This is a post-peer-review, pre-copyedit version of an article published in Educational Assessment, Evaluation and Accountability. The final authenticated version is available online at: https://doi.org/10.1007/s11092-023-09421-6

Grade retention in Spain: the right way?

Luis Alejandro Lopez-Agudo[1]

Claudia Prieto-Latorre[2]

Oscar D. Marcenaro-Gutierrez[3]

Abstract

There is a wide debate on the convenience of grade retention for students' future cognitive development. Nevertheless, due to the endogenous characteristics of grade retention to explain students' academic performance, most of the literature fails in capturing its actual influence. In the present research we intend to get as close as possible to the causal effect of grade retention on lower secondary education students' academic competences. In order to do this, we employ rich administrative census and longitudinal data from the most populated region of Spain (Andalusia) using Ordinary Least Squares (OLS) estimation method with student fixed-effects within-students between-academic years. Our results show a positive influence of grade retention on students' academic competences in the short-term in mathematics, but a null one in reading.

Keywords: grade retention; student fixed-effects; lower secondary education; reading; mathematics.

JEL Codes: I20, I21, I28.

Conflict of Interest forms/ethical statements: The authors do not have any conflict of interest.

Data availability statement: The data was provided by the Andalusian Agency of Education Assessment to the authors for their exclusive use.

[1] Departamento de Economía Aplicada (Estadística y Econometría), Facultad de Ciencias Económicas y Empresariales, Universidad de Málaga, C/ Ejido, 6, 29013, Málaga, España. Tel.: +34 95 213 7003; Fax: +34 95 213 7262. Email: lopezagudo@uma.es. ORCID: 0000-0002-0906-3206

[2] Departamento de Economía Aplicada (Estadística y Econometría), Facultad de Ciencias Económicas y Empresariales, Universidad de Málaga, C/ Ejido, 6, 29013, Málaga, España. Tel.: +34 95 213 7003; Fax: +34 95 213 7262. Email: claudiaprieto@uma.es. ORCID: 0000-0002-6510-3057

[3] Departamento de Economía Aplicada (Estadística y Econometría), Facultad de Ciencias Económicas y Empresariales, Universidad de Málaga, C/ Ejido, 6, 29013, Málaga, España. Tel.: +34 95 213 7003; Fax: +34 95 213 7262. E-mail: odmarcenaro@uma.es. ORCID: 0000-0003-0939-5064

1. Introduction

There is a wide debate on the way to manage the academic progression of those students who do not reach the minimum level of knowledge required for the grade they are taking, an issue which is differently faced by each country. Some countries bet for automatic promotion and support courses in future grades to compensate these students' potential lack of skills or knowledge; others stand up for alternative academic or vocational tracks, while others support grade retention as a way to help these students to acquire the skills they missed the first time they took that grade. For instance, there are some countries like Norway or Iceland where grade retention does not exist and students are automatically promoted to the next grade; there are others like the United Kingdom in which students do not repeat but are placed – only in exceptional circumstances – in "out of year-group", while in many others students may repeat depending on the amount of subjects they failed – e.g. Greece, Cyprus, Portugal, France, Spain, etc. – (Eurydice, 2011). It is also interesting to note that a high variation in the likelihood of grade retention that students present is influenced by country level factors – i.e., national educational policy factors, traditions and societal beliefs regarding the benefits of grade (Goos, Schreier, Knipprath, De Fraine, Van Damme, & Trautwein, 2013). In fact, there are wide differences in terms of grade retention rates by country (e.g., following PISA 2015, Greece presents a grade retention rate of 5%, Cyprus 4.7%, Portugal 31.2%, France 22.1% and Spain 31.3%; OECD, 2016). To the extent that Spain exhibits one of the greatest grade retention rates among European countries, the grade retention policy of this education system will be the focus of the present research.

In spite of its relevance, grade retention is not an easy topic to study when it comes to analysing its influence on students' academic performance. This is because of the endogenous characteristics that grade retention presents in this relationship, which makes difficult its identification as a cause or as an effect of students' lower academic performance. In this sense, most studies fail in determining whether grade retention is the "effect" of certain students' characteristics that make them perform lower, fail and repeat or, alternatively, a "cause" of students' lower academic performance. Especially, the lack of longitudinal data to analyse this issue makes it even more difficult to solve.

Most of the existing literature points towards a negative association between grade retention and students' academic performance, presenting a long-lasting influence that may be translated into school dropout. In this way, Andrew (2014) analysed grade retention in primary education for United States students and found that it negatively influenced high school completion (due to higher dropout) and that students who overcame this drawback did it earlier than those who did not. Glick and Sahn (2010) also analysed this issue of grade retention for primary education students in Senegal, finding that students are more likely to dropping out their studies when they are hold back. Manacorda (2012) employed a discontinuity in grade failure legislation (automatic grade retention with three failed subjects) for students in Uruguay and found that it increased dropout and produced long lasting lower academic achievement. Likewise, Jacob and Lefgren (2009) found that grade retention did not influence the likelihood of high school completion of students in Chicago, but making low-achieving students in primary education repeat increased their likelihood of dropping out high school. For students in this same city, Ou and Reynolds (2010) analysed the relationship between grade retention and participation in postsecondary education, finding that those who repeated later had higher likelihood of not finishing postsecondary education. Hwang and Cappella (2018) analysed United States students using propensity score matching and

found that those students who repeated performed lower in reading six or seven years later, although grade retention did not have any influence in other academic or psychological outcomes.

In addition, grade retention has been found to have influence on other outputs. In this way, Pagani, Tremblay, Vitaro, Boulerica, and McDuff (2001) analysed the influence that grade retention could have on students in Quebec from primary education to secondary education, finding a short and long-term negative influence on students' academic performance, increasing also children's anxiety and disruptive behaviour. Other authors as Jimerson and Ferguson (2007) analysed a longitudinal sample of western United States students from kindergarten to secondary education and found that grade retention did not improve students' academic outcomes and that retained students showed more aggressive behaviour. Özek (2015) analysed third grade students in Florida using regression discontinuity design, finding that grade retention was positively associated with higher likelihood of disciplinary incidents and suspensions in the short run, but not in the long term. Other authors such as Martorell and Mariano (2018) employed a fuzzy regression discontinuity design for elementary and middle school students in New York, and found that grade retention does not seem to increase behavioral problems as measured by absences and suspensions. Mariano, Martorell, and Tsai (2018) employed regression discontinuity for New York students in middle school and found that grade retention increased dropout rates and reduced the likelihood of completing Regents exam (i.e. exams to graduate from high school) graduation requirements. Diaz, Grau, Reyes, and Rivera (2021) analysed grade retention for primary education students in Chile using regression discontinuity, finding that early grade retention had a negative influence on the probability of committing a crime, while grade retention in late primary education increased this probability.

