Is retention beneficial to low-achieving students? Evidence from Portugal[1]

Luis Catela Nunes*,[a]

Ana Balcão Reis*,[b]

Carmo Seabra*,[c]

*Nova School of Business and Economics, Universidade Nova de Lisboa, Lisboa, Portugal

February 2018

Accepted for publication in Applied Economics

Abstract

The role of retention as an educational tool to overcome under-achievement is a hotly debated issue, especially given that the results in the literature are not consensual. The Portuguese case is particularly well suited to study this issue: all students must take standardized national exams at specific grades. Moreover, the available dataset tracks the performance of students over time. Therefore, we are able to measure the impact of students' retention on their subsequent academic performance since we can control for each student's initial level of ability at the moment of retention. We use a propensity score matching approach, in which retained and promoted 4th grade students are matched according to their socioeconomic characteristics and the scores obtained in national exams. To address potentially remaining endogeneity biases we also use the culture of retention at school level as an instrumental variable. The results suggest that in some situations retentions may have on average a positive impact on future achievement. However, in the cases where statistically significant impacts are found, the estimated magnitudes are relatively small. Our results are relevant for countries with high retention rates that are considering alternative educational policies to promote students' achievement.

Keywords: Grade retention; Low-achieving students; Students' performance; Standardized national exams

JEL codes: I21, I28, H42

---

[1] We acknowledge financial support from Fundação Francisco Manuel dos Santos and from National Funds through Fundação para a Ciência e Tecnologia (UID/ECO/00124/2013, PTDC/IIM-ECO/6813/2014), POR Lisboa (LISBOA-01-0145-FEDER007722 and Social Sciences Data Lab project 22209), and POR Norte (Social Sciences Data Lab project 22209). We thank DGEEC-Ministry of Education for providing the data. We are indebted to two anonymous reviewers who helped to improve the paper.

[a] Corresponding author: lcnunes@novasbe.pt; [b] abr@novasbe.pt; [c] carmo.seabra@novasbe.pt; phone: +351213801600 Postal Address: Nova SBE, Universidade Nova de Lisboa, Campus de Campolide, 1099-032 Lisboa, Portugal

1. Introduction

According to OECD (2014), 12% of the 15-year-old students across OECD countries reported having repeated a grade at least once during compulsory schooling, with 7% of the students having repeated a grade, at least once, in primary school. The incidence of retention is quite heterogeneous across countries, ranging from no repetitions at all in some countries, like Japan and Norway, to repetition rates above 25%, at least once before the age of 15, in many countries, including the Netherlands, France, Portugal, and Brazil (see OECD, 2013). With regard to primary education, we also observe a strong heterogeneity across countries, in which Portugal stands out with a repetition rate above 25% at the primary education level (see Eurydice, 2011).[2]

These figures highlight the controversy around the impact of retention/promotion decisions on the academic, professional, and social path of low-achieving students. Proponents of policies that encourage the retention of low-achieving students believe that retention gives students an opportunity to mature and master materials and contents – which were not duly learned during the failed year – before having to take up more challenging materials. Moreover, they argue that retention can improve the balance between low-achieving students' abilities and the abilities of their peers, thereby preventing them from feeling frustrated on a daily basis. This increased homogeneity in class composition can also allow teachers to better target the specific characteristics of their students. Critics of retention, or social promotion advocates, however, fear that retained students might suffer from stigmatization, reduced expectations toward their academic performance, by both teachers and parents, and perceiving themselves

[2] ISCED 1 corresponds to 1st and 2nd cycles in Portugal (from the 1st to the 6th grade; from 6 up to 12 years old, in normal circumstances).

as being less competent and having less potential, besides having to deal with the challenges of adjusting to a new peer group. They believe that together these factors could suppress any benefits arising from grade repetition and increase students´ anxiety, misbehaviour, and disengagement from school.[3]

The real impact of retention/promotion decisions on students' subsequent achievement has been the subject of many empirical studies. The majority of earlier research found that retention has a negative impact on students' subsequent academic achievement. Jimerson (2001) provides a summary of 22 empirical results published until that date, including the results of a meta-analysis of 44 empirical studies carried out by Holmes and Matthews (1984). While the results are far from unanimous, most of these early studies conclude that retention does not benefit students.

These conclusions remain controversial as these studies typically fail to take into account the endogeneity present in the decision to retain a student, making it difficult to isolate the causal effect of retention on students' academic performance at later stages. The challenge with the empirical analysis of the consequences of retention/promotion decisions resides in the fact that these decisions are usually related to unobservable characteristics of the students that, nonetheless, affect their future achievement, such as their ability, motivation, maturity, parental involvement, and aspirations. Many empirical studies ignore this unobserved heterogeneity, so that their results are likely to be influenced by a severe selection bias, the statistical estimates of grade retention benefits being biased downward.

[3] The financial costs of keeping a student at school for one more year (to students, their families, and society in general) should also be taken into consideration when comparing the possible benefits of retention with the costs of such decision.

More recent research takes advantage of the availability of more detailed longitudinal datasets containing students' microdata and is able to apply different identification strategies that increase the comparability between retained and promoted students, allowing researchers to establish the causal effect of retention. Allen et al. (2009) carried out a meta-analysis of post-1990 studies and found that the subset of empirical studies with more rigorous econometric designs conclude that retention has fewer negative effects on students' subsequent academic achievement.

