0002-0906-3206

[3] Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: odmarcenaro@uma.es. Tel.: +34 952137003. ORCID: 0000-0003-0939-5064

1. Introduction

Childhood is a milestone in a person's cognitive and non-cognitive development. During the first years of life we develop the foundations for thinking and behaviour. As a result, children's development is more malleable and more susceptible to any contextual intervention. In this sense, early environments can become predictors of later outcomes, favouring not only academic performance, but also individuals' future productivity (Heckman et al., 2006). Thus, the training and education given to young people in early ages can have a positive impact on their later academic performance and in their entry into the labour market (Bakken et al., 2017).

In this context, early childhood/pre-primary education in Spain – subdivided into first stage (or kindergarten) for ages 0-3 and second stage (or preschool) for ages 3-6 – can be considered as a key transition period; early education can serve as a foundation for acquiring basic – cognitive and non-cognitive – skills in order to develop more advanced ones later on (Cunha & Heckman, 2008)[4]. Bearing in mind that skills are acquired over time, children who are in more favourable environments and can, e.g., develop language earlier, may have a clear advantage over the rest in the educational process. Consequently, early investments in children's education can be used to reduce the academic achievement gap between the more advantaged and their less advantaged counterparts (Heckman et al., 2006; Cunha & Heckman, 2007; Quirk et al., 2013).

In this sense, and given that in recent years there has been a general tendency to design policies to facilitate childcare access, some studies have tried to analyse the impact that early childhood education may have on educational attainment, with the vast majority of them pointing in the same direction: early childhood education enhances students' future academic performance (Heckman et al., 2010; Havnes & Mogstad, 2011; Keys et al., 2013; Duncan & Magnuson, 2013; Auger et al., 2014; Felfe et al., 2015; Ansari et al., 2017; McCoy et al., 2017; Nakajima et al., 2019; Bai et al., 2020; among others). Besides cognitive stimulation and academic gains, attending preschool also seems to have a positive impact on the non-cognitive sphere. Learning begets less tangible outcomes as well, and early childhood education can also foster social skills and behavioural patterns (Heckman et al., 2010; Bakken et al., 2017).

Given how important it may be for policymaking to assess whether early investment in children pays off later, throughout this paper we have analysed the influence that attending preschool has on Spanish students' reading performance. Analysing reading achievement and the role that early education plays in it is particularly relevant and justifies the aim of our study. This is worth of analysis because, although children can learn almost spontaneously skills related to language and reading such as letter recognition (Ehri et al., 2001), access to early education may foster this literacy development. Moreover, as it has been extensively reported in the literature, there is a noticeable gender gap in reading in favour of girls (Marinak & Gambrell, 2010; McGeown et al., 2012; Mullis et al., 2017). Consequently, analysing reading achievement and the role that early education plays in it may be particularly important for understanding these potential gender differences in academic achievement and ultimately

[4] The Spanish education system is structured in 5 levels/stages: early childhood education, primary education, secondary education, preparation for higher education and higher education. Early childhood education is divided into 2 stages. The first stage goes from ages 0 to 3 and is usually held in kindergartens. The second stage, on which we have based our study, goes from ages 3 to 6.

explain the overrepresentation of females in reading-related degrees (Flabbi, 2012; Breda & Napp, 2019; Card & Payne, 2021).

In particular, the research question we want to answer is:

Does preschool attendance have a positive impact on the reading achievement of Spanish primary education boys and girls?

Compulsory education in Spain starts at the age of 6. However, parents who wanted to enrol their children in preschool education could do it in primary schools that also offered pre-primary education. Preschool education offered by Spanish schools, which at first could be public, semiprivate[5] or private, changed as a result of the educational reforms that took place at the beginning of the 2000s. Following the footsteps of Scandinavian countries and most OECD countries that had already introduced policies to make childcare cheaper or easier to access, the 2006 Spanish education reform sought to promote free second stage preschool education (between 3 and 6 years) by guaranteeing a sufficient supply in public and semiprivate centres (LOE, art. 15.2; BOE 2006).

This paper tries to get advantage of this reform and to exploit this legal change to establish the causal impact that attending preschool has on reading achievement when students reach 4th grade. In other words, as main novelty we will go beyond simple correlation analyses and approach causality using that reform as instrumental variable. To this aim, we use the rich data from the 2011 and 2016 Progress in International Reading Literacy Study (PIRLS), which contain information on children attending primary school born before and after the reform[6].

This is not the first paper focused on Spain that has analysed how early education affects academic performance; the most recent ones are Hidalgo-Hidalgo and García-Pérez (2012), Santin and Sicilia (2015), González-Betancor and López-Puig (2015) and Felfe et al. (2015). However, to the best of our knowledge, there is no evidence in Spain that has approached causality, and eventually none considering the 2006 reform that guarantees free education in the second period of ISCED 0 (from 3 to 6 years old) in public and semiprivate schools. Thus, this paper makes a valuable contribution to the literature to date.

The article is structured as follows. We briefly review some of the relevant literature in the next section. We then present the main characteristics of the dataset. Next, we describe the methodology employed. The theoretical framework shaping the results obtained are the focus of section 5. Lastly, we comment and discuss the main conclusions.

2. Literature Review

Early childhood education is an essential policy tool, not only because it boosts a country's productivity, but also because it reduces issues in the education system. Consequently, numerous authors have explored the impact of preschool on academic attainment. For instance, Andrews et al. (2012) found that a large-scale public preschool programme implemented in Texas had a positive effect on mathematics and reading achievement test scores of students from low socio-economic backgrounds by 0.05 to 0.1 standard deviations (SD). In addition, it decreased student's likelihood of receiving special education services and of holding back a year. In the same line, Ramey et al.

[5] Private schools receiving public funds.

[6] Using previous or more recent data from PIRLS may artificially dilute the potential influence of preschool, so we have employed the closest data of students who are affected by the reform (PIRLS 2016) and the previous closest one (PIRLS 2011), as surveyed students are aged 9-10 when they take the test.

(2000) conducted a randomised controlled trial to evaluate the impact that a programme of comprehensive education, health care and family support could have on students in North Carolina. Running the intervention from preschool to age 5, the authors found that children in the intervention group significantly outperformed their counterparts in academic achievement and were less likely to be retained or to receive special education. Moreover, those mothers whose children received childcare during preschool years improved their own educational attainment and employment.

Likewise, Havnes and Mogstad (2011) used a differences-in-differences (DID) approach to obtain causal relationships between childcare and children's long-term achievements. The authors found that the reform implemented in Norway in 1975, which universalised access to childcare, had large positive effects on children's outcomes in adulthood. Access to childcare not only increased the likelihood of completing high school and attending university, but also increased labour market attachment. The authors also reported that children with low-educated mothers were the biggest beneficiaries of the education effect, while girls were the greatest recipients of the labour market attachment effect, which not only favoured intergenerational mobility, but also contributed to reducing the gender wage gap.

Ansari et al. (2017) reported that preschool programmes in Miami-Dade County provided sustained benefits over time for low-income Latino children. In particular, children who attended this programme performed better in mathematics and literacy than their peers by the time they reached 3rd grade, even when controlling for baseline characteristics. In the same vein, Bai et al. (2020) found that the early education programme implemented in North Carolina had a positive effect on mathematics and reading outcomes. Besides decreasing the likelihood of special education placement and grade repetition, the authors found that the positive effects derived from the programme were sustained through 3rd to 8th grades.

Datta and Simonsen (2016) analysed how non-parental childcare affected educational outcomes. The authors used data from Danish registers and, as an instrument, whether the municipality in which the child lives offers guaranteed access to childcare. They found that children who were enrolled in childcare performed better in Danish language scores by 0.2 SD at the end of compulsory education. Cortázar (2015) evaluated Chile's public early childhood education programmes' impact on fourth grade academic achievement, finding that those children who attended the programme performed an average of 0.19 SD higher in reading and 0.23 SD higher in mathematics.