There is another strand of the literature which indicates that grade retention does not seem to influence students' academic outcomes. For instance, Allen, Chen, Willson, and Hughes (2009) performed a meta-analysis of 22 studies on grade retention and found that it seems to have a null association with academic performance. Likewise, Jimerson (2001) performed a meta-analysis on 20 studies and indicated that grade retention does not show better results than automatic promotion, and that alternative remedial strategies have to be found – this was also indicated in further research by Jimerson and Kaufman (2003) and Jimerson, Pletcher, Graydon, Schnurr, Nickerson, and Kundert (2005). However, the latter two meta-analyses present some issues which may make their results difficult to go beyond correlation: they gather research for students from different grades (kindergarten to 12th grade), with a mixed research quality in terms of methods, and most studies come from two or three decades ago.

Belot and Vandenberghe (2011) analysed a policy reform in Belgium which reinstated the possibility that students repeat – using data from the Programme for International Student Assessment (PISA) – which may suppose a threat for students who have low scores; nevertheless, they found that this exogenous legislation did not have any influence on students' scores. Similarly, Goos, Pipa, and Peixoto (2021) performed a meta-analysis of 84 studies from 2000 to 2019 and also indicated a null influence of grade retention on students' academic performance. In terms of other outputs, Eide and Showalter (2001) studied the influence of grade retention on dropping out and labour market earnings of United States students using an instrumental variable

approach (using as instrument an exogenous variation across states in kindergarten entry dates); they found that it did not have any significant influence on these outcomes.

Finally, research works such as that by Schwerdt, West, and Winters (2017) analysed third grade students from Florida using regression discontinuity and found that grade retention seemed to have a positive influence on students' academic outcomes in reading and mathematics, reducing the remedial courses that students had to take in high school, although grade retention did not have any influence on their probability of graduating. Similarly, Tafreschi and Thiemann (2016) analysed higher education students in Switzerland using regression discontinuity analysis, and found that grade retention was positively associated with a 10% higher likelihood of dropout, but for those who did not dropout their scores were 0.5 standard deviations higher after grade retention.

As we can see, many research works have analysed the influence of grade retention on students' outcomes, finding mixed results depending on the country and the methodology employed, which range from correlational to almost causal. However, although evidence is increasing in the case of Spain, unfortunately, it is based mostly on correlational research, due to a lack of longitudinal education data in this country. In this sense, Calero, Choi, and Waisgrais (2010) studied PISA 2006 Spanish data for secondary education students using a logistic multilevel model and found that grade retention is positively associated with school dropout. Carabaña (2013) analysed PISA 2012 data for Spanish secondary education students using Ordinary Least Squares analysis and found that grade retention is associated with cognitive factors (academic performance) and with non-cognitive factors (e.g. different languages in the classroom and the school, age, etc.). He also found that grade retention in primary education is the best predictor of grade retention in secondary education. Gortazar (2019) used descriptive analysis and highlighted that there is a relationship between grade retention and students' dropout in Spain which negatively influenced students' future career development. To the extent that the Spanish education system presents its own characteristics in terms of grade retention (as it will be described below) more research – that goes beyond correlation – on this issue would be needed.

Nevertheless, there are some relevant approaches which intended to overcome this lack of longitudinal Spanish education data. The first is that of García-Pérez, Hidalgo-Hidalgo, and Robles-Zurita (2014), who analysed this issue for secondary education students using PISA 2009 – cross sectional – data, implementing a switching regression model using the quarter of birth as instrument. They indicated that grade retention presents a negative influence on students' academic achievement, which becomes higher when students have repeated both in primary and secondary education. However, the imputation procedure employed by PISA to provide cognitive test scores in every subject to all students (even those which they did not take) may influence the validity[4] of the employed instrument[5] – the quarter of birth. In addition, Buckles and Hungerman (2013) indicated that seasonality of birth may make the quarter of birth to fail the exclusion restriction when used as instrumental variable, as family background characteristics may have strong relation with both season of birth and later outcomes. Choi, Gil, Mediavilla, and Valbuena (2016) combined cross-sectional data from PISA 2012 and the Progress in International Reading Literacy Study 2006 (PIRLS) to create a pseudo-panel. They found a negative association of grade retention on Spanish students'

[4] This validity requirement means that the instrument does not have to be correlated with the error term.

[5] More information on this imputation issue can be found in Jerrim, Lopez-Agudo, Marcenaro-Gutierrez, and Shure (2017).

academic performance; nevertheless, they employed an imputation procedure – based on students with similar background characteristics – to match PISA 2012 students with students' scores from PIRLS 2006, as the same students cannot be followed in PISA 2012. Therefore, matching both databases is based on a set of variables which are common to both and do not account for all students' characteristics (e.g. students' unobserved ability) which, joined to a different sampling procedure of both databases[6], would make this identification of students very difficult to work out. In addition, both studies do not employ the appropriate counterfactual (the former does not use same-age students and the latter does not follow the same student) when analysing grade retention effects, which may also undermine their contribution.

In spite of these two examples for Spain, which present some issues in their methodology, the lack of longitudinal data in the Spanish context has made almost impossible to analyse the influence that grade retention may have on students' academic performance due to the endogenous characteristics of the grade retention variable in this relationship. Due to the difficulties in dealing with this issue, most of the research focused on the Spanish education system intends to explore the determinants of grade retention; nevertheless, these research works present similar econometric problems due to using cross-sectional data. For instance, as grade retention determinants, some research works point towards differences in early-acquired skills or socio-economic characteristics (Peraita & Pastor, 2000; Agasisti & Cordero, 2017); others point to a combination of socio-economic characteristics and immigrant status (Calero, Choi, & Waisgrais, 2010) or to socio-economic status and birth date (Pedraja-Chaparro, Santín, & Simancas, 2015; González-Betancor & López-Puig, 2016).

In this particular context, the current research focuses on analysing the influence of grade retention on students' academic achievement for the Spanish region of Andalusia. This region presents some characteristics which make its study of great interest. First, it is the most populated region in Spain and also one of the worst performing according to international large-scale assessment tests. Specifically, PISA 2015 showed that Andalusian students performed 19 standardised points under the Spanish average and 11.7 points under the OECD average (OECD, 2016). In addition, Andalusia presented one of the highest grade retention rates of all Spanish regions in PISA 2015 (38% of students repeated at least once in primary, lower or upper secondary education, which is 7% higher than the rate of Spain and 25% than the OECD). Moreover, Andalusia presented a dropout rate in compulsory education of 24.9% in 2015, overcoming that of Spain in 4.9% (IECA, 2019). More information on Andalusia and a comparison with Spain and the OECD countries can be found in Table A1 (Appendix). The Spanish compulsory education system consists of two stages: primary (1st to 6th grade) and secondary education (7th to 10th grade), which go from ages 6 to 15. However, to the extent that grade retention is allowed in Spain, students may finish their compulsory education at an older age in case they repeat one or more grades. In particular, students can repeat only once in primary education and twice in secondary education (BOE, 2006). Regarding the conditions for grade retention in Spain, the decision of whether a student has to repeat or not is based on an objective criterion (as indicated by BOE, 2006): students failing a minimum of three subjects (i.e. not achieving a minimum score of "5", also called a "pass") will have to repeat (or when failing two subjects, if they are reading

[6] PISA uses a two-stage sampling procedure based on randomly choosing schools and then random students within schools, whereas PIRLS presents a two-stage sampling procedure based on randomly choosing schools and then random and complete classes within the school

and mathematics). Parents and teachers cannot influence this result, as it is based on an objective criterion, so there is not any possibility of veto of the grade retention decision by parents; therefore, teachers gather students' results and communicate this grade retention in case students fulfil its conditions.