One of the econometric identification strategies used to address the endogeneity issues is the regression discontinuity analysis. This method can be used whenever there is a test score cut-off that defines the conditions for promotion. The existence of that discontinuity allows the estimation of the impact of retention for students near the cut-off. Roderick and Nagaoka (2005), and Jacob and Lefgren (2004) used this approach for 3rd and 6th grades in Chicago Public Schools. They found a positive effect of retention on the first year that vanished in the second year for 3rd graders and no effect for 6th graders. Jacob and Lefgren (2009) extended this research to study the impact of retention on the likelihood of high school completion. They concluded that retention had no impact among 6th grade students, and that retaining 8th grade students decreased their probability of completing high school. Greene and Winters (2007) used a similar design to study the impact of retention of 3rd grade students in Florida and determined that retained students slightly outperformed socially promoted students regarding their reading skills in the year following retention, and that the gains increased substantially in the second year. Schwerdt et al. (2015), also using a regression discontinuity analysis and data for Florida, found that 3rd grade retention substantially improved students' achievement in the short run and reduced the probability of retention in subsequent grades. In

line with Jacob and Lefgren's (2009) results for Chicago, they concluded that 3rd grade retentions had no impact on high school completion. Overall, these results suggest that, for students with scores near the cut-off that determines retention, the causal impact of early grade retention (3rd grade) may be positive or, in some cases, null. The causal impact of retention at later stages seems to be more negative. Furthermore, even positive short-run impacts seem to fade out after a few years.

Whenever this type of discontinuity does not hold, as is the case with our data set, dealing with endogeneity requires other techniques. Pereira and Reis (2014), and Garcia-Pérez et al. (2014) study the impact of retention in Portugal and Spain, respectively, using the PISA dataset. Both papers use the date of birth as an instrumental variable in the context of regression models that account for the endogeneity bias. These methods allow for the estimation of the average impact of retention on retained and promoted students. In both cases, the authors found that grade retention has a negative impact on educational outcomes, and that it is stronger for retentions occurring during primary education. The advantage of using the PISA dataset is the richness of the available controls on students' characteristics at the age of 15, when the test is applied, as well as on schools' characteristics. However, a major drawback of the PISA dataset is the lack of information about the students' level of educational attainment at the moment of their initial retention. In fact, as mentioned by Allen et al. (2009), the studies that compare promoted and retained students matched according to previous academic achievement identify more positive effects due to the retention decision than the studies that compare promoted and retained students matched according to non-academic variables.

Other researchers have resorted to different matching strategies, comparing the population of retained students with a subset of promoted students that have similar baseline scores. Greene and Winters (2009) consider a set of 3rd grade students with test scores below the threshold set by the retention policy of Florida. They take advantage of Florida's exemptions system to compare retained and promoted students, matched according to their test scores. They conclude that retention has a positive impact on the reading skills of low-achieving students.

In our study we focus on the Portuguese case to analyse the impact of retention on the 4th grade. The main strategy to isolate the causal effect of retention is to compare promoted and retained students matched according to their baseline scores. Moreover, we make use of several methodological approaches to address potentially remaining endogeneity issues: multiple regression with a large set of statistical controls, propensity score matching, and instrumental variable estimation.

The remainder of this paper is organized as follows: in Section 2 we describe our methodological approach and the data, in Section 3 we present the results regarding the determinants of retentions and the impact of retentions, and in Section 4 we discuss the results and present our conclusions.

2. Methodology and Data

Our objective is to measure the impact of retention decisions on students' future academic performance. We wish to answer the following question: "How does the observable future performance of retained students compare with the performance they could obtain if they had not been retained?" A naïve approach to answering this question would merely compare the

performance of retained and promoted students. However, besides their retained or promoted status, there are other characteristics that differentiate these two groups of students that have an impact on the future performance of each group, leading to an endogeneity problem. Therefore, any differences in performance between the two groups cannot be attributed solely to retention decisions. We deal with this problem by using a matched sample of low-achieving students.

The Portuguese case is particularly well suited to study this issue. There is an administrative dataset following all the student population throughout their schooling trajectory, and during the period under analysis, all students had to take standardized national exams at the end of the 4th and 6th grades, making it possible to analyse the impact of retention at the 4th grade on the scores obtained on the 6th-grade exams, controlling for the level of ability at the moment of retention.[4] Moreover, although the 4th and 6th grade compulsory national exams are scored on a scale of 1 to 5, and a score of 1 or 2 is considered a negative evaluation, these scores do not determine whether a student repeats or is promoted to the next grade. That decision falls on teachers and on their evaluation of each student. [5]

We focus our analysis on the 4th grade low-achieving students, the ones who obtained negative scores on both Portuguese and Mathematics 4th grade national exams, some of whom were retained while others were promoted. These two groups of students – retained and promoted – are much more alike in our sample than in the whole population, so that the endogeneity problem is attenuated. We check for this homogeneity by looking at several descriptive statistics of the whole population and our sample.

[4] Other relevant consequences studied in the literature, such as the probability of drop-out or the participation in the labour market, are not considered in our study.
[5] These decisions may be subject to parents' and school pedagogical council agreement.

Our dataset comes from an administrative database managed by the Portuguese Ministry of Education containing information about students in public schools. For each student, there is information on the grade, track of studies, gender, age, nationality, social support, and availability of a computer and internet at home, as well as parents' nationality and education. We also have information on the scores obtained by the students on the national exams and whether they were promoted or retained at the end of an academic year. In our study we consider 4th grade students enrolled in the academic year of 2006/2007 and we follow their progress until 2009/2010.[6]

In Table 1 we present some descriptive statistics of the sample of students who obtained negative scores on both Portuguese and Mathematics 4th grade national exams and compare them with the total population of 4th grade students in the same year (2006/2007). As expected, the retention percentage in our sample (29%) is much higher than the retention percentage in the whole population (6%). The percentage of boys in our sample (62%) is also higher than in the population (52%). Many students in our sample (65%) are older than they should be. Some may have entered school later, but most of them are likely to have been retained before. The equivalent percentage for the population is only 25%. The percentage of students coming from other Portuguese-speaking countries is much higher in the sample. These results were expected, since the population of immigrants generally comes from low-income countries. A large percentage of the students in our sample (56%) have a low educational background, only 2% have parents with higher education.[7] Finally, students that

[6] The databases used are MISI and *Júri Nacional de Exames* from the Portuguese Ministry of Education.