In the context of Southern European countries, Brilli et al. (2016) used the database provided by the Italian Institute for the Evaluation of the Educational System (IVALSI) for 2008-2009, which contained information about 2nd and 5th year primary school students, to explore the impact of public childcare coverage on mothers' working status and children's school performance. The authors concluded that the availability of childcare, in addition to increasing the likelihood of mothers to participate in the labour market, positively contributed to students' cognitive development by having a positive effect of 0.18 SD on language scores. Likewise, Corazzini et al. (2021) analysed the impact of early childcare attendance on the educational outcomes of second-generation immigrant students. They used the database of 5th grade students collected by the INVALSI for the academic years 2014-2015 to 2016-2017. Using cross-sectional and longitudinal variation in early childcare supply in Italian municipalities as an instrument, the authors found that there was a positive effect of 0.6 SD between childcare and language test scores for immigrant students. However, this effect was -0.17 SD for native students.

In the context of Portugal, Pinto et al. (2013) examined the joint effects of the quality of centre-based childcare and family environment on children's language, communication and early literacy development. Assessing 95 preschool children attending early childhood education classrooms in Porto, the authors concluded that preschool quality may be an important moderator between family environment quality and children's skills, having a positive effect on their early literacy.

Previous meta-analyses also seem to find a positive relationship between preschool education and academic achievement. Magnuson et al. (2016) found that early childhood education had significant and roughly equal effects for both boys and girls on educational achievement. Conducting a meta-analysis of 23 early childhood education programmes, they found a positive influence of 0.20 SD on students' academic performance, although there were no significant effects on adult outcomes like employment. Similarly, van Huizen and Plantenga (2018) found a positive relationship between the two phenomena, pointing out that early childhood education seemed to benefit children more in the cognitive domain rather than in the non-cognitive one.

On the other hand, although the effect of early childhood education on academic performance seems to be positive, we find a few studies that add some nuances to this conclusion. Morabito et al. (2018), conducted a randomised control trial to compare the effects of preschool education on a sample of children in Mauritius according to whether preschool education was of high quality or not. They found that the quality of preschool education had no significant effect on children's overall achievement. Using the same approach, Lipsey et al. (2018) found significant gains in mathematics and reading achievement for children attending preschool compared to the control group. However, those positive effects on academic performance faded over time, and the control group's performance was better on some performance measures by the time they reached 2nd grade. The fact that early childhood education's positive effects on academic performance disappear over time has also been supported by other authors (see e.g., Magnuson et al., 2007; Heckman et al., 2010; Miller & Bassok, 2019). In contrast, we also find few authors who have reported no significant effects of early childhood education on academic performance (Goldfeld et al., 2012; Zucker et al., 2013). In particular, Li et al. (2020) found no effect of preschool attendance on the reading ability of Zhuang ethnic minority adolescents after controlling for confounding factors.

Besides boosting academic achievement, preschool education can also improve behavioural skills such as discipline or classroom participation (Berlinski et al., 2009), which may be especially important in the student's future (Blau & Currie, 2006). In this sense, Bakken et al. (2017) analysed a sample of children from low socio-economic background who received high-quality early education in the United States. The authors found that the early education received not only improved students' performance in reading and mathematics in 4th grade, but also improved their social and emotional status compared to the control group. The positive effect of preschool education on social skills is also supported by other authors like Osakwe (2009) or Heckman et al. (2010). Although social skills are first acquired through parenting (Ölçer & Aytar, 2014), continuous interactions with peers and teachers in the early school environment will allow students to put them into practice. In this sense, early childhood education can not only improve academic performance, but also foster some basic social skills (Knight & Hughes, 1995) and form complex social networks (Martin et al., 2005; DeLay et al., 2016) that prepare them for school environment.

Focusing on Spain, several authors have also tried to analyse the effect that early education has on academic achievement. Some, such as Santin and Sicilia (2015) or

González-Betancor and López-Puig (2015) have focused on the first level of ISCED 0, from a correlational perspective. For instance, Santin and Sicilia (2015) found that early childhood education from 0 to 2 years seemed to have a positive and significant effect on students' grades in 4th grade, but only for mathematics. On the other hand, Hidalgo-Hidalgo and García-Pérez (2012), using the 2011 TIMSS-PIRLS database for Spain, tried to study the impact of early childhood education attendance on students' academic outcomes in fourth grade. Using Ordinary Least Squares (OLS) and the potential population of students per preschool in each region as a variable to control for possible endogeneity, they concluded that those students who attended preschool for at least three years scored about 16 points higher in reading than those who attended for fewer years. Similarly, Felfe et al. (2015) analysed how the education reform of the early 1990s affected the academic performance of 15-year-old students. The reform divided pre-school education into two levels – the first level included children up to 3 years old, while the second level included children from 3 to 5 years old – and established the possibility for each region to establish agreements with semiprivate schools to promote free education (LOGSE, art. 111; BOE, 1990). Using data from the 2003, 2006 and 2009 PISA waves and DID approach, they found that the reform led to an increase in reading and mathematics scores.

There is a different strand of the literature, focused on the factors affecting how students learn to read. In summary, e.g., Jiménez and O'Shanaha (2008) state that this process requires the development of cognitive skills such as mastery of grapheme-phoneme conversion rules or the development of phonological awareness, which depend to a large extent on the writing system to be learnt. In this sense, and in contrast to English, the analogy approach (i.e., the way of teaching reading that teaches students unfamiliar words in order to establish analogies with known words) is not ideal for learning Spanish. In particular, Jiménez and O'Shanaha (2008) report that children do not rely on this type of linguistic unit for visual word recognition in Spanish. On the other hand, we also find studies focusing on the results obtained in reading learning conditioned on the type of instruction. For example, Rendón Romero (2019) carried out a study with the aim of determining which linguistic factors influence the development of reading skills and influence reading learning. To do so, they used two samples of pre-primary school students in two schools in the Spanish city of Seville. The authors concluded that phonological processing, letter and pseudo-word reading were determinant in reading learning compared to the speed of naming colours, pictures or numbers. Unfortunately, the PIRLS dataset does not contain information to deal with the factors affecting how students learn to read, which will be the focus for future research[7].

In any case, more research is needed. While the international literature on this topic is more prolific, the evidence in Spain is still scarce and far from causal.

3. Data and institutional background

As mentioned in the introduction, we have used data from PIRLS 2011 and 2016. PIRLS is conducted by the International Association for the Evaluation of Educational Achievement (IEA) every five years to assess the reading literacy[8] of 4th grade students,

[7] Likewise, PIRLS does not contain any information regarding dyslexia, speech and language problems or learning difficulties, which could affect students' progression in terms of reading competences. Furthermore, it does not include information regarding the quality of preschool education, which may differ depending on e.g. physical and human conditions, territories' specific conditions or in the assessment methodologies of students' performance.

[8] The definition of the reading literacy competence is the following: "Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual.

which will be the outcome of interest in the present study. In addition to the reading assessment, PIRLS offers school, student, teacher and home questionnaires, which provide information about home and school contextual factors that may influence the teaching-learning process. Therefore, in addition to parents' educational level and occupation, students' characteristics, as well as variables that can serve as a proxy for students' socioeconomic level among other variables, the questionnaires give us information on whether or not students attended preschool. Specifically, the PIRLS 2016 household questionnaire includes the following question:

Did your child attend the following before first grade? Preschool education for children age 3 or older, including kindergarten.