The present work is novel to the extent that we analyse the influence that grade retention has on Andalusian lower secondary education students' academic competences in the short-term going further from correlation and getting as close as possible to its causal effect. To do this, census and longitudinal data, together with Ordinary Least Squares (OLS) estimation method with student fixed-effects within-student between-academic years (2011-12/2012-13) methodology, have been employed, using the student in different academic years as his/her own counterfactual. Hence, the research question that we try to answer is:

Can grade retention put Spanish students who failed in lower secondary education again on track in the short-term?

The rest of the paper is structured as follows: first, we describe the data and methodology employed for the present research; then, the main results are presented, discussed and closed by some conclusions.

2. Data

The data we use are based upon a census assessment of Andalusian children conducted by the Andalusian Agency of Education Assessment (AGAEVE from now on) to evaluate students' competences. This supposes that each child completes a diagnostic assessment (*Evaluación de Diagnóstico*) evaluating their competences in Spanish language (reading from now on)[7] and mathematics[8]. The use of competences instead of test scores administrative data and the correction of cognitive tests by external teachers are actually positive elements of this data, to the extent that we measure students' skills and avoid potential inflation on students' test scores by teachers. This is especially relevant in our research, as we are focusing on repeaters and teachers may be tempted to increase their scores because they cannot repeat twice the grade under analysis (8th grade; BOE, 2006), so teachers are, in some way, "obligated" to pass these repeater students to 9th grade. In addition, these competence scores are used to report regional education results and not for the decision of making students pass to the next grade or not.

The data employed in the present research are those of the cohort of students in 8th grade (the second year of secondary school) in the academic years 2011/12 and 2012/13. The use of these two cohorts has been done in order to identify those students who were in 8th grade in 2011-12 but failed that grade and repeated it in 2012-13 (which is key for our approach, as we will see in the Methodology section)[9]. Furthermore, all students and their parents completed background questionnaires.

[7] This competence is defined as "the use of language as an instrument of oral and written communication, of presentation, interpretation and comprehension of reality; to construct and communicate the knowledge, to organize and to auto-regulate thinking, emotions and behaviour" (AGAEVE, 2009, p. 7).

[8] This competence is defined as "the ability to use and relate numbers, their basic operations, symbols and expression forms and mathematic reasoning, to produce and interpret different types of information and to increase knowledge on quantitative and spatial aspects of reality and to solve problems related to daily life and to the labour world" (AGAEVE, 2009, p. 7).

[9] Data for these academic years have been used as 2012/13 is the latest available wave of this census data; therefore, our research study will focus on it and in the previous academic year.

From a total census of 78,706 8th grade students in 2011-12, we can identify 6,594 students who failed 8th grade in 2011-12 and repeated this grade in 2012-13. The main descriptive statistics and test of mean differences (in 2011-12)[10], comparing those students who passed 8th grade in 2011-12 with those who failed 8th grade in 2011-12 and repeated it in 2012-13, are presented in Table A2 (Appendix)[11]. As expected, we can see that students who repeated obtained lower scores in reading and mathematics in 2011-12. In addition, they also had lower socio-economic status[12] than those who did not repeat.

With the objective of showing the methodology employed by most of the research works on this issue, in Table A3 (Appendix) we have estimated the influence of grade retention on 8th grade students' academic competences in 2012-13 including both non-repeaters and repeaters, employing grade retention as an explanatory variable and using ordinary least squares (OLS) without student fixed-effects. We find that being a repeater in 8th grade seems to negatively influence students' academic competences. Nevertheless, as we previously highlighted, this is due to the endogenous characteristics of grade retention when used to explain students' competences. Therefore, in the following we propose the use of Ordinary Least Squares (OLS) estimation method with student fixed-effects within-students between-academic years (2011-12/2012-13) to overcome this issue and try to get as close as possible to the effect of grade retention on students' academic competences.

3. Methodology

The methodology employed in the present research is based on the use of Ordinary Least Squares (OLS) estimation method with student fixed-effects within-student between-academic years (2011-12/2012-13). Particularly, we focus on the sample of students who failed 8th grade in 2011-12 and repeated it in 2012-13. Hence, after controlling by a set of variables which may change (i.e. are not fixed) between 2011-12 and 2012-13, the rest may be kept the same, so students' difference in terms of competences may be driven by grade retention. Particularly, we estimate the following model:

$$Y_{ijct} = \alpha + \beta GR_{ijct} + \gamma X_{ijct} + \delta SCH_{jct} + \varphi CLSS_{ct} + \varepsilon_{ijct} \quad (1)$$

Where i represents the student, j the school, c the classroom and t the time (2011-12 or 2012-13). Y_{ijct} are students' standardised competence scores (alternatively) in reading or mathematics[13]; GR_{ijct} is grade retention (which takes the value "1" in $t = 2012 - 13$ and "0" in $t = 2011 - 12$); X_{ijct} are students' characteristics; SCH_{jct} are school characteristics; $CLSS_{ct}$ are classroom characteristics; α is a constant term and ε_{ijct} is the idiosyncratic error term.

[10] The test of mean differences has been performed using the t-student test (Gosset, 1908) for mean comparisons between populations, with unknown but equal variances. This test has been widely used over a century for this purpose and was designed by student (pseudonym of "Gosset").

[11] These statistics were calculated using the information of 8th grade students in 2011-12.

[12] This dataset contains a cultural and socio-economic index created by AGAEVE to have mean 0 and standard deviation 1. Some socio-economic variables from parents' questionnaire were used to create it: the highest level of education of the parents, the highest parental occupation, the number of books at home and the level of home resources.

[13] Students' scores in both reading and mathematics have been standardised to have mean 0 and standard deviation 1 (using the census mean and standard deviations for this purpose), with the objective of interpreting the coefficients of our estimations as effect sizes, i.e. the influence of the variables in the estimations is interpreted as standard-deviations variations on students' standardised scores. This facilitates the comparison of our results with other international studies.