[7] We do not present the information concerning fathers as the values are very similar to the data for mothers but with many more values missing.

perform poorly on the national exams are less likely to own a computer or have access to the internet at home. These results are in line with the literature on the determinants of success.

Table 1: Descriptive statistics for the population and sample of 4th grade students in 2006/07 (in %)

		Population	Sample (Negative Scores in Exams)
No. of students		106,469	6,039
Retained in Total		6	29
Males		52	62
Year of Birth	up to 1995 (aged 12)	8	29
	1996 (aged 11)	17	36
	1997 on (aged 10)	74	35
Student's Nationality	Other Portuguese-speaking Countries	3	5
Mother's Nationality	Other Portuguese-speaking Countries	4	9
Mother's Academic Background	Primary ($\leq$ ISCED 1)	44	56
	Secondary	46	42
	Higher Education	10	2
Student's Social Support		13	24
Computer at home		49	32
Internet at home		30	16

Table 2 includes data for the whole population of retained and promoted 4th grade students. In the population, 6% of the students were retained in the 4th grade. According to their age, only 54% of the retained students have never been retained before. The percentage of students from other Portuguese-speaking countries is higher among retained students. The same pattern is true for their parents' nationality. The academic background of the mother exhibits the anticipated pattern, i.e., a greater percentage of children of less educated mothers is

observed in the group of retained students. The percentage of students benefiting from social support (low-income families) is much higher in the retained group than in the promoted one. Finally, retained students do not have as many computers and access to the internet as their promoted colleagues. Overall, we conclude that there are significant differences between promoted and retained 4th grade students in the whole population.

Table 2: Descriptive statistics for the population of 4th grade students in 2006/2007: Promoted vs. Retained (in %)

		Promoted	Retained
No. of students		99,817	6,652
% of Males		52	59
Year of Birth	up to 1995 (aged 12)	8	18
	1996 (aged 11)	17	28
	1997 on (aged 10)	76	54
Student's Nationality	Other Portuguese-speaking Countries	2	6
Mother's Nationality	Other Portuguese-speaking Countries	4	9
Mother's Academic Background	Primary (≤ ISCED 1)	43	51
	Secondary	47	47
	Higher	10	2
Student's Social Support		12	22
Computer at home		50	32
Internet at home		31	19

Table 3 includes data for the sample population of retained and promoted 4th grade students with two negative scores on national exams. In this sample 29% of the students were retained. Both groups (retained and promoted students) are very similar in most dimensions. The percentage of males, the academic background of the mother, the economic background (as proxied by the level of social support), the existence of a computer, and access to the internet

are almost the same in both groups. The variable with a greater difference in distribution between the two groups is age, which reflects whether a student has been previously retained or not. These results suggest that our sample of low-achieving students is quite homogeneous in almost all socio-economic characteristics, except for their age. We will thus focus our analysis on our sample and on a subsample of students with the "correct" age, that is, with the expected age for their grade if they have not been previously retained.

Table 3: Descriptive statistics for the sample of 4th grade students with two negative scores in 2006/2007: Promoted vs. Retained (in %)

		Promoted	Retained
No. of students		4,313	1,726
% of Males		61	62
Year of Birth	up to 1995 (aged 12)	35	13
	1996 (aged 11)	37	34
	1997 on (aged 10)	28	53
Students' Nationality	Other Portuguese-speaking Countries	4	8
Mother's Nationality	Other Portuguese-speaking Countries	7	12
Mother's Academic Background	Primary ($\leq$ ISCED 1)	57	55
	Secondary	41	43
	Higher	2	2
Student's Social Support		24	24
Computer at home		32	33
Internet at home		16	17

To analyse the impact of retention on students' future performance we conduct a regression analysis including a number of covariates that reflect the characteristics of the students and their families. As an alternative to this parametric econometric regression analysis approach we also perform a propensity score matching in order to estimate the average effect of

retaining a student on future performance. This methodology limits the comparison to retained and promoted students with the same probability of retention. We estimate both the average treatment effect (ATE) and the average treatment effect on the treated (ATET), in which the treatment, in our case, corresponds to retaining a student, and the treated group corresponds to the set of retained students. The ATET measures the estimated impact of retention on the group of students who were retained, while the ATE measures its average effect on both retained and promoted students. We also check for heterogeneous effects over different levels of the propensity score using the Smoothing-Difference estimator of Xie, Brand and Jann (2012). To ensure the robustness of our conclusions, we implement an instrumental variable approach that accounts for the possibility of some unobservable characteristics affecting students' initial retention decisions and their future performance. We discuss the validity of our proposed instrument – the culture of retention at the school level – and present the results of the regressions using the instrumental variable approach.