In this question parents had to answer "Yes" or "No". It should be noted that due to a change in the structure of the question in the PIRLS 2011 questionnaire, preschool attendance was determined for this sample by having attended ISCED Level 0 for three or more years. This change was probably due to the revision of ISCED and subsequent adoption of ISCED 2011, which subdivided early childhood education into two levels/stages (from 0 to 3 years and from 3 to the start of primary education). Nevertheless, both questions are similar, so the fact that we are measuring the same variable is assured.

In total we have a sample of 19,337 students. We have included a set of missing flag variables to control for those students who did not answer to any of the questions included in the analysis, and we have standardised the reading scores to mean 0 and standard deviation 1; this allows us to make international comparisons easier by interpreting our results as effect size.

Table A1 (Appendix) shows descriptive statistics. As we can observe, most of the children attended preschool (approximately 75%). Far from being surprising, this seems to be in line with the fact that Spain has one of the highest enrolment rates in preschool education among OECD countries (OECD, 2018). It is also observed that 21% of the mothers of children who attended preschool have the highest level of education, while this percentage is only 11 % for mothers whose children did not enrol. In the same vein, it seems that mothers of children who attended preschool are more likely to be in positions of responsibility. This may underline how important the availability of early childhood education and care can be for women's participation in the labour market.

4. Methodology

4.1. Ordinary Least Squares

First, in order to analyse the influence that attending preschool has on reading achievement, we define our baseline model as follows:

$$R_{ijt} = \alpha + \beta P_{ijt} + \gamma X_{ijt} + \delta F_{ijt} + \varepsilon_{ijt} \quad (1)$$

where i is the student, j the school and t the year ($t = 0$ for 2011 and $t = 1$ for 2016); R_{ij} are students' standardised scores in reading; P_{ij} is a dummy variable that takes value "1" if the student attended preschool and "0" otherwise; X_{ijt} are students' background characteristics; F_{ijt} are family characteristics (parents' level of education, parents' occupation and number of books at home); ε_{ijt} is the idiosyncratic error term.

Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment" (Mullis & Martin, 2015, p. 12).

By estimating the model using ordinary least squares (OLS) we obtain the β coefficient that measures the influence that early childhood education has on students' reading achievement when they reach 4th grade. However, although we have controlled for certain variables within this model, it is virtually impossible to fully control for all unobservables that may affect academic performance as, e.g., students' preferences or ability. Since all omitted variables are included in ε_{ijt}, our β coefficient may be biased. In order to solve this problem, we have implemented an instrumental variable approach by using two-stage least squares (2SLS).

4.2. Two-stage least squares

To apply this methodology, first we must identify an instrument (Z_{ijt}) and use other control variables (X_{ijt}, F_{ijt}) that allow us to eliminate confounding factors correlated with both the treatment (preschool attendance) and the outcome (students' reading achievement). In this sense, the instrument to be used is the 2006 education reform. As mentioned in the Introduction section, the 2006 education reform has turned early childhood education upside down. Before the reform, second stage preschool education was exclusively free of charge in public schools with the possibility for each Autonomous Community (region) to reach certain agreements with semiprivate schools. Since the reform, the second level of ISCED 0 is universally free of charge, with the sole exception of those who voluntarily choose to enrol their children in private schools. In this context, data from PIRLS 2011 and PIRLS 2016 provides us with information on students who benefited from the reform and those who did not. Our instrument will be a binary variable (denoted as Z_{ijt}) that takes the value "1" for students from the PIRLS 2016 database (i.e. who benefited from free education) and "0" for those from the PIRLS 2011 database (i.e. who could not benefit from the reform).

This methodology can be properly applied if the chosen instrument follows these conditions:

(a) *Relevance* or *first stage condition*. The instrument should be strongly associated with the treatment variable (P_{ijt}). This correlation seems evident in our study. Better access to early childhood education from 3 to 6 years is a clear incentive to enrol children in preschool. A greater number of available schools offering free early childhood education may encourage parents to enrol their children in schools rather than, e.g., using relatives as childcare givers. In order to check whether this condition is met, we have also applied Stock and Yogo's (2005) weak instrument test, which is presented in the results section.

(b) *Independence/exogeneity assumption*. This condition indicates that the instrument has to be randomly assigned or "as good as randomly assigned", implying that the instrument is not correlated with the omitted variables we want to control for. The introduction of the education reform could be seen as a natural experiment that exogenously assigns individuals to different treatments (those who benefit from the reform and those who do not) and therefore creates a "natural" and exogenous variation in the data.

(c) *Exclusion restriction*. This condition states that there is a single channel (through P_{ijt}) for the influence of the instrument (Z_{ijt}) on the outcome (students' standardised reading scores, R_{ijt}). This requires that the independence assumption is satisfied to the extent that other potential channels have already been controlled for. Consequently, the exclusion restriction is satisfied in our research study when controlling for (X_{ijt}, F_{ijt}).

(d) *Monotonicity property*, also called as the *no defiers assumption*. Applying it to our study this assumption would mean that there are no students who have decided not to attend preschool even though their parents have decided to enrol them and, similarly, there are no students whose parents have decided not to enrol them in preschool even though they decided to attend. In this sense, and given the nature of our data, monotonicity property is always satisfied, since children could not make a different decision from the one chosen by their parents, as they are very young when their parents had to take the decision to enrol them in preschool.

Having identified the instrument, the model of equation (1) can be estimated by two-stage least squares (2SLS). The *first stage* is defined as:

$$P_{ijt} = \pi_0 + \pi_1 Z_{ijt} + \pi_2 X_{ijt} + \pi_3 F_{ijt} + \omega_{ijt} \quad (2)$$

being ω_{ijt} the idiosyncratic error term. Once this equation is estimated, we obtain a prediction of the dependent variable "attend preschool" ($\hat{P}_{ijt}$), which will be used in the model defined by equation (1) in place of P_{ijt} to create the *reduced form* (i.e. henceforth our base model):

$$R_{ijt} = \alpha + \beta \hat{P}_{ijt} + \gamma X_{ijt} + \delta F_{ijt} + \varepsilon_{ijt} \quad (3)$$

The β coefficient captures the influence of preschool attendance on students' reading achievement, which has been estimated using student weights, jackknife repeated replication weights and five plausible values (Martin & Mullis, 2011; Martin et al., 2017) to adjust the sample size to the population size.

The β coefficient estimated with the reduced form should not be biased, but we remain cautious and interpret our results as conditional associations rather than causal effects.

5. Results

5.1. Theoretical framework

Before providing the main results obtained in the empirical models, it is relevant to understand the potential theoretical mechanisms behind the figures reported in subsection 5.2. Specifically, from the field of psychology, the Ecological Theory (Bronfenbrenner, 2005) asserts that the key issue in children's development is the environment, which is a mix of children's interactions both with the classroom and family environments, being this theory consistent with the skills formation technology argument. Therefore, it is worth noting that cognitive and non-cognitive skills are determinants of school and socio-economic success and are not only conditioned by genetics, but also by the environment. This environment might be modified through early childhood interventions, such as preschool attendance, which may reduce educational gaps between students from different socio-economic backgrounds if a positive influence of preschool is found.

5.2. Main Results

Following the methodology explained in the previous section, we have first estimated the baseline model (Equation 1) using OLS. Table 1 shows that attending preschool has a positive impact on 4th grade students' reading achievement, being 0.16 SD for girls and 0.13 SD for boys, which seems to be in line with the evidence collected in previous literature (van den Brook et al., 2011; Cortázar, 2015; Bakken et al., 2017). Parents' educational level also has a positive impact on reading achievement, with a greater impact as education increases, which is far from surprising. Moreover, the number

of books available at home, as a proxy for students' socio-economic status, is also positively related to reading achievement. As for parents' occupation, it seems that in general mother's occupation has a stronger influence on reading achievement, while father's occupation only affects it when it comes to high responsibility positions.