Then, we take time differences within-student between-academic years, obtaining the following Ordinary Least Squares (OLS) estimation with student fixed-effects (herein the base model):

$$\Delta Y = \beta \Delta GR + \gamma \Delta X + \delta \Delta SCH + \varphi \Delta CLSS + \Delta \varepsilon \quad (2)$$

Where β is the influence of grade retention on students' competences in lower secondary education, which is our coefficient of interest. The use of student fixed-effects lets us control by every variable which is the same within-student between-academic years (so, for these variables, $\Delta X = 0$, $\Delta SCH = 0$ and $\Delta CLSS = 0$). The idea of using this approach is to analyse the influence of grade retention on students' competence scores reducing the potential confounders which may arise when many years pass between the academic year in which the student failed and that in which the student repeats, so two consecutive academic years have been employed for this purpose. Nevertheless, we add certain variables which may vary between-academic years to control by the potential bias on the β coefficient. These variables are, at student level: students' socio-economic index, frequency reading books per week, daily time on homework, daily time watching TV, daily time playing video games or computer games[14], relationship with classmates and relationship with the teachers; at school level we control by school change and the funding of the school of origin and destiny; at classroom level, we control by the average socio-economic status of the classroom, the proportion of students in the classroom with a very good/good relationship with classmates, the proportion of students in the classroom with a very good/good relationship with teachers and student age differences with the average age of students in the classroom[15,16]. The latter variable is of special relevance. Particularly, it takes from the β coefficient the influence that higher age (i.e., potential higher mental maturity due to age and experience), compared to peers, can have in the relationship of grade retention with students' academic competences. Thus, after all these controls, we can get as close as possible to the causal effect of grade retention in lower secondary education on students' competences. However, Ordinary Least Squares with student fixed-effects may not capture all the effects of potential unobservable variables, so we are cautious in the interpretation of our results. Robust estimations have been obtained using a first-order Taylor-series linearization method, which is also named as Huber/White/sandwich method.

We analysed the robustness of our results by dividing our estimations by whether students repeated previously to 2011-12 or not, due to potential differences in the

[14] The complementarity of the time use variables "frequency reading books per week", "daily time on homework", "daily time watching TV", "daily time playing video games or computer games" has been checked using them alternatively in the estimations and results do not change. These estimations will be provided upon request to the authors.

[15] This variable has been created by calculating the difference between each student's age and the average age of the students in the classroom in days, then converting it into weeks.

[16] Controlling for age is crucial in this research study. However, age cannot be included as regressor in the OLS estimations with student fixed-effects, to the extent that all students have aged the same between 2011-12 and 2012-13, so the age variation obtained in the estimations presents a null coefficient. However, this is not an issue, because fixed-coefficients, although not present in the estimation, are controlled by this methodology. Therefore, age differences with the average age of students in the classroom have been controlled for. In addition, to the extent that early and late school starts are not possible in Spain, as the Spanish education law indicates that students have to begin compulsory education the year they get 6 years old, without any possibility of advancing or delaying this start, age differences of more than one year would be only due to grade retention.

education production function by these students (as indicated by e.g. Carabaña, 2013; García-Pérez, Hidalgo-Hidalgo, & Robles-Zurita, 2014).

4. Results

In the following we present the main estimations of our base model of equation (2) in Table 1, adding the variables in a stepwise procedure. First, in specification I, the focus is only placed on students' grade retention; then, in specification II, we add students' age differences with the average age of students in the classroom and, in specification III, we add all the previously indicated covariates, reaching to our base model of equation (2)[17].

-Insert Table 1 here-

In the view of the results in specification I, it may seem that grade retention actually has a positive influence on students' reading and mathematics competences (of 0.144 and 0.109 standard deviation – SD –, respectively). However, in specification II, we control by differences with the average age of students in the classroom to account for the potential higher mental maturity of repeater students (due to higher age and experience) and we find that grade retention does not seem to positively influence students' reading competence, but still influences mathematics in 0.080 SD. In fact, a higher difference with the average age of students in the classroom is significant and presents a positive influence of 0.004 SD per week of difference for the reading competence, gathering the influence of grade retention on this competence. Finally, we intend to control by any other potential confounding variables which may bias this influence, so we include a set of variables which may change between 2011-12 and 2012-13 (those indicated in the Methodology section). Then, when adding these controls in specification III, we find that grade retention is still not significant for reading but, in the case of mathematics, the influence is increased to 0.145 SD. Thus, it seems that students improve their competences in mathematics when repeating.

As we previously indicated, students may present differences by previous grade retention before the academic year 2011-12 (or not) in this grade retention influence. In order to check this, we re-estimate our base model (the one in Table 1 specification III) differentiating by previous grade retention before the academic year 2011-12 (or not), and present the results in Table 2.

-Insert Table 2-

Regarding previous grade retention, as we can see, it seems that the positive influence of grade retention on students' mathematics competence is kept for both students who did not repeat before 2011-12 (0.144 SD) and those who repeated before 2011-12 (0.146 SD), so there are not any differences due to previous repetition in mathematics. Moreover, as shown in Table 1, we can see that the positive influence of grade retention is taken by age differences with the average age of the classroom in the case of reading.

5. Discussion and conclusions

In the present research we intend to add to the ongoing debate on the influence that grade retention may have on the development of students' academic competences in the short-term. To do this, we analyse lower secondary education students going further from correlation and getting as close as possible to a causal effect. This is achieved by

[17] A missing flag methodology to avoid losing observations has been used.

the use of census and longitudinal data for the Spanish region of Andalusia, together with Ordinary Least Squares (OLS) estimation method with student fixed-effects within-students between-academic years. Our results have shown that grade retention seems to have a null influence in reading and a positive influence in mathematics in the short-term. Particularly, we have found that these results do not differ conditioned on previous repetition.

There is also a relevant variable which has been included in our research and seems to influence the relationship between grade retention and students' competence: students' age differences with the average age of the classroom. In fact, our results seem to indicate that being older supposes a higher level of reading competence; this is logical to the extent that older students may be potentially more mentally mature due to age and experience (or maybe because they have been exposed to a higher amount of time reading texts, due to their age).

Nevertheless, in spite of the null influence that grade retention has on students' academic competences in reading, the students under analysis automatically passed to 9th grade in 2013-14 (they cannot repeat twice 8th grade due to the Spanish education legislation; BOE, 2006). Hence, these students may accumulate this gap on their competences compared to students of their age, which may be translated into higher competence deficiencies or even dropout (Spanish students can drop out at the age of 16 years old).

Additionally, these results are interesting to the extent that we analyse students' competences, i.e., students' skills, and not content-based knowledge. It seems that grade retention does not negatively influence these competences, a result which may seem counterintuitive. It could be argued that teachers may be inflating the test scores obtained by these students, maybe because students cannot repeat the same grade twice; however, two factors guarantee that this does not happen: (a) we are measuring students' competences (and not content-based knowledge) and (b) these tests are scored by external teachers (and not considered for the decision of passing the students to the next grade or not). In addition, the cumulative learning process in the learning of competences also supports that students may gain competences and do not lose them (Illeris, 2009).

Then, grade retention seems to help some students to compensate the potential lack of competences that they may present in mathematics. Nevertheless, as we did not find that grade retention is positive for reading, its implementation has to be made in a precise way. Maybe, this grade retention should not be limited to repeating the same grade again, but also including some support for repeater students in the form of, e.g., complementary lessons (Allen, Chen, Willson, & Hughes, 2009). Furthermore, preventing this grade retention can be another way of helping students to follow their academic track without losing one year, together with avoiding unnecessary public expense on these students. This has also been indicated by many authors, who agree that early identification of potential repeaters should be done in order to implement remedial measures to improve students' competences and prevent them from repeating (e.g., Jimerson, 2001; Jimerson & Kaufman, 2003; Jimerson, Pletcher, Graydon, Schnurr, Nickerson, & Kundert, 2005; Jimerson & Ferguson, 2007; Glick & Sahn, 2010; Choi, Gil, Mediavilla, & Valbuena, 2016; Agasisti & Cordero, 2017; among others).