3. Empirical Results

3.1. Determinants of retention of low-achieving students

Before studying the impact of retention, we focus on the determinants of retention of low-achieving students. In Tables 4 and 5 we present the estimation results of logit models in which the dependent variable is Retention (equal to 1 in case of retention and 0 otherwise). In each table we show the results for our sample of low-achieving students and for the sub-sample with no previous retentions. In addition to the estimated coefficients and p-values, we also present the average marginal effect of each variable on the probability of retention. In Table 4 the results include all explanatory variables, while in Table 5 we exclude the mother's

education, since there are many missing values for this variable. All models include municipality fixed effects to control for each student's place of residence.

Table 4: Determinants of retention of low-achieving students - Logit model

Dependent Variable: Retention - 4th grade 06/07	**Sample (2 Negatives)**			**Sub-Sample (2 Negatives and No Previous Retentions)**		
	Coef.	*d* Prob / *d* x	p-value	Coef.	*d* Prob / *d* x	p-value
No previous retention	1.21	0.22	0.00			
Male	0.29	0.05	0.00	0.30	0.06	0.02
Nationality: Portuguese-speaking country	0.57	0.10	0.03	0.37	0.08	0.39
Mother's nationality: Portuguese-speaking country	0.16	0.03	0.45	0.38	0.08	0.27
Mother's education: Primary	0.37	0.07	0.00	0.61	0.13	0.00
Mother's education: Higher	0.10	0.02	0.72	0.19	0.04	0.63
Social support	0.11	0.02	0.30	0.43	0.09	0.02
Computer at home	0.09	0.02	0.37	0.15	0.03	0.34
Internet at home	-0.16	-0.03	0.23	-0.20	-0.04	0.30
Intercept	-1.86		0.01	-1.17		0.24
Municipality Fixed Effects		Yes			Yes	
No. of Observations		3616			1319	
Pseudo R^2		0.1144			0.1074	

The results obtained with our sample show that: teachers have a tendency not to retain students that have been previously retained; boys are retained more than girls, being from another Portuguese-speaking country increases the probability of being retained, and having a mother with a primary education level increases the probability of retention. In turn, benefiting from social support, having a computer, and having access to the internet do not seem to influence the probability of retention. When we restrict the analysis to children that

were not previously retained, receiving social support increases the probability of retention, while the nationality of the student no longer affects the probability of retention for the model in Table 4. The other results still hold.

Given that students in our sample and sub-sample are all low-achieving, as per their national exams scores, these results suggest that when it comes to deciding about repeating the 4th grade, there is some discrimination associated with the characteristics captured by the statistically significant variables in the logit regressions.

Table 5: Determinants of retention of low-achieving students
Logit model - drop mother's education

Dependent Variable: Retention - 4th grade	**Sample (2 Negatives)**			**Sub-Sample (2 Negatives and No Previous Retentions)**		
	Coef.	d Prob / d x	p-value	Coef.	d Prob / d x	p-value
No previous retention	1.16	0.22	0.00			
Male	0.18	0.03	0.01	0.23	0.05	0.02
Nationality: Portuguese-speaking country	0.50	0.09	0.00	0.66	0.15	0.03
Mother's nationality: Portuguese-speaking country	0.09	0.06	0.55	0.19	0.04	0.43
Social support	0.06	0.02	0.45	0.40	0.09	0.00
Computer at home	-0.07	-0.01	0.41	-0.07	-0.02	0.59
Internet at home	-0.09	-0.02	0.40	-0.09	-0.02	0.54
Intercept	-1.77		0.00	-0.46		0.53
Municipality Fixed Effects		Yes			Yes	
No. of Observations		5727			1970	
Pseudo R^2		0.0957			0.0737	

3.2. The Impact of retention on the score of subsequent exams

The output measure used to analyse the impact of retaining a low-achieving student is the score obtained on the 6th grade national exams of Portuguese and Mathematics. The students of our sample are 4th grade students enrolled in 2006/2007 who took the 6th grade exams in 2009 if they never failed, in 2010 if they were retained once, or later if they were retained more than once. Therefore, to use this score to compare the performance of students who took the exam in different years, we must guarantee an equivalent level of difficulty of the national exams across those years. The data analysis of the whole population of students, in Table 6, shows that the distribution of scores is similar in 2009 and 2010. However, that is not the case in 2011. For that reason, we restricted our sample to students who took the exams in 2009 and 2010.

Table 6: Distribution of the 6th grade national exam scores (in %)

Portuguese	1	2	3	4	5	Total
2009	0.9	10.7	52.3	28.2	7.9	100
2010	0.7	10.9	58.2	26.2	4.0	100
2011	0.4	16.6	40.0	37.4	5.6	100
Mathematics	**1**	**2**	**3**	**4**	**5**	**Total**
2009	1.7	19.6	51.2	20.3	7.2	100
2010	1.3	21.7	47.7	20.8	8.5	100
2011	3.3	33.1	30.9	25.7	7.0	100

In Table 7 we separate the sample of low-achieving students into two groups: promoted students and retained students and then compare the performance of both groups as measured by their scores on the 6th grade national exams.

Table 7: Effect of retention on performance: Scores on 6th grade national exams - Sample

Average score on the 6th grade national exams Sample (2 negatives)			
	Promoted in 06/07	**Retained in 06/07**	**Difference**
Mathematics	2.13	2.31	0.18 (0.14, 0.23)
Portuguese	2.33	2.54	0.21 (0.17, 0.25)

Note: 95% confidence intervals are presented in parentheses.

Students who repeated the 4th grade obtained slightly higher scores on the subsequent Mathematics and Portuguese exams (0.18 and 0.21, respectively, on a scale of 1 to 5). In Table 8, we restrict the analysis to the subgroup of 4th grade students who had never repeated a grade. The benefit of retention decreases to 0.10 and 0.08. For both the sample and sub-sample the differences in scores between retained and promoted, although relatively small, are statistically significant.