-Insert Table 1 here-

In any case and, as mentioned above, when estimating the model using OLS an omitted variable bias is unavoidable. Our estimates may not be accurate and we may be overestimating or underestimating the extent to which these variables, especially attending preschool, affect reading achievement. In this sense, 2SLS estimates (Equation 3) are shown in Table 2. As we can observe, the influence that attending preschool has on reading achievement is much larger than what we obtained by OLS. Early childhood education affects girls' reading achievement by 1.44 SD and boys' by 0.94 SD, which is in line with other studies reporting greater benefits of early childhood education for girls than boys (Havnes & Mogstad, 2011; Felfe et al., 2015). Furthermore, the results of Stock and Yogo's (2005) test of weak instruments indicate that the instrument is strong. As for the influence of parents' educational level or occupation on academic performance, it seems to remain the same, except for father's occupation, which appears to have even a smaller influence than when we estimated the model by OLS.

-Insert Table 2 here-

Traditionally, and as a result of gender stereotypes that attribute greater mathematics skills to boys than to girls (Spencer et al. 1999; Spencer et al., 2016), girls tend to have a higher linguistic self-concept than boys (Heyder et al., 2017), which favours their reading performance. In this gender gap, early childhood education also seems to play an important role. Early interaction with peers and teachers may be a major transmission mechanism by which younger students may acquire gender stereotypes. If teachers maintain an outdated view of the society, hold a traditional gender role attitude and express their own beliefs about gender, children can quickly learn and reproduce them as they progress through the educational process (Tiedemann, 2002; Wolter et al., 2014; Retelsdorf et al., 2015).

5.3. Robustness Check

To check the robustness of our results, we have re-estimated our model (Equation 3) distinguishing between those students who have ever repeated a grade and those who have not (see Table A2, Appendix).

Although PIRLS database does not give us information on whether the student has repeated or not, it does give us students' age. As in Spain students born in the same calendar year start school in the same academic year, we can easily identify those students who are above the age that by law they should be in 4^{th} grade and, therefore, who should be those that have ever repeated. As we can observe, if we look at the results of Stock and Yogo's (2005) test of weak instruments, the instrument appears to be strong, even when we distinguish between repeaters and non-repeaters.

Preschool attendance positively affects 4^{th} grade students' reading achievement who have ever repeated in 1.89 SD, while this influence is 1.12 SD for those students who have not repeated. Based on the latest Programme for International Student Assessment (PISA) report (2018), 28.7% of students had repeated at least one grade, which is nearly three times the OECD average (OECD, 2020). Therefore, if early childhood education has a greater impact on the academic performance of those students who have been held back, providing quality early childhood education may have a

sufficient impact on these students to avoid this scenario. In short, the results are very similar to those obtained in our main model with early childhood education, also having a positive influence on academic performance, which proves the consistency of our results.

6. Discussion and conclusions

In this study, we have analysed the research question regarding whether preschool education (from 3 to 6 years old) has influence on Spanish boys' and girls' academic performance in reading when they are in 4th grade. To do so, we have used the database provided by PIRLS 2011 and 2016, using an instrumental variables approach to get closer to causality. As instrument, we have used the 2006 educational reform that took place in Spain, which promoted the universalisation of the second stage of ISCED 0 through free education in semiprivate and public schools.

In this context, our results show that going to preschool has a positive impact on reading achievement even when students reach primary school, with the impact being 1.44 SD for girls and 0.93 SD for boys. These results seem to be consistent with previous studies (Hidalgo-Hidalgo & García-Perez, 2012; Felfe et al., 2015; Cortázar, 2015; Datta & Simonsen, 2016; Ansari et al., 2017; Bai et al., 2020) that have tried to shed some light on whether early childhood education and early investment in children have benefits on academic performance not only in the short but also in the medium term. This positive relationship also holds when we distinguish between repeaters and non-repeaters and the corresponding socio-economic characteristics attached to them, which only reaffirms the consistency of our results. In this sense, preschool education has significant influence on academic performance throughout elementary school and can serve as a tool to identify those students who are at risk of falling behind in the educational process and to improve their performance for the future, acting as a driver of equal opportunities in education.

Therefore, the positive influence of preschool attendance can serve as an incentive to expand free access to the first stage of ISCED 0 in Spain. Beyond what this study can offer from an academic point of view, we are talking about a measure that has been the subject of political debate for years. The virtually free access to early childhood education promoted by the Spanish education system not only favours the process of skills formation, which guarantees future returns to human capital, but also confirms the effectiveness of a slightly controversial educational policy that boosts Spanish social progress. In other words, to the extent that the development of reading and oral language skills is highly dependent on the richness of the environment in which children grow up (Strickland et al., 2004; Kuhl, 2011), while parents can assist in the development of oral vocabulary and print awareness through everyday conversations or activities involving shared reading, the attendance to preschool will provide more specific instruction. In this sense, preschool contributes to the effectiveness of reading skills by ensuring that children have frequent and meaningful language interactions in early childhood. Access to books, organising daily interactive story times and creating many opportunities for conversation between adults and children are some of the ways in which preschool attendance can make students less likely to experience reading and academic difficulties later in life.

Taking into account how the environment (as a mix of children's interactions both with the classroom and family environments) works, early childhood interventions such as preschool attendance are supported by a theoretical framework that justifies this type of measures and constitutes a key research area for public policy interventions to promote a more egalitarian society.

In this sense, it would also be interesting to shed further light in future research works on whether the positive influence of free education, in first stage of ISCED 0, persists or, on the contrary, fades away after compulsory education. Furthermore, the reported evidence allows us to assert that, for those students who did not attend preschool, the educational system should provide compensation measures at an early age in primary education, to promote their development in terms of reading competence.

To conclude, with the present study we have contributed to the literature by providing quasi-experimental evidence on the influence of preschool attendance on reading achievement. However, it may have some limitations. First, there may be some unobservable variables that we have not been able to control for, which may influence reading achievement, so we should interpret our results as conditional associations rather than pure causal effects. Second, this research study has internal validity for Spain, but not external validity for other countries, so the particular case of each country has to be analysed. Third, PIRLS does not contain any information regarding dyslexia, speech and language problems or learning difficulties, which could affect students' progression in terms of reading competences. Furthermore, it does not include information regarding the quality of preschool education, which may differ depending on, e.g., physical and human conditions, territories' specific conditions or in the assessment methodologies of students' performance. Fourth, although the two-stage random sample design of PIRLS' assessment (i.e. firstly randomly sampling schools within countries and then random and complete classes within the schools) can be accounted by the use of student weights, jackknife repeated replication weights and five plausible values (as it is done in the present research study), it may be subject to survey data limitations (as any other survey), i.e. nonsampling errors (population coverage limitations, nonresponse bias, and measurement error, as well as data collection, processing, and reporting procedures), sampling errors (there is a degree of uncertainty associated with statistics estimated from a sample, which is usually expressed as the standard errors) and missing data (NCES, 2015).

References

Andrews, R., Jargowsky, P., & Kuhne, K. (2012). *The Effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance*. NBER Working Paper 18598, 1–20. doi: 10.3386/w18598

Ansari, A., López, M., Manfra, L., Bleiker, C., Dinehart, L., Hartman, S. C., & Winsler, A. (2017). Differential Third-Grade Outcomes Associated With Attending Publicly Funded Preschool Programs for Low-Income Latino Children. *Child development*, *88*(5), 1743–1756. doi: 10.1111/cdev.12663

Auger, A., Farkas, G., Burchinal, M. R., Duncan, G. J., & Vandell, D. L. (2014). Preschool center care quality effects on academic achievement: An instrumental variables analysis. *Developmental Psychology*, *50*(12), 2559–2571. doi: 10.1037/a0037995

Bai, Y., Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2020). Long-term effects of early childhood programs through eighth grade: Do the effects fade out or grow? *Children and Youth Services Review, 112,* 104890. doi: 10.1016/j.childyouth.2020.104890.