This research work presents the limitation that, although the use of census longitudinal data and Ordinary Least Squares with student fixed-effects within-students between-academic years let us get very close to the causal effect of grade retention on lower secondary education students' academic competences, there may still be some

potential unobservable variables whose influence has not been captured by this methodology. In addition, we have found this relationship for Andalusian lower secondary education students in 8th grade; hence, although our results have high internal validity, this research has to be reproduced for other countries to check its external validity (Goos, Schreier, Knipprath, De Fraine, Van Damme, & Trautwein, 2013).

References

AGAEVE (2009). *Evaluación de Diagnóstico. Curso 2008-2009*. Andalucía: Consejería de Educación, Junta de Andalucía.

Agasisti, T., & Cordero, J. M. (2017). The determinants of repetition rates in Europe: Early skills or subsequent parents' help? *Journal of Policy Modeling, 39*(1), 129–146. doi: 10.1016/j.jpolmod.2016.07.002

Allen, C. S., Chen, Q., Willson, 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*(4), 480–499. doi: 10.3102/0162373709352239

Andrew, M. (2014). The Scarring Effects of Primary-Grade Retention? A Study of Cumulative Advantage in the Educational Career. *Social Forces, 93*(2), 653–685. doi: 10.1093/sf/sou074

Belot, M., & Vandenberghe, V. (2011). Evaluating the "threat" effects of grade repetition: exploiting the 2001 reform by the French-Speaking Community of Belgium. *Education Economics, 22*(1), 73–89. doi: 10.1080/09645292.2011.607266

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

Buckles, K. S., & Hungerman, D. M. (2013). Season of birth and later outcomes: Old questions, new answers. *The Review of Economics and Statistics, 95*(3), 711–724. doi: 10.1162/REST_a_00314

Calero, J., Choi, A., & Waisgrais, S. (2010). Determinants of the school failure risk in Spain: a multilevel logistic model approach to PISA-2006. *Revista de Educación, Extraordinary vol.*, 225–256.

Carabaña, J. (2013). Repetición de curso y puntuaciones pisa ¿cuál causa cuál? In INEE (Eds.), *PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Análisis secundario* (pp. 32–66). Madrid: Ministerio de Educación, Cultura y Deporte, Instituto Nacional de Evaluación Educativa.

Choi, A., Gil, M., Mediavilla, M., & Valbuena, J. (2016). *Double toil and trouble: grade retention and academic performance*. Document de treball de l'IEB 2016/17, 1–28.

Diaz, J., Grau, N., Reyes, T., & Rivera, J. (2021). *The Impact of Grade Retention on Juvenile Crime*. Working Papers wp513, University of Chile, Department of Economics, 1–58.

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. doi: 10.1016/s0272-7757(00)00041-8

Eurydice (2011). *Grade Retention during Compulsory Education in Europe: Regulations and Statistics*. Brussels: Education, Audiovisual and Culture Executive Agency. doi: 10.2797/50570

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. doi: 10.1080/00036846.2013.872761

Glick, P., & Sahn, D. E. (2010). Early Academic Performance, Grade Repetition, and School Attainment in Senegal: A Panel Data Analysis. *The World Bank Economic Review, 24*(1), 93-120. doi: 10.2307/40647422

González-Betancor, S. M., & López-Puig, A. J. (2016). Grade Retention in Primary Education Is Associated with Quarter of Birth and Socioeconomic Status. *PLOS ONE, 11*(11), e0166431. doi: 10.1371/journal.pone.0166431

Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: A systematic review and meta-analysis. *Educational Research Review, 34*, 100401. doi: 10.1016/j.edurev.2021.100401

Goos, M., Schreier, B. M., Knipprath, H. M. E., De Fraine, B., Van Damme, J., & Trautwein, U. (2013). 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. doi: 10.1086/667655

Gortazar, L. (2019). *¿Favorece el sistema educativo español la igualdad de oportunidades?* FEDEA, Estudios sobre la Economía Española no. 2019-17, Madrid.

Gosset [Student, pseud.], W. S. (1908). The probable error of a mean. Biometrika, 6, 1–25. doi: 10.2307/2331554

Hwang, S. H. J., & Cappella, E. (2018). Rethinking Early Elementary Grade Retention: Examining Long-Term Academic and Psychosocial Outcomes. *Journal of Research on Educational Effectiveness, 11*(4), 559–587. doi: 10.1080/19345747.2018.1496500

IECA (2019). Premature dropout rate by sex. Last time accessed July 2023 from http://www.juntadeandalucia.es/institutodeestadisticaycartografia/indsoc/indicadores/1038.htm

Illeris, K. (2009). Competence, learning and education: how can competences be learned, and how can they be developed in formal education? In K. Illeris (Eds.), *International Perspectives on Competence Development Developing Skills and Capabilities* (83–98). London: Routledge. doi: 10.4324/9780203523032

INE (2019). Annual household net income per capita, by Autonomous Community. Retrieved from http://www.ine.es/jaxiT3/Tabla.htm?t=9949 (Last time accessed July 2023).

Jacob, B. A., & Lefgren, L. (2009). The Effect of Grade Retention on High School Completion. *American Economic Journal: Applied Economics, 1*(3), 33–58. doi: 10.1257/app.1.3.33

Jerrim, J., Lopez-Agudo, L. A., Marcenaro-Gutierrez, O. D., & Shure, N. (2017). What happens when econometrics and psychometrics collide? An example using the PISA data. *Economics of Education Review, 61*, 51–58. doi: 10.1016/j.econedurev.2017.09.007

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., & Kaufman, A. M. (2003). Reading, Writing, and Retention: A Primer on Grade Retention Research. *The Reading Teacher, 56*(7), 622–635.

Jimerson, S. R., Pletcher, S. M. W., Graydon, K., Schnurr, B. L., Nickerson, A. B., & Kundert, D. K. (2005). Beyond grade retention and social promotion: Promoting the social and academic competence of students. *Psychology in the Schools, 43*(1), 85–97. doi: 10.1002/pits.20132

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. doi: 10.1037/1045-3830.22.3.314

Manacorda, M. (2012). The Cost of Grade Retention. *Review of Economics and Statistics, 94*(2), 596–606. doi: 10.1162/rest_a_00165

Mariano, L. T., Martorell, P., & Tsai, T. (2018). *The Effects of Grade Retention on High School Outcomes: Evidence from New York City Schools*. RAND Education, WR-1259-DEIES, 1–48.