Table 8: Effect of retention on performance: Scores on 6th grade national exams – Sub-sample

Average score on the 6th grade national exams Sub-Sample (2 negatives and no previous retentions)			
	Promoted in 06/07	**Retained in 06/07**	**Difference**
Mathematics	2.30	2.40	0.10 (0.04, 0.16)
Portuguese	2.58	2.66	0.08 (0.03, 0.14)

Note: 95% confidence intervals are presented in parentheses.

We also estimate the effect of retention on students' performance applying a linear regression in which the dependent variable is the 6th grade exam score and the independent variable of

interest is a dummy variable for retention in the initial year 2006/2007 (equal to 1 for retained students and 0 for promoted students). We include a set of control variables, namely, whether the student has or has not been previously retained, student's gender, student's and mother's nationality, mother´s education (the default being secondary), whether or not the student has social support, a computer, and access to the internet at home. We also include municipality fixed effects in the regressions. The estimates for both groups are shown in Table 9.

The results suggest that repeating the 4th grade increases students' scores on the subsequent 6th grade national exams of Portuguese and Mathematics by 0.08 and 0.10 for the sample (on a scale of 1 to 5). For the sub-sample the impact of retention is not statistically significant. Also, students with no previous retentions obtain better scores on both exams; being a male or having a mother with a primary education level decreases the scores in Portuguese but does not affect the scores in Mathematics.

As before, we repeat the estimation excluding the mother´s education variable. These results are shown in Table 10. We now obtain positive but small effects of retention on the scores of the subsequent national exams for both the sample and sub-sample.[8]

[8] We also estimated the model considering only the observations with information on mother's education but dropping this variable from the list of controls. The results are similar to those presented in Table 9.

Table 9: Effect of retention on performance: Scores on 6th grade national exams - Least squares estimation

Dependent Variable: 6th Grade Exam Score	Sample (2 negatives)				Sub-sample (2 negatives and no previous retentions)			
	Portuguese		Mathematics		Portuguese		Mathematics	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Retained in 06/07	0.08	0.00	0.10	0.00	0.05	0.21	0.04	0.28
No previous retention	0.33	0.00	0.26	0.00				
Male	-0.12	0.00	0.01	0.66	-0.07	0.03	-0.01	0.74
Nationality: Portuguese-speaking country	0.04	0.58	0.06	0.40	-0.04	0.75	0.03	0.80
Mother's nationality: Portuguese-speaking country	0.02	0.70	-0.12	0.03	0.02	0.86	-0.10	0.28
Mother's education: Primary	-0.10	0.00	-0.05	0.12	-0.09	0.02	-0.07	0.12
Mother's education: Higher	0.11	0.15	0.09	0.34	0.13	0.23	0.13	0.33
Social support	-0.04	0.15	0.00	0.95	-0.05	0.32	0.01	0.84
Computer at home	-0.02	0.39	0.02	0.46	-0.09	0.04	-0.08	0.08
Internet at home	-0.01	0.85	-0.04	0.36	0.05	0.35	0.01	0.87
Intercept	2.50	0.00	2.12	0.00	2.73	0.00	2.36	0.00
Municipality Fixed Effects	Yes		Yes		Yes		Yes	
No of observations	2825		2824		1244		1244	
R^2	0.2173		0.1558		0.2109		0.1988	
95% Conf. Interval for Impact of Retention	(0.03, 0.13)		(0.05, 0.16)		(-0.03, 0.12)		(-0.04, 0.12)	

Table 10: Effect of retention on performance: Scores on 6th grade national exams – Least squares estimation - drop mother's education

Dependent Variable: 6th Grade Exam Score	Sample (2 negatives)				Sub-sample (2 negatives and no previous retentions)			
	Portuguese		Mathematics		Portuguese		Mathematics	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Retained in 06/07	0.10	0.00	0.12	0.00	0.07	0.03	0.07	0.05
No previous retention	0.35	0.00	0.27	0.00				
Male	-0.11	0.00	0.01	0.72	-0.06	0.05	0.03	0.44
Nationality: Portuguese-speaking country	0.01	0.83	0.00	0.95	0.00	0.99	-0.02	0.82
Mother's nationality: Portuguese-speaking country	0.05	0.26	-0.07	0.18	0.03	0.75	-0.09	0.32
Social support	-0.05	0.06	-0.01	0.81	-0.04	0.32	-0.01	0.87
Computer at home	-0.01	0.81	0.02	0.45	-0.05	0.17	-0.05	0.21
Internet at home	0.01	0.73	-0.02	0.51	0.05	0.21	0.02	0.65
Intercept	2.33	0.00	1.93	0.00	2.58	0.00	2.19	0.00
Municipality Fixed Effects	Yes		Yes		Yes		Yes	
No of observations	3691		3690		1554		1554	
R^2	0.1989		0.1392		0.1691		0.1577	
95% Conf. Interval for Impact of Retention	(0.06, 0.15)		(0.07, 0.17)		(0.01, 0.13)		(0.00, 0.14)	

To further check the robustness of our results, we carry out a propensity score matching. In Tables 11 and 12 we present the corresponding estimates of the impact of retention on the scores of the 6th grade national exams of Portuguese and Mathematics, for the sample and the sub-sample as before, with and without the mother´s education variable, and the corresponding confidence intervals.

The results are very similar: There is an increase of about 0.10 points in students' scores following retention. This effect increases slightly to 0.12 when the mother´s education variable is dropped and the number of observations increases. The effect is smaller for the sub-sample, meaning that any benefits of retention are less clear for the students that have not been retained before. These estimates are statistically significant in all cases.