Bakken, L., Brown, N. & Downing, B. (2017) Early Childhood Education: The Long-Term Benefits. *Journal of Research in Childhood Education, 31*(2), 255–269. doi: 10.1080/02568543.2016.1273285

Berlinski, S., Galiani, S., & Gertler, P. (2009). The Effect of Pre-primary Education on Primary School Performance. *Journal of Public Economics, 93* (1-2), 219–234. doi: 10.1016/j.jpubeco.2008.09.002

Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education, 2*, 1163–1278. doi: 10.1016/S1574-0692(06)02020-4

BOE (1990). *Organic Law 1/1990, 3rd October, of the General Organisation of the Education System (LOGSE*). Nº 238. 4th October 1990, 28927 –28942. Spain.

BOE (2006). *Organic Law 2/2006, 3rd May, on Education (LOE*). Nº 106. 4th May 2006, 17158 – 17207. Spain.

Brilli, Y., Del Boca, D., & Pronzato, C. D. (2016). Does childcare availability play a role in maternal employment and children's development? Evidence from Italy. *Review of Economics of the Household,* 14, 27–51. doi: 10.1007/s11150-013-9227-4

Bronfenbrenner, U. (2005). The developing ecology of human development: Paradigm lost or paradigm regained. In U. Bronfenbrenner (Ed.), *Making human beings human: Bioecological perspectives on human development* (pp. 94–105). Thousand Oaks, CA: Sage.

Breda, T., & Napp, C. (2019). Girls' comparative advantage in reading can largely explain the gender gap in math-related fields. *Proceedings of the National Academy of Sciences*, *116*(31), 15435–15440. doi: 10.1073/pnas.1905779116

Card, D., & Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, *59*(1), 9–28. doi: 10.1111/ecin.12934

Corazzini, L., Meschi, E., & Pavese, C. (2021). Impact of early childcare on immigrant children's educational performance. *Economics of Education Review, 85*, 102181. doi: 10.1016/j.econedurev.2021.102181

Cortázar, A. (2015). Long-term effects of public early childhood education on academic achievement in Chile. *Early Childhood Research Quarterly, 32*, 13–22. doi: 10.1016/j.ecresq.2015.01.003

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American economic review, 97*(2), 31–47. doi: 10.1257/aer.97.2.31

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources, 43*(4), 738–782. doi: 10.3368/jhr.43.4.738

Datta, N., & Simonsen, M. (2016). Academic performance and type of early childhood care. *Economics of Education Review, 53*, 217–229. doi: 10.1016/j.econedurev.2016.03.013

DeLay, D., Hanish, L. D., Martin, C. L., & Fabes, R. A. (2016). Peer effects on Head Start children's preschool competency. *Developmental Psychology*, *52*(1), 58–70. doi: 10.1037/dev0000066

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *Journal of economic perspectives*, *27*(2), 109–132. doi: 10.1257/jep.27.2.109

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read:

Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*(3), 250–287. doi: 10.1598/RRQ.36.3.2

Felfe, C., Nollenberger, N., & Rodríguez-Planas, N. (2015). Can't buy mommy's love? Universal childcare and children's long-term cognitive development. *Journal of population economics, 28*(2), 393–422 (2015). doi:10.1007/s00148-014-0532-x

Flabbi, L. (2012). Gender Differences in Education, Career Choices and Labor Market Outcomes on a Sample of OECD countries. World Development *Report 2012*. World Development Report Background Papers.

Goldfeld, S., Quach, J., Nicholls, R., Reilly, S., Ukoumunne, O. C., & Wake, M. (2012). Four-year-old outcomes of a universal infant-toddler shared reading intervention: the Let's read trial. *Archives of pediatrics & adolescent medicine, 166*(11), 1045–1052. doi: 10.1001/archpediatrics.2012.1099

González-Betancor, S. M., & López-Puig, A. J. (2015). Escolarización temprana, trimestre de nacimiento y rendimiento educativo en primaria. *Revista de Educación, 369*, 159–183. doi: 10.4438/1988-592X-RE-2015-369-294

Havnes, T., & Mogstad, M (2011). No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy, 3*(2), 97–129. doi: 10.1257/pol.3.2.97

Heckman, J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative economics, 1*(1), 1–46. doi: 10.3982/qe8

Heckman, J. J., Stixrud, J., & Urzua, S. S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics, 24*(3), 411– 482. doi: 10.1086/504455

Heyder, A., Kessels, U., & Steinmayr, R. (2017). Explaining academic–Track boys' underachievement in language grades: Not a lack of aptitude but students' motivational beliefs and parents' perceptions? *British Journal of Educational Psychology, 87*(2), 205–223. doi: 10.1111/bjep.12145

Hidalgo-Hidalgo, M., & García-Pérez, J. I. (2012). Impacto de la asistencia a educación infantil sobre los resultados académicos del estudiante en primaria. In MECD (Ed.), *PIRLS - TIMSS 2011. Volumen II: Informe Español. Análisis Secundario* (pp. 83–113). Madrid: Instituto Nacional de Evaluación Educativa.

Jiménez, J., & O'Shanahan, I. (2008). Enseñanza de la lectura: de la teoría y la investigación a la práctica educativa. *Revista iberoamericana de educación, 45(*5), 1–22. doi: 10.35362/rie4552032

Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., Ruzek, E. A., & Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child development, 84*(4), 1171–1190. doi: 10.1111/cdev.12048

Knight, B. A., & Hughes, D. (1995). Developing social competence in the preschool years. *Australasian Journal of Early Childhood*, *20*(2), 13–19. doi: 10.1177/183693919502000204

Kuhl, P. K. (2011). Early language learning and literacy: neuroscience implications for education. *Mind, Brain, and Education, 5*(3), 128–142. doi: 10.1111/j.1751-228X.2011.01121.x

Li, L., Chen, X., Wu, D., & Li, H. (2020). Effects of attending preschool on adolescents' reading literacy: Evidence from the ethnic minority children in China. *Children and Youth Services Review*, *116*, 105211. doi: 10.1016/j.childyouth.2020.105211

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly, 45*, 155–176. doi: 10.1016/j.ecresq.2018.03.005.

Magnuson, K. A., Kelchen, R., Duncan, G. J., Schindler, H. S., Shager, H., & Yoshikawa, H. (2016). Do the effects of early childhood education programs differ by gender? A meta-analysis. *Early Childhood Research Quarterly, 36*, 521–536. doi: 10.1016/j.ecresq.2015.12.021

Marinak, B. A., & Gambrell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and instruction*, *49*(2), 129–141.doi: 10.1080/19388070902803795

Martin, C. L., Fabes, R. A., Hanish, L. D., & Hollenstein, T. (2005). Social dynamics in the preschool. *Developmental Review*, *25*(3-4), 299–327. doi: 10.1016/j.dr.2005.10.001

Martin, M. O., & Mullis, V. S. (2011). *Methods and Procedures in TIMSS and PIRLS 2011*. Boston: Lynch School of Education.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (2017). *Methods and Procedures in PIRLS 2016*. Boston: Lynch School of Education.