Martorell, P., & Mariano, L. T. (2018). The Causal Effects of Grade Retention on Behavioral Outcomes. *Journal of Research on Educational Effectiveness, 11*(2), 192–216. doi: 10.1080/19345747.2017.1390024

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

Ou, S.-R., & Reynolds, A. J. (2010). Grade Retention, Postsecondary Education, and Public Aid Receipt. *Educational Evaluation and Policy Analysis, 32*(1), 118–139. doi: 10.3102/0162373709354334

Özek, U. (2015). Hold Back To Move Forward? Early Grade Retention And Student Misbehavior. *Education Finance and Policy, 10*(3), 350–377. doi: 10.1162/edfp_a_00166

Pagani, L., Tremblay, R. E., Vitaro, F., Boulerica, B., & McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology, 13*(2), 297–315. doi: 10.1017/s0954579401002061

Pedraja-Chaparro, F., Santín, D., & Simancas, R. (2015). Determinants of grade retention in France and Spain: Does birth month matter? *Journal of Policy Modeling, 37*(5), 820–834. doi: 10.1016/j.jpolmod.2015.04.004

Peraita, C., & Pastor, M. (2000). The Primary School Dropout in Spain: The Influence of Family Background and Labor Market Conditions. *Education Economics, 8*(2), 157–168. doi: 10.1080/096452900410721

Schwerdt, G., West, M. R., & Winters, M. A. (2017). The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. *Journal of Public Economics, 152*, 154–169. doi: 10.1016/j.jpubeco.2017.06.004

Tafreschi, D., & Thiemann, P. (2016). Doing it twice, getting it right? The effects of grade retention and course repetition in higher education. *Economics of Education Review, 55*, 198–219. doi: 10.1016/j.econedurev.2016.10.003

Appendix

-Insert Table A1-

-Insert Table A2-

-Insert Table A3-

Table 1. The influence of grade retention on students' competences in 8th grade

	Specification I		Specification II		Specification III	
Variables	Reading	Mathematics	Reading	Mathematics	Reading	Mathematics
Failed 8th grade in 2011-12 and repeating in 2012-13 (Ref.: Non-repeater)	0.144***	0.109***	-0.035	0.080**	0.028	0.145***
	(0.012)	(0.010)	(0.038)	(0.033)	(0.043)	(0.037)
Standardised socio-economic index	-	-	-	-	0.025	0.006
					(0.017)	(0.015)
Standardised socio-economic index Missing flag	-	-	-	-	-0.164***	-0.092***
					(0.037)	(0.028)
Frequency reading books per week (Ref.: Never or almost never)						
Every or almost every day	-	-	-	-	0.072**	0.039
					(0.035)	(0.030)
Once or twice a week	-	-	-	-	0.074***	0.025
					(0.026)	(0.022)
Once or twice a month	-	-	-	-	0.061***	0.001
					(0.023)	(0.020)
Frequency reading books. Missing flag	-	-	-	-	-0.000	-0.023
					(0.051)	(0.042)
Daily time on homework (Ref.: 0 hours)						
More than 3 hours	-	-	-	-	0.233***	0.207***
					(0.059)	(0.052)
3 hours	-	-	-	-	0.245***	0.219***
					(0.049)	(0.043)
2 hours	-	-	-	-	0.238***	0.194***
					(0.041)	(0.037)
1 hour	-	-	-	-	0.167***	0.103***
					(0.038)	(0.034)
Daily time on homework. Missing flag	-	-	-	-	0.243***	0.132*
					(0.080)	(0.070)
Daily time watching TV (videos, DVD) (Ref.: 0 hours)						
5 hours or more	-	-	-	-	-0.116*	-0.018
					(0.060)	(0.051)
Between 3 and 5 hours	-	-	-	-	-0.074	0.004
					(0.052)	(0.044)
Between 1 and 3 hours	-	-	-	-	-0.121**	-0.004
					(0.048)	(0.041)
Until 1 hour	-	-	-	-	-0.095**	0.014
					(0.047)	(0.041)
Time watching TV. Missing flag	-	-	-	-	-0.241***	-0.049
					(0.089)	(0.074)

Daily time playing video games or computer games (Ref.: 0 hours)						
5 hours or more	-	-	-	-	-0.035	-0.033
					(0.036)	(0.032)
Between 3 and 5 hours	-	-	-	-	0.016	-0.003
					(0.035)	(0.030)
Between 1 and 3 hours	-	-	-	-	0.003	0.006
					(0.031)	(0.027)
Until 1 hour	-	-	-	-	-0.022	0.025
					(0.031)	(0.026)
Time playing video games or computer games. Missing flag	-	-	-	-	-0.013	0.027
					(0.078)	(0.060)
Relationship with classmates (Ref.: Bad)						
Very good	-	-	-	-	0.033	0.104*
					(0.068)	(0.055)
Good	-	-	-	-	0.055	0.101*
					(0.067)	(0.054)
Not good	-	-	-	-	0.045	0.085
					(0.068)	(0.055)
Relationship with classmates. Missing flag	-	-	-	-	-0.095	0.000
					(0.103)	(0.083)
Relationship with teachers (Ref.: Bad)						
Very good	-	-	-	-	0.209***	0.134***
					(0.055)	(0.051)
Good	-	-	-	-	0.163***	0.088*
					(0.050)	(0.046)
Not good	-	-	-	-	0.072	0.041
					(0.049)	(0.044)
Relationship with teachers. Missing flag	-	-	-	-	0.300***	0.151*
					(0.091)	(0.078)
School change (Ref.: Did not change school)						
Public to semi-private/private school	-	-	-	-	0.199*	0.543***
					(0.109)	(0.117)
Semi-private/private to public school	-	-	-	-	-0.131	-0.058
					(0.097)	(0.109)
Public to public school	-	-	-	-	0.020	-0.071
					(0.075)	(0.061)
Semi-private/private to semi-private/private school	-	-	-	-	0.062	-0.012
					(0.148)	(0.121)
Average standardised socio-economic index of the classroom	-	-	-	-	0.023	0.016
					(0.037)	(0.033)
Average standardised socio-economic index of the classroom. Missing flag	-	-	-	-	0.184*	0.062

					(0.110)	(0.085)
Age difference with the average age of students in the classroom (in weeks)	-	-	0.004***	0.001	0.003***	-0.000
			(0.001)	(0.001)	(0.001)	(0.001)
Proportion of students in the classroom with a very good/good relationship with classmates	-	-	-	-	0.142	0.183**
					(0.103)	(0.089)
Proportion of students in the classroom with a very good/good relationship with classmates. Missing flag	-	-	-	-	0.427*	0.330*
					(0.246)	(0.192)
Proportion of students in the classroom with a very good/good relationship with teachers	-	-	-	-	0.130	0.126*
					(0.084)	(0.070)
Constant	-0.663***	-0.661***	-0.658***	-0.660***	-1.187***	-1.261***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.136)	(0.108)
Observations	11,086	11,154	11,086	11,154	11,086	11,154
R-squared	0.746	0.756	0.747	0.756	0.755	0.763
Number of students	5,543	5,577	5,543	5,577	5,543	5,577

Notes: Standard errors in parentheses and robust (using a first-order Taylor-series linearization method, which is also named as Huber/White/sandwich method). Sample of students who failed 8th grade in 2011-12 and repeated in 2012-13.
Estimation method: Ordinary Least Squares (OLS) estimation method with student fixed-effects.
Dependent variable: Students' standardised scores using population's mean and standard deviation.
Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Table 2. The influence of grade retention on students' competences in 8th grade by previous grade retention