Table 11: Effect of retention on performance- Scores on 6th grade
Propensity score matching

Dependent Variable: 6th Grade Exam Score		**Sample (2 negatives)**				**Sub-Sample (2 Negatives and no previous retentions)**			
		Portuguese		**Mathematics**		**Portuguese**		**Mathematics**	
		Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Matching on All Variables	ATET	0.10	0.00	0.10	0.00	0.07	0.04	0.09	0.01
	ATE	0.09	0.00	0.10	0.00	0.05	0.12	0.08	0.03
Matching on All Variables Except Mother's Academic Background	ATET	0.12	0.00	0.12	0.00	0.09	0.00	0.11	0.00
	ATE	0.12	0.00	0.12	0.00	0.08	0.00	0.10	0.00

Note: ATE - Average treatment (i.e. retention) effect, ATET- ATE on the treated (i.e. on the retained)

As a further robustness check we have included the district of residence and the scores obtained on the 4th grade national exams of Mathematics and Portuguese in the list of matching variables, to calculate the propensity score matching estimates of the ATE and ATET.[9] The conclusions remain the same.

[9] It was not possible to include the municipality of residence since the common support assumption of the propensity score matching could not be verified.

Table 12: Effect of retention on performance – Scores on 6th grade
Propensity score matching – 95% confidence intervals

Dependent Variable: 6th Grade Exam Score		Sample (2 Negatives)		Sub-Sample (2 Negatives and No Previous Retentions)	
		Portuguese	Mathematics	Portuguese	Mathematics
Matching on All Variables	ATET	(0.05 0.16)	(0.04, 0.15)	(0.00, 0.14)	(0.02, 0.16)
	ATE	(0.03, 0.14)	(0.04, 0.16)	(-0.01, 0.12)	(0.01, 0.15)
Matching on All Variables Except Mother's Academic Background	ATET	(0.08, 0.16)	(0.08, 0.17)	(0.04, 0.15)	(0.05, 0.17)
	ATE	(0.07, 0.16)	(0.07, 0.17)	(0.03, 0.14)	(0.04, 0.16)

Note: ATE - Average treatment (i.e. retention) effect, ATET- ATE on the treated (i.e. on the retained)

3.3. Allowing for further heterogeneous effects of retention

Given the small magnitude of the positive effect of retention detected above and the financial costs of having students for an additional year in school, an obvious policy implication would be for the government to implement a recommendation for public schools to decrease the average rate of retention. Such a measure would lead to the promotion of the students in our sample with a lower propensity to be retained. This possibility raises the question of whether the effect detected above holds for the particular subset of students with a lower likelihood to be retained. To check if this is a plausible assumption we analyse whether the impact of retention varies systematically with the propensity score.

We use the Smoothing-Difference estimator of heterogeneous treatment effects of Xie, Brand and Jann (2012). This analysis requires that the following conditions hold: conditional on the propensity score, there is no pre-treatment heterogeneity bias and no treatment effect

heterogeneity bias, i.e., for any particular value of the propensity score, the average potential outcomes without treatment and the average treatment effects are the same for the retained and non-retained students, respectively. To implement this analysis we first estimate separate local polynomial regressions of the outcome variable on the propensity score for the treated and non-treated groups. Heterogeneous treatment effects are then calculated as differences between the two regression lines at different levels of the propensity score over the common support.

In Figures 1 and 2 in Appendix A we present the results for the 6th grade Portuguese and Mathematics exam scores, respectively, using our sample of low-achieving students. The propensity score is estimated as before but excluding the mother's academic background to ensure a larger sample. For the case of the Portuguese exam, we observe two main features. First, for low propensity scores the local average treatment effect is always significant although small in magnitude, around 0.2. Second, the amplitude of the confidence interval in the right tail of the distribution of the propensity score is too high to guarantee a statistically significant effect of retention for those students. In the case of Mathematics the pattern is very similar. These results suggest that a reduction in retentions that affects first those with lower propensity scores would decrease their performance on subsequent exams, but this decrease would be very slight, of around 0.2 points on a scale of 1.0 to 5.0.

3.4. Instrumental variable estimation

Although we are using a matched sample and controls in the regression, there may subsist some endogeneity, since the "Retained 06/07" variable may still be influenced by the (unobserved) innate ability of the students or by the students' and their families' motivation,

which also affect students' subsequent performance. To eliminate this potential endogeneity problem, we use an instrumental variable estimation approach. As instrumental variable (IV) we use the "culture of retention" at the school level. We proxy this variable by the percentage of retained students in the 4th grade in each school. The validity of this instrument rests on the assumption that the probability of a child being retained depends largely on the cultural value of retention that prevails in the student's 4th grade school and on the fact that this "culture of retention" does not affect the student's future achievement. This assumption is justified by the fact that 4th grade is the last year of the 1st cycle in Portugal, and students in public schools move to a different school at this stage. Also, it is reasonable to assume that scores on the 6th grade national exams, which are determined by a blind-classification procedure defined by the central government, are not related to school cultural values concerning retention at the 4th grade.[10] Tables 13 and 14 show the results of the IV estimation. In all cases the first stage F-statistic is high, which supports the validity of our approach.

When we instrument retention with the prevailing "culture of retention" in the school, most of the benefits of retention estimated above disappear, with the exception of Portuguese for the sub-sample of students with no previous retentions. The results obtained for the other control variables remain unchanged. As before, when the equations are re-estimated dropping the mother's education variable, we obtain similar results for Mathematics while for Portuguese we observe benefits from retentions in both the sample and the sub-sample.