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., Yang, R., Koepp, A. E., & Shonkoff, J. P. (2017). Impacts of early childhood education on medium-and long-term educational outcomes. *Educational Researcher*, *46*(8), 474–487. doi: 10.3102/0013189X17737739

McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: Does sex or gender identity provide a better account? *Journal of Research in Reading*, *35*(3), 328–336. doi: 10.1111/j.1467-9817.2010.01481.x

Miller, L. C., & Bassok, D. (2019). The effects of universal preschool on grade retention. *Education Finance and Policy*, *14*(2), 149–177. doi: 10.1162/edfp_a_00248

Morabito, C., Figueroa, J. L., & Vandenbroeck, M. (2018). Effects of high versus low-quality preschool education: A longitudinal study in Mauritius. *Economics of Education Review*, *65*, 126–137.doi: 10.1016/j.econedurev.2018.06.006

Mullis, I. V. S., & Martin, M. O. (2015). *PIRLS 2016 Assessment Framework. 2nd Edition*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading.* Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/international-results/

Nakajima, N., Hasan, A., Jung, H., Brinkman, S., Pradhan, M., & Kinnell, A. (2019). Investing in school readiness: A comparison of different early childhood education pathways in rural Indonesia. *International Journal of Educational Development*, *69*, 22–38. doi: 10.1016/j.ijedudev.2019.05.009

NCES (2015). Trends in International Mathematics and Science Study (TIMSS). Retrieved from https://nces.ed.gov/timss/timss15technotes_datalim.asp

OECD (2018). *Education Policy Outlook 2018: Putting Student Learning at the Centre*. Paris: OECD Publishing. Retrieved from https://doi.org/10.1787/9789264301528-en.

OECD (2020). *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*. Paris: OECD Publishing.

Ölçer, S., & Aytar, A. G. (2014). A comparative study into social skills of five-six year old children and parental behaviors. *Procedia-Social and Behavioral Sciences, 141*, 976–995. doi: 10.1016/j.sbspro.2014.05.167

Osakwe, R. N. (2009). The effect of early childhood education experience on the academic performances of primary school children. *Studies on Home and Community Science*, *3*(2), 143–147.doi: 10.1080/09737189.2009.11885290

Pinto, A. I., Pessanha, M., & Aguiar, C. (2013). Effects of home environment and center-based childcare quality on children's language, communication, and literacy outcomes. *Early Childhood Research Quarterly*, *28*(1), 94–101. doi: 10.1016/j.ecresq.2012.07.001

Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with Grade 2 achievement. *Early Childhood Research Quarterly*, *28*(2), 437–449. doi: 10.1016/j.ecresq.2012.11.002

Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent Effects of Early Childhood Education on High-Risk Children and Their Mothers. *Applied Developmental Science, 4*(1), 2–14. doi: 10.1207/s1532480xads0401_1

Rendón Romero, S. I., García Jiménez, E., & Navarro Pablo, M. (2019). La lectura de palabras: la influencia del procesamiento fonológico y del método lectoescritor. *Revista Fuentes, 21*(1), 11–24. doi: 10.12795/revistafuentes.2018.v21.i1.01

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of educational psychology*, *107*(1), 186–194. doi: 10.1037/a0037107

Santin, D., & Sicilia, G. (2015). *El Impacto de la Educación Infantil en los Resultados de Primaria: Evidencia para España a partir de un Experimento Natural*. Madrid: Fundación Ramón Areces.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, *67*, 415–437. doi: 10.1146/annurev-psych-073115-103235

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28. doi: 10.1006/jesp.1998.1373

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80–108). New York: Cambridge University Press.

Strickland, D. S., Morrow, L. M., Neuman, S. B., Roskos, K., Schickedanz, J. A., & Vukelich, C. (2004). The role of literacy in early childhood education. *Reading Teacher, 58*(1), 86–103. doi: 10.1598/RT.58.1.9

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in mathematics*, *50*(1), 49–62. doi: 10.1023/A:1020518104346

van den Brook, P., Kendeou, P., Lousberg, S., & Visser, G. (2011). Preparing for reading comprehension: Fostering text comprehension skills in preschool and early elementary school children. *International Electronic Journal of Elementary Education*, *4*(1), 259-268.

van Huizen, T., & Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review*, *66*, 206–222. doi: 10.1016/j.econedurev.2018.08.001

Wolter, I., Glüer, M., & Hannover, B. (2014). Gender-typicality of activity offerings and child–teacher relationship closeness in German "Kindergarten". Influences on the development of spelling competence as an indicator of early basic literacy in boys and girls. *Learning and Individual Differences*, *31*, 59–65. doi: 10.1016/j.lindif.2013.12.008

Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., & Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Developmental psychology*, *49*(8), 1425–1439. doi: 10.1037/a0030347

Appendix

-Insert Table A1 here-

-Insert Table A2 here-

Table 1. Influence of preschool attendance on 4th graders' reading scores. Ordinary least squares

Variables	Total Sample	Females	Males
Female (ref.: male)			
Yes	0.07**	-	-
	(0.03)	-	-
Attendance to preschool (ref.: no)			
Yes	0.15***	0.16***	0.13***
	(0.03)	(0.04)	(0.04)
Father level of education (ref.: less than ISCED 2)			
ISCED 2	0.11***	0.09	0.12**
	(0.04)	(0.05)	(0.05)
ISCED 3	0.25***	0.23***	0.26***
	(0.04)	(0.06)	(0.06)
ISCED 5b	0.33***	0.36***	0.30***
	(0.04)	(0.06)	(0.06)
ISCED 5a, first degree	0.24***	0.28***	0.20***
	(0.05)	(0.07)	(0.07)
ISCED 5a, beyond first degree	0.42***	0.46***	0.38***
	(0.05)	(0.07)	(0.08)
Father education. Missing Flag	0.06	0.08	0.04
	(0.05)	(0.06)	(0.06)
Mother level of education (ref.: less than ISCED 2)			
ISCED 2	0.06*	0.02	0.10*
	(0.04)	(0.06)	(0.06)
ISCED 3	0.20***	0.16***	0.24***
	(0.04)	(0.06)	(0.07)
ISCED 5b	0.19***	0.15**	0.23***
	(0.05)	(0.07)	(0.08)
ISCED 5a, first degree	0.25***	0.21**	0.29***
	(0.06)	(0.09)	(0.08)
ISCED 5a, beyond first degree	0.39***	0.37***	0.40***
	(0.06)	(0.09)	(0.08)
Mother education. Missing Flag	0.03	0.01	0.06
	(0.04)	(0.07)	(0.07)
Father's occupation (ref.: has never worked for pay)			
Small business owner	0.21	-0.02	0.39
	(0.15)	(0.16)	(0.24)
Clerical worker	0.16	0.02	0.26
	(0.16)	(0.16)	(0.25)
Service or sales worker	0.25*	-0.01	0.45*
	(0.15)	(0.15)	(0.24)
Skilled agricultural or fishery worker	0.15	-0.10	0.33
	(0.16)	(0.17)	(0.26)
Craft or trade worker	0.31**	0.09	0.48**
	(0.13)	(0.15)	(0.21)
Plant or machine operator	0.14	-0.06	0.27
	(0.14)	(0.15)	(0.22)
General labourers	0.22	0.00	0.39
	(0.16)	(0.16)	(0.25)
Corporate manager or senior official	0.25*	0.02	0.44*
	(0.14)	(0.17)	(0.22)
Professional	0.34**	0.08	0.54**
	(0.15)	(0.15)	(0.24)