Variables	Students who did not repeat before 2011-12		Students who repeated before 2011-12	
	Reading	Mathematics	Reading	Mathematics
Failed 8th grade in 2011-12 and repeating in 2012-13 (Ref.: Non-repeater)	0.068	0.144***	0.002	0.146***
	(0.064)	(0.053)	(0.059)	(0.051)
Age difference with the average age of students in the classroom (in weeks)	0.002	0.000	0.004***	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Additional controls	✓	✓	✓	✓
Observations	6,376	6,412	4,710	4,742
R-squared	0.734	0.751	0.768	0.770
Number of students	3,188	3,206	2,355	2,371

Notes: Standard errors in parentheses and robust (using a first-order Taylor-series linearization method, which is also named as Huber/White/sandwich method). Sample of students who failed 8th grade in 2011-12 and repeated in 2012-13. The think " ✓ " indicates that it has been also controlled by students' socio-economic index, frequency reading books per week, daily time on homework, daily time watching TV, daily time playing video games or computer games, relationship with classmates and relationship with the teachers, school change and the funding of the school from origin and destiny, the socio-economic status of the classroom, the proportion of students in the classroom with a very good/good relationship with classmates and the proportion of students in the classroom with a very good/good relationship with teachers. These estimations are available upon request to the authors.
Estimation method: Ordinary Least Squares (OLS) estimation method with student fixed-effects.
Dependent variable: Students' standardised scores using population's mean and standard deviation.
Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Table A1. Comparison of Andalusian, Spanish and OECD education figures in 2015

		Andalusia	**Spain**	**OECD**
PISA 2015 Mean Scores	Reading	479	496	487
	Mathematics	466	486	478
	Sciences	473	493	488
Percentage of students repeating at least one grade by 12th grade		38%	31%	13%
Father's education	University studies	32%	43%	42%
	High school studies	15%	18%	33%
	Secondary education studies	29%	23%	15%
	Primary education studies	18%	12%	7%
	Less than primary education studies	7%	4%	3%
Mother's education	University studies	30%	43%	43%
	High school studies	17%	21%	33%
	Secondary education studies	30%	23%	14%
	Primary education studies	18%	11%	7%
	Less than primary education studies	5%	3%	3%
Annual household net income per capita, in PPPs		16,276$	20,367$	28,443$

Source: Authors' own calculations from PISA 2015, OECD (2016) and INE (2019).

Table A2. Descriptive statistics and test of mean differences by grade retention when finishing 8th grade in 2011-12

		Students who passed 8th grade in 2011-12			Students who did not pass 8th grade in 2011-12 and repeated it in 2012-13		
Variables		**Obs.**	**Mean**	**S.d.**	**Obs.**	**Mean**	**S.d.**
Standardised scores in reading		68,407	0.11^{D}	0.97	6,161	-0.68^{D}	0.86
Standardised scores in mathematics		68,758	0.10^{D}	0.99	6,163	-0.67^{D}	0.71
Standardised socio-economic status index		69,876	0.05^{D}	1.00	6,374	-0.46^{D}	0.85
Frequency reading books	**Every or almost every day**	70,525	0.14^{D}	0.35	6,365	0.09^{D}	0.29
	Once or twice a week	70,525	0.28^{D}	0.45	6,365	0.23^{D}	0.42
	Once or twice a month	70,525	0.31^{D}	0.46	6,365	0.30^{D}	0.46
	Never or almost never	70,525	0.27^{D}	0.44	6,365	0.38^{D}	0.49
Daily time on homework	**More than 3 hours**	70,984	0.09^{D}	0.29	6,471	0.05^{D}	0.21
	3 hours	70,984	0.17^{D}	0.38	6,471	0.11^{D}	0.31
	2 hours	70,984	0.42^{D}	0.49	6,471	0.37^{D}	0.49
	1 hour	70,984	0.28^{D}	0.45	6,471	0.40^{D}	0.49
	0 hours	70,984	0.04^{D}	0.20	6,471	0.07^{D}	0.25
Time watching TV	**5 hours or more**	71,345	0.05^{D}	0.22	6,495	0.08^{D}	0.27
	Between 3 and 5 hours	71,345	0.15^{D}	0.35	6,495	0.17^{D}	0.38
	Between 1 and 3 hours	71,345	0.46^{D}	0.50	6,495	0.43^{D}	0.50
	Until 1 hour	71,345	0.30^{D}	0.46	6,495	0.28^{D}	0.45
	0 hours	71,345	0.04	0.20	6,495	0.04	0.20
Time playing video games or computer games	**5 hours or more**	70,847	0.06^{D}	0.24	6,465	0.10^{D}	0.29
	Between 3 and 5 hours	70,847	0.12^{D}	0.32	6,465	0.14^{D}	0.35
	Between 1 and 3 hours	70,847	0.29^{D}	0.45	6,465	0.27^{D}	0.45
	Until 1 hour	70,847	0.32^{D}	0.47	6,465	0.30^{D}	0.46
	0 hours	70,847	0.21^{D}	0.41	6,465	0.19^{D}	0.39
Relationship with classmates	**Very good**	71,534	0.49^{D}	0.50	6,518	0.45^{D}	0.50
	Good	71,534	0.42	0.49	6,518	0.42	0.49
	Not good	71,534	0.08^{D}	0.28	6,518	0.11^{D}	0.32
	Bad	71,534	0.01^{D}	0.11	6,518	0.02^{D}	0.12
Relationship with teachers	**Very good**	71,511	0.27^{D}	0.44	6,508	0.15^{D}	0.36
	Good	71,511	0.57^{D}	0.50	6,508	0.54^{D}	0.50
	Not good	71,511	0.14^{D}	0.35	6,508	0.27^{D}	0.44
	Bad	71,511	0.02^{D}	0.14	6,508	0.04^{D}	0.20
Average standardised socio-economic index of the classroom		70,977	0.01^{D}	0.61	6,534	-0.19^{D}	0.51
Proportion of students in the classroom with a very good/good relationship with classmates		72,100	0.90^{D}	0.08	6,594	0.89^{D}	0.09
Proportion of students in the classroom with a very good/good relationship with teachers		72,112	0.83^{D}	0.11	6,594	0.80^{D}	0.12

Notes: "Obs." stands for "Observations" or "S.d." stands for "standard deviations". The superscript "D" indicates that there are significant mean differences between students who passed 8th grade in 2011-12 and students who did not pass 8th grade in 2011-12 and repeated it in 2012-13 (t-student test has been employed, Gosset, 1908). The data is that of 8th grade in 2011-12.
Source: Authors' own calculations.