[10] Appendix B shows that the variance of the school retention rate explained by the characteristics of the students' population is very low, which justifies the use of that variable without further adjustments.

Table 13: Effect of retention on performance – Scores on 6th grade national exams – Instrumental variable: Culture of retention

Dependent Variable: 6th Grade Exam Score	Sample (2 Negatives)				Sub-Sample (2 Negatives and No Previous Retentions)			
	Portuguese		Mathematics		Portuguese		Mathematics	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Retained in 06/07	0.12	0.12	0.10	0.21	0.20	0.02	0.02	0.84
No previous retention	0.32	0.00	0.25	0.00				
Male	-0.12	0.00	0.01	0.67	-0.08	0.01	-0.01	0.67
Nationality: Portuguese-speaking country	0.03	0.62	0.06	0.37	-0.06	0.62	0.04	0.76
Mother's nationality: Portuguese-speaking country	0.02	0.67	-0.12	0.03	0.01	0.93	-0.10	0.25
Mother's education: Primary	-0.10	0.00	-0.05	0.11	-0.11	0.00	-0.07	0.08
Mother's education: Higher	0.11	0.13	0.09	0.32	0.12	0.23	0.13	0.28
Social support	-0.05	0.12	-0.00	0.92	-0.07	0.15	0.01	0.86
Computer at home	-0.03	0.33	0.02	0.47	-0.09	0.02	-0.08	0.04
Internet at home	-0.01	0.88	-0.04	0.33	0.05	0.24	0.01	0.86
Intercept	2.50	0.00	2.13	0.00	2.68	0.00	2.38	0.00
Municipality Fixed Effects	Yes		Yes		Yes		Yes	
No. of Observations	2818		2817		1240		1240	
R^2	0.2161		0.1539		0.1971		0.1984	
1st stage F-statistic	90.73		90.71		60.23		60.23	
95% Conf. Interval for Impact of Retention	(-0.03, 0.27)		(-0.06, 0.25)		(0.04, 0.36)		(-0.16, 0.19)	

Table 14: **Effect of retention on performance – Scores on 6th grade national exams - Instrumental variable: Culture of retention – Drop mother's education**

Dependent Variable: 6th Grade Exam Score	**Sample (2 Negatives)**				**Sub-Sample (2 Negatives and No Previous Retentions)**			
	Portuguese		**Mathematics**		**Portuguese**		**Mathematics**	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Retained in 06/07	0.15	0.04	0.11	0.12	0.20	0.02	0.04	0.59
No previous retention	0.34	0.00	0.27	0.00				
Male	-0.11	0.00	0.01	0.79	-0.06	0.02	0.02	0.45
Nationality: Portuguese-speaking country	0.01	0.91	0.00	0.94	-0.01	0.90	-0.02	0.84
Mother's nationality: Portuguese-speaking country	0.05	0.24	-0.07	0.16	0.02	0.80	-0.08	0.30
Social support	-0.05	0.04	-0.01	0.77	-0.06	0.14	-0.01	0.86
Computer at home	-0.01	0.70	0.02	0.49	-0.05	0.12	-0.05	0.16
Internet at home	0.02	0.60	-0.02	0.50	0.06	0.16	0.02	0.64
Intercept	2.31	0.00	1.94	0.00	2.52	0.00	2.21	0.00
Municipality Fixed Effects	Yes		Yes		Yes		Yes	
No. of Observations	3680		3679		1548		1548	
R^2	0.20		0.14		0.16		0.16	
1st stage F-statistic	116.29		116.28		40.85		40.85	
95% Conf. Interval for Impact of Retention	(0.01, 0.30)		(-0.03, 0.25)		(0.03, 0.36)		(-0.11, 0.19)	

4. Discussion and Conclusions

Our first results concern the determinants of retention within the group of low-achieving students. We show that teachers tend not to retain students who have been retained before,

that boys are retained more than girls, and that being from another Portuguese-speaking country or having a mother with only a primary education level increases the probability of retention. These results are in accordance with previous empirical studies on the determinants of retention, namely Pereira and Reis (2014) for Portugal and García-Pérez et al. (2014) for Spain. The relevance of the immigrant status on the probability of retention has also been found by Greene and Winters (2009). LiCalsi et al. (2017) also obtain that mother's education is an important determinant of retention in the context of Florida's mandatory grade retention policy. Given that we consider only students with similar levels of performance on the national exams, these results suggest that there is discrimination associated with each of these variables.

Regarding the consequences of retention of low-achieving students at the 4th grade, we find that there is a statistically significant, although small, positive impact of retention on the scores obtained at the 6th grade. This result is obtained when we match retained and promoted 4th grade students according to socioeconomic characteristics and the scores obtained on national exams. When allowing for heterogeneous average treatment effects, we find that a reduction in retentions that affects first those students with lower propensity scores would decrease slightly their performance on subsequent exams.

Finally, when we control for possible endogeneity due to unobservables, using the culture of retentions at the 4th grade school as an instrumental variable, we find that the overall effect of retaining low-achieving 4th grade students on their 6th grade exams scores is not statistically significant for Mathematics. In the case of Portuguese, the impact is statistically significant only for students with no previous retentions, although it is small, with an estimated magnitude of around 0.2 on a scale of 1 to 5.

The above results of either small benefits or no statistically significant effects of retention are in line with the findings of no significant effects in the meta-analysis of Allen et al. (2009) for the subset of studies using more solid methodological designs, and of slightly positive effects determined by Greene and Winters (2009) and Roderick and Nagaoka (2005) for 3rd graders. Schwerdt et al. (2015) find positive effects in the first three years after retention (that become insignificant within five years). Notice, however, that in these latter studies, conducted for Florida, retentions were accompanied by additional educational support measures after the retention took place, which is not the standard in Portugal. Therefore it is natural that the impact of retentions is stronger in their case.