Technician or associate professional	0.32**	0.08	0.51*
	(0.16)	(0.16)	(0.26)
Father's occupation. Missing flag	0.12	-0.18	0.36
	(0.14)	(0.15)	(0.23)
Mother's occupation (ref.: has never worked for pay)			
Small business owner	0.21***	0.27**	0.14
	(0.07)	(0.11)	(0.09)
Clerical worker	0.30***	0.28***	0.32***
	(0.07)	(0.09)	(0.10)
Service or sales worker	0.32***	0.34***	0.30***
	(0.06)	(0.08)	(0.09)
Skilled agricultural or fishery worker	0.43***	0.43***	0.44***
	(0.11)	(0.14)	(0.15)
Craft or trade worker	0.13	0.09	0.18
	(0.14)	(0.21)	(0.17)
Plant or machine operator	0.20*	0.26*	0.16
	(0.12)	(0.14)	(0.18)
General labourers	0.23***	0.22**	0.26**
	(0.07)	(0.09)	(0.11)
Corporate manager or senior official	0.26***	0.29***	0.24*
	(0.08)	(0.10)	(0.14)
Professional	0.35***	0.35***	0.36***
	(0.07)	(0.10)	(0.10)
Technician or associate professional	0.34***	0.36***	0.33***
	(0.08)	(0.10)	(0.11)
Mother's occupation. Missing flag	0.14*	0.14	0.13
	(0.07)	(0.09)	(0.10)
Number of books at home (ref.: less than 11)			
11 to 25 books	0.36***	0.22***	0.46***
	(0.04)	(0.06)	(0.07)
26 o 100 books	0.57***	0.44***	0.67***
	(0.05)	(0.06)	(0.07)
101 o 200 books	0.63***	0.54***	0.69***
	(0.05)	(0.07)	(0.07)
More than 200 books	0.59***	0.49***	0.65***
	(0.06)	(0.07)	(0.08)
Number of books at home. Missing Flag	-0.26**	-0.34	-0.20
	(0.10)	(0.21)	(0.15)
Constant	-1.56***	-1.14***	-1.83***
	(0.16)	(0.16)	(0.24)
Observations	19,337	9,762	9,575
R-squared	0.19	0.20	0.19

Notes: Standard errors are in parenthesis and PIRLS recommended practices have been applied.
Estimation method: Ordinary Least Squares (OLS).
Dependent variable: Students' standardised scores in reading.
Coefficient: *** significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Table 2. Influence of preschool attendance on 4th graders' reading scores. Instrumental variables

Variables	Total Sample	Females	Males
Female (ref.: male)			
Yes	0.08***	-	-
Attendance to preschool (ref.: no)			
Yes	1.18***	1.44***	0.94***
	(0.24)	(0.26)	(0.31)
Father level of education (ref.: less than ISCED 2)			
ISCED 2	0.12**	0.11	0.13**
	(0.05)	(0.08)	(0.05)
ISCED 3	0.24***	0.16**	0.29***
	(0.05)	(0.08)	(0.07)
ISCED 5b	0.27***	0.28***	0.27***
	(0.05)	(0.08)	(0.06)
ISCED 5a, first degree	0.24***	0.29***	0.19**
	(0.06)	(0.09)	(0.08)
ISCED 5a, beyond first degree	0.40***	0.40***	0.39***
	(0.07)	(0.10)	(0.08)
Father education. Missing Flag	0.11**	0.11	0.11*
	(0.06)	(0.08)	(0.06)
Mother level of education (ref.: less than ISCED 2)			
ISCED 2	0.06	0.05	0.07
	(0.04)	(0.07)	(0.06)
ISCED 3	0.16***	0.16**	0.17**
	(0.04)	(0.07)	(0.07)
ISCED 5b	0.08	0.00	0.14*
	(0.06)	(0.09)	(0.08)
ISCED 5a, first degree	0.16**	0.10	0.21***
	(0.06)	(0.10)	(0.08)
ISCED 5a, beyond first degree	0.28***	0.25**	0.31***
	(0.06)	(0.11)	(0.08)
Mother education. Missing Flag	0.05	0.07	0.04
	(0.05)	(0.08)	(0.07)
Father's occupation (ref.: has never worked for pay)			
Small business owner	0.13	0.04	0.26
	(0.14)	(0.14)	(0.23)
Clerical worker	0.06	0.04	0.11
	(0.15)	(0.15)	(0.25)
Service or sales worker	0.17	0.03	0.33
	(0.13)	(0.13)	(0.23)
Skilled agricultural or fishery worker	0.14	0.02	0.26
	(0.15)	(0.15)	(0.25)
Craft or trade worker	0.25**	0.11	0.40**
	(0.12)	(0.13)	(0.19)
Plant or machine operator	0.05	-0.02	0.13
	(0.13)	(0.14)	(0.21)
General labourers	0.12	-0.01	0.26
	(0.14)	(0.15)	(0.23)
Corporate manager or senior official	0.20	0.13	0.30
	(0.13)	(0.14)	(0.21)
Professional	0.24*	0.09	0.40*

	(0.14)	(0.13)	(0.24)
Technician or associate professional	0.27**	0.14	0.42*
	(0.14)	(0.15)	(0.24)
Father's occupation. Missing flag	0.07	-0.07	0.24
	(0.13)	(0.12)	(0.22)
Mother's occupation (ref.: has never worked for pay)			
Small business owner	0.26***	0.33**	0.19*
	(0.09)	(0.13)	(0.11)
Clerical worker	0.27***	0.25**	0.29**
	(0.08)	(0.10)	(0.12)
Service or sales worker	0.34***	0.38***	0.31***
	(0.07)	(0.09)	(0.10)
Skilled agricultural or fishery worker	0.40***	0.37**	0.42***
	(0.11)	(0.17)	(0.14)
Craft or trade worker	0.13	0.12	0.15
	(0.14)	(0.21)	(0.19)
Plant or machine operator	0.15	0.22	0.10
	(0.13)	(0.15)	(0.20)
General labourers	0.26***	0.29***	0.25**
	(0.09)	(0.11)	(0.11)
Corporate manager or senior official	0.25***	0.28**	0.22
	(0.08)	(0.12)	(0.15)
Professional	0.33***	0.32***	0.34***
	(0.08)	(0.12)	(0.11)
Technician or associate professional	0.30***	0.31***	0.29**
	(0.09)	(0.11)	(0.12)
Mother's occupation. Missing flag	0.17**	0.19*	0.15
	(0.08)	(0.10)	(0.11)
Number of books at home (ref.: less than 11)			
11 to 25 books	0.30***	0.18***	0.40***
	(0.05)	(0.06)	(0.08)
26 o 100 books	0.51***	0.38***	0.62***
	(0.06)	(0.07)	(0.09)
101 o 200 books	0.54***	0.45***	0.62***
	(0.06)	(0.07)	(0.08)
More than 200 books	0.51***	0.42***	0.58***
	(0.07)	(0.07)	(0.09)
Number of books at home. Missing Flag	-0.35***	-0.41*	-0.30*
	(0.10)	(0.22)	(0.15)
Constant	-2.17***	-2.04***	-2.22***
	(0.17)	(0.21)	(0.21)
Observations	19,337	9,762	9,575
Stock and Yogo (2005) test of weak instruments	314.67***	139.40***	178.70***

Notes: Standard errors are in parenthesis and PIRLS recommended practices have been applied.
Estimation method: Two-Stage Least Squares (2SLS). The instrument is the educational reform. The null hypothesis of the Stock and Yogo (2005) test of weak instruments is that the instrument is weak.
Dependent variable: Students' standardised scores in reading.
Coefficient: *** significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Table A1. Descriptive statistics