Table A3. The influence of grade retention on 8th grade students' competences in 2012-13

Variables	Reading	Mathematics
Failed 8th grade in 2011-12 and repeating in 2012-13 (Ref.: Non-repeater)	-0.585***	-0.621***
	(0.012)	(0.011)
Constant	0.059***	0.063***
	(0.004)	(0.004)
Observations	78,318	78,789
R-squared	0.024	0.027

Notes: Standard errors in parentheses and robust (using a first-order Taylor-series linearization method, which is also named as Huber/White/sandwich method). Complete population of students in 8th grade in 2012-13.
Estimation method: Ordinary Least Squares (OLS).
Dependent variable: Students' standardised scores using population's mean and standard deviation in 2012-13.
Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Online supplemental material

Table S1. The influence of grade retention on students' competences in 8th grade by previous grade retention (Table 2 in main text)

Variables	Students who did not repeat before 2011-12		Students who repeated before 2011-12	
	Reading	Mathematics	Reading	Mathematics
Failed 8th grade in 2011-12 and repeating in 2012-13 (Ref.: Non-repeater)	0.068	0.144***	0.002	0.146***
	(0.064)	(0.053)	(0.059)	(0.051)
Standardised socio-economic index	0.036	-0.008	0.015	0.024
	(0.024)	(0.020)	(0.025)	(0.022)
Standardised socio-economic index Missing flag	-0.168***	-0.090**	-0.153***	-0.089**
	(0.053)	(0.040)	(0.051)	(0.040)
Frequency reading books per week (Ref.: Never or almost never)				
Every or almost every day	0.108**	0.054	0.011	0.013
	(0.046)	(0.041)	(0.054)	(0.044)
Once or twice a week	0.086**	0.025	0.059	0.027
	(0.034)	(0.029)	(0.040)	(0.034)
Once or twice a month	0.068**	-0.002	0.052	0.001
	(0.031)	(0.026)	(0.035)	(0.030)
Frequency reading books. Missing flag	0.077	-0.042	-0.081	-0.007
	(0.067)	(0.061)	(0.078)	(0.058)
Daily time on homework (Ref.: 0 hours)				
More than 3 hours	0.207**	0.181**	0.256***	0.217***
	(0.083)	(0.080)	(0.088)	(0.070)
3 hours	0.184***	0.249***	0.330***	0.157**
	(0.069)	(0.066)	(0.073)	(0.061)
2 hours	0.211***	0.182***	0.264***	0.210***
	(0.061)	(0.061)	(0.057)	(0.048)
1 hour	0.155***	0.109*	0.174***	0.092**
	(0.058)	(0.059)	(0.050)	(0.041)
Daily time on homework. Missing flag	0.263**	0.145	0.199	0.122
	(0.106)	(0.094)	(0.125)	(0.113)
Daily time watching TV (videos, DVD) (Ref.: 0 hours)				
5 hours or more	-0.074	-0.071	-0.137	0.030
	(0.084)	(0.076)	(0.084)	(0.066)
Between 3 and 5 hours	-0.065	-0.033	-0.081	0.040
	(0.071)	(0.063)	(0.075)	(0.059)
Between 1 and 3 hours	-0.122*	-0.019	-0.104	0.011
	(0.066)	(0.059)	(0.070)	(0.055)
Until 1 hour	-0.055	-0.008	-0.132*	0.036
	(0.064)	(0.059)	(0.069)	(0.053)
Time watching TV. Missing flag	-0.208*	0.044	-0.269**	-0.175
	(0.125)	(0.100)	(0.129)	(0.112)
Daily time playing video games or computer games (Ref.: 0 hours)				
5 hours or more	-0.015	0.026	-0.063	-0.093**
	(0.049)	(0.045)	(0.053)	(0.044)
Between 3 and 5 hours	0.046	0.014	-0.025	-0.013
	(0.047)	(0.041)	(0.053)	(0.043)
Between 1 and 3 hours	0.015	0.038	-0.009	-0.030
	(0.042)	(0.036)	(0.047)	(0.041)
Until 1 hour	-0.010	0.051	-0.041	0.002

	(0.041)	(0.036)	(0.046)	(0.038)
Time playing video games or computer games. Missing flag	-0.060	-0.030	0.056	0.065
	(0.111)	(0.088)	(0.108)	(0.077)
Relationship with classmates (Ref.: Bad)				
Very good	-0.057	0.163**	0.111	0.024
	(0.097)	(0.074)	(0.095)	(0.080)
Good	-0.011	0.178**	0.104	-0.009
	(0.095)	(0.071)	(0.094)	(0.079)
Not good	0.003	0.163**	0.058	-0.026
	(0.098)	(0.075)	(0.094)	(0.079)
Relationship with classmates. Missing flag	-0.203	0.068	0.004	-0.090
	(0.145)	(0.116)	(0.145)	(0.112)
Relationship with teachers (Ref.: Bad)				
Very good	0.175**	0.184***	0.239***	0.061
	(0.071)	(0.068)	(0.087)	(0.075)
Good	0.157**	0.134**	0.163**	0.014
	(0.065)	(0.063)	(0.078)	(0.068)
Not good	0.071	0.073	0.065	-0.018
	(0.064)	(0.060)	(0.078)	(0.063)
Relationship with teachers. Missing flag	0.463***	0.229**	0.096	0.020
	(0.123)	(0.109)	(0.136)	(0.107)
School change (Ref.: Did not change school)				
Public to semi-private/private school	0.172	0.378***	0.238	0.811***
	(0.133)	(0.144)	(0.192)	(0.191)
Semi-private/private to public school	-0.140	-0.020	-0.173	-0.076
	(0.147)	(0.162)	(0.132)	(0.146)
Public to public school	0.065	-0.079	-0.022	-0.066
	(0.100)	(0.088)	(0.112)	(0.083)
Semi-private/private to semi-private/private school	0.115	-0.193	-0.011	0.193
	(0.213)	(0.161)	(0.188)	(0.176)
Average standardised socio-economic index of the classroom	0.096*	0.027	-0.065	0.011
	(0.049)	(0.046)	(0.056)	(0.046)
Average standardised socio-economic index of the classroom. Missing flag	0.001	0.047	0.415**	0.097
	(0.143)	(0.119)	(0.171)	(0.113)
Age difference with the average age of students in the classroom (in weeks)	0.002	0.000	0.004***	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Proportion of students in the classroom with a very good/good relationship with classmates	0.080	0.160	0.210	0.204
	(0.142)	(0.126)	(0.150)	(0.126)
Proportion of students in the classroom with a very good/good relationship with classmates. Missing flag	-0.127	0.182	0.886***	0.514*
	(0.431)	(0.250)	(0.279)	(0.285)
Proportion of students in the classroom with a very good/good relationship with teachers	0.128	0.085	0.159	0.182*
	(0.115)	(0.093)	(0.124)	(0.106)
Constant	-0.894***	-1.242***	-1.611***	-1.241***
	(0.190)	(0.155)	(0.200)	(0.155)
Observations	6,376	6,412	4,710	4,742
R-squared	0.734	0.751	0.768	0.770
Number of students	3,188	3,206	2,355	2,371

Notes: Standard errors in parentheses and robust (using a first-order Taylor-series linearization method, which is also named as Huber/White/sandwich method). Sample of students who failed 8th grade in 2011-12 and repeated in 2012-13.
Estimation method: Ordinary Least Squares (OLS) estimation method with student fixed-effects.

Dependent variable: Students' standardised scores using population's mean and standard deviation.
Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.