In summary, the main finding of this paper is that the impact of early retention is either not statistically significant or of a small positive magnitude. Taking into account the high costs of maintaining students in school for one extra year, the small benefit from retention we obtained suggests that repetition is an ineffective tool to deal with under-achievement at early stages. Thus it would be interesting to implement experiments to evaluate and compare the impact of alternative measures to promote the success of low-achieving students, such as extra hours of teacher support, mentoring, summer schools, and preferential assignment to high performing teachers. These results are especially important for countries with high retention rates that are considering alternative educational policies to promote students' success.

References

Allen, C. S., Chen, Q., Willson, V.L., and Hughes, J.N., 2009, "Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-analytic, Multi-level Analysis" Educational Evaluation and Policy Analysis, 31(4), pp. 480-499.

Eurydice, 2011, "Grade Retention During Compulsory Education in Europe: Regulations and Statistics", Education, Audiovisual and Culture Executive Agency (EACEA P9 Eurydice).

García-Pérez, J.I., Hidalgo-Hidalgo, M., and Robles-Zurita, J.A., 2014, "Does grade retention affect students' achievement? Some evidence from Spain". Applied Economics, 46(12), pp.1373-1392.

Greene, J. P. and Winters, M. A., 2007, "Revisiting grade retention: An evaluation of Florida's test-based promotion policy". Education Finance and Policy, 2(4), pp. 319-340.

Greene, J. P. and Winters, M.A., 2009, "The effects of exemptions to Florida's test-based promotion policy: Who is retained? Who benefits academically?" Economics of Education Review, 28(1), pp. 135-142.

Holmes, C.T. and Matthews, K.M., 1984, "The effects of non-promotion on elementary and junior high school pupils: A meta-analysis." Review of Educational Research 54(2), pp. 225-236.

LiCalsi, C., Ozek, U. and Figlio, D., 2017, "The Uneven Implementation of Universal School Policies: Maternal Education and Florida's Mandatory Grade Retention Policy." Education Finance and Policy, 1-53.

Jacob, B.A. and Lefgren, L., 2004, “Remedial Education and Student Achievement: A Regression-Discontinuity Analysis”, The Review of Economics and Statistics, 86(1), pp. 226-244.

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

Jimerson, S. R., 2001, “Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century”, School Psychology Review, 30(3), pp. 420-437.

OECD, 2013, “PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)”, PISA, OECD Publishing.

OECD, 2014, “Are disadvantaged students more likely to repeat grades?”, PISA in Focus, 43, 2014/09 (September), 1-4.

Pereira, M.C. and Reis, H., 2014, “Grade retention during basic education in Portugal: determinants and impact on student achievement”, Economic Bulletin, Banco de Portugal, June 2014, pp. 61-82.

Roderick, M. and Nagaoka, J., 2005, “Retention Under Chicago's High-Stakes Testing Program: Helpful, Harmful, or Harmless?” Educational Evaluation and Policy Analysis, 27(4), pp. 309-340.

Schwerdt, G., West, M. R., and Winters, M. A., 2015, “The Effects of Test-Based Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida”, NBER Working-Paper 21509, August.

Xie, Y., Brand, J. E., and Jann, B., 2012, “Estimating heterogeneous treatment effects with observational data”, Sociological Methodology, 42(1), pp. 314-347.

Appendix A

Figure 1. Heterogeneous effect of retention on the 6th grade Portuguese exam score: Heterogeneous ATE = E(score if retained - score if promoted | propensity score) for the sample.

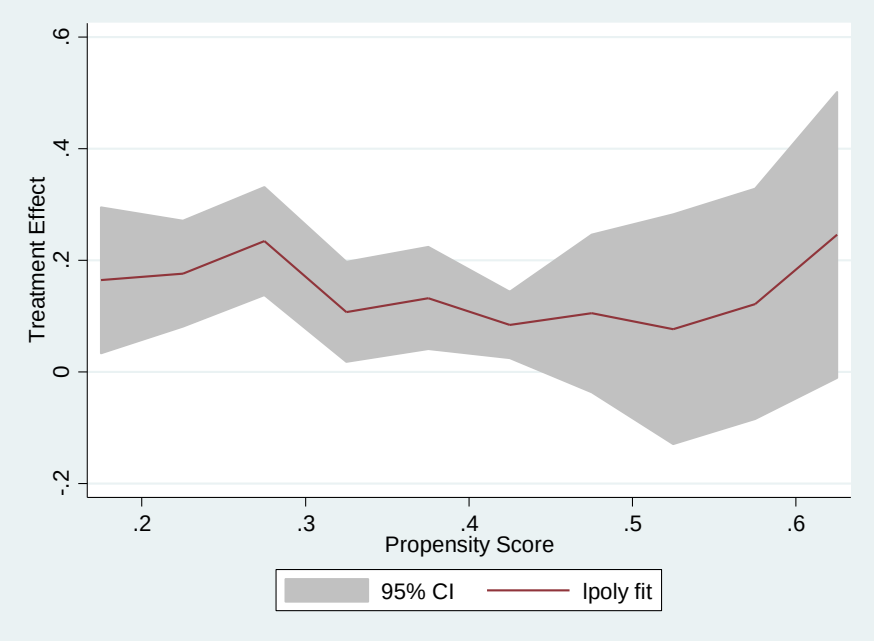