Variables		Total Sample		Attending Preschool		Not attending Preschool	
		Mean	SD	Mean	SD	Mean	SD
Standardised reading scores		0	1	0.118*	0.973	-0.236*	1.01
Preschool Attendance	Yes	0.747	0.435	-	-	-	-
	No	0.253	0.435	-	-	-	-
Sex of the student	Male	0.506	0.50	0.50	0.50	0.489	0.50
	Female	0.494	0.50	0.499	0.50	0.511	0.50
Repeater	Yes	0.095	0.294	0.072*	0.265	0.124*	0.306
	No	0.905	0.294	0.928*	0.265	0.876*	0.306
Level of education of the father	Did not go to school	0.010	0.10	0.008*	0.091	0.018*	0.123
	ISCED 1	0.142	0.349	0.126*	0.339	0.175*	0.352
	ISCED 2	0.234	0.424	0.222*	0.425	0.295*	0.422
	ISCED 3	0.251	0.433	0.251	0.443	0.257	0.404
	ISCED 5b	0.094	0.293	0.099*	0.306	0.063*	0.225
	ISCED 5a, first degree	0.107	0.309	0.113*	0.323	0.089*	0.264
	ISCED 5a, beyond first degree	0.161	0.367	0.180*	0.393	0.103*	0.281
Level of education of the mother	Did not go to school	0.01	0.098	0.009*	0.093	0.015*	0.112
	ISCED 1	0.120	0.325	0.098*	0.304	0.159*	0.338
	ISCED 2	0.204	0.403	0.188*	0.40	0.279*	0.415
	ISCED 3	0.242	0.428	0.238*	0.436	0.264*	0.408
	ISCED 5b	0.098	0.298	0.101*	0.308	0.062*	0.225
	ISCED 5a, first degree	0.147	0.354	0.161*	0.375	0.107*	0.287
	ISCED 5a, beyond first degree	0.179	0.383	0.205*	0.412	0.112*	0.292
Occupation of the father	Has never worked for pay	0.010	0.101	0.008*	0.091	0.017*	0.119
	Small business owner	0.145	0.353	0.142	0.357	0.152	0.331
	Clerical worker	0.048	0.213	0.051*	0.226	0.038*	0.177
	Service or sales worker	0.114	0.318	0.113	0.324	0.125	0.305
	Skilled agricultural or fishery worker	0.036	0.186	0.031*	0.177	0.052*	0.205
	Craft or trade worker	0.171	0.377	0.16*	0.375	0.209*	0.375
	Plant or machine operator	0.114	0.318	0.11	0.32	0.114	0.293
	General labourers	0.070	0.255	0.067	0.257	0.078	0.247
	Corporate manager or senior official	0.084	0.277	0.087	0.288	0.074	0.242
	Professional	0.16	0.366	0.18*	0.393	0.097*	0.273
	Technician or associate professional	0.048	0.214	0.049	0.221	0.044	0.189
Occupation of the mother	Has never worked for pay	0.053	0.225	0.049*	0.222	0.071*	0.238
	Small business owner	0.089	0.285	0.081*	0.279	0.113*	0.292
	Clerical worker	0.178	0.383	0.189*	0.40	0.148*	0.328
	Service or sales worker	0.195	0.396	0.179*	0.392	0.233*	0.391
	Skilled agricultural or fishery worker	0.017	0.130	0.015	0.125	0.020	0.13
	Craft or trade worker	0.016	0.125	0.015	0.125	0.018	0.123
	Plant or machine operator	0.035	0.183	0.034	0.185	0.032	0.162
	General labourers	0.128	0.334	0.115*	0.326	0.174*	0.35
	Corporate manager or senior official	0.035	0.184	0.038*	0.196	0.028*	0.153
	Professional	0.197	0.397	0.222*	0.425	0.122*	0.303
	Technician or associate professional	0.057	0.232	0.062*	0.247	0.040*	0.18
Books at home	0-10	0.098	0.298	0.077*	0.274	0.131*	0.313
	11-25	0.277	0.447	0.262*	0.450	0.307*	0.427
	26-100	0.339	0.473	0.349	0.488	0.33	0.436
	101-200	0.163	0.369	0.175*	0.389	0.131*	0.313
	More than 200	0.123	0.329	0.137*	0.352	0.10*	0.278

Notes: PIRLS recommended practices have been applied. "S.D." indicates "standard deviation". Test for mean differences between: attending preschool and not attending preschool. ***Significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Table A2. Influence of preschool attendance on 4th graders' reading scores distinguishing between repeaters and non-repeaters. Instrumental variables

Variables	Repeaters	Non-Repeaters
Female (ref.: male)		
Yes	-0.04	0.08***
Attendance to preschool (ref.: no)		
Yes	1.89***	1.12***
	(0.36)	(0.25)
Father level of education (ref.: less than ISCED 2)		
ISCED 2	0.48***	0.07
	(0.16)	(0.05)
ISCED 3	0.56***	0.18***
	(0.17)	(0.06)
ISCED 5b	0.93***	0.21***
	(0.26)	(0.06)
ISCED 5a, first degree	1.05***	0.16***
	(0.28)	(0.06)
ISCED 5a, beyond first degree	0.88***	0.33***
	(0.31)	(0.07)
Father education. Missing Flag	0.58***	0.06
	(0.20)	(0.06)
Mother level of education (ref.: less than ISCED 2)		
ISCED 2	-0.04	0.04
	(0.16)	(0.05)
ISCED 3	-0.18	0.15***
	(0.19)	(0.05)
ISCED 5b	-0.01	0.07
	(0.21)	(0.06)
ISCED 5a, first degree	-0.20	0.15**
	(0.25)	(0.06)
ISCED 5a, beyond first degree	-0.13	0.28***
	(0.31)	(0.06)
Mother education. Missing Flag	-0.23	0.06
	(0.19)	(0.06)
Father's occupation (ref.: has never worked for pay)		
Small business owner	-0.08	0.06
	(0.37)	(0.14)
Clerical worker	-0.35	-0.02
	(0.59)	(0.15)
Service or sales worker	0.42*	0.08
	(0.24)	(0.14)
Skilled agricultural or fishery worker	0.67*	0.03
	(0.36)	(0.16)
Craft or trade worker	0.48*	0.15
	(0.26)	(0.14)
Plant or machine operator	0.55*	-0.08
	(0.28)	(0.14)
General labourers	0.32	0.02
	(0.31)	(0.16)
Corporate manager or senior official	0.36	0.11
	(0.29)	(0.14)
Professional	0.10	0.16
	(0.35)	(0.15)
Technician or associate professional	0.16	0.19

	(0.35)	(0.14)
Father's occupation. Missing flag	0.23	0.00
	(0.27)	(0.14)
Mother's occupation (ref.: has never worked for pay)		
Small business owner	0.79**	0.18**
	(0.35)	(0.09)
Clerical worker	0.64**	0.19**
	(0.27)	(0.08)
Service or sales worker	0.76***	0.25***
	(0.25)	(0.07)
Skilled agricultural or fishery worker	0.49*	0.34***
	(0.29)	(0.12)
Craft or trade worker	-0.20	0.19
	(0.50)	(0.14)
Plant or machine operator	0.50	0.08
	(0.52)	(0.13)
General labourers	0.53**	0.20**
	(0.26)	(0.09)
Corporate manager or senior official	0.35	0.18**
	(0.34)	(0.09)
Professional	0.58	0.26***
	(0.36)	(0.08)
Technician or associate professional	0.40	0.23**
	(0.44)	(0.09)
Mother's occupation. Missing flag	0.45**	0.10
	(0.18)	(0.08)
Number of books at home (ref.: less than 11)		
11 to 25 books	0.38***	0.26***
	(0.14)	(0.05)
26 o 100 books	0.38**	0.49***
	(0.15)	(0.06)
101 o 200 books	0.73***	0.50***
	(0.16)	(0.06)
More than 200 books	0.15	0.50***
	(0.20)	(0.06)
Number of books at home. Missing Flag	-0.84**	-0.24*
	(0.36)	(0.14)
	-3.40***	-1.86***
Constant	(0.26)	(0.22)
Observations	1,633	17,680
Stock and Yogo (2005) test of weak instruments	53.12***	266.47***

Notes: Standard errors are in parenthesis and PIRLS recommended practices have been applied.
Estimation method: Two-Stage Least Squares (2SLS). The instrument is the educational reform. The null hypothesis of the Stock and Yogo (2005) test of weak instruments is that the instrument is weak.
Dependent variable: Students' standardised scores in reading.
Coefficient: *** significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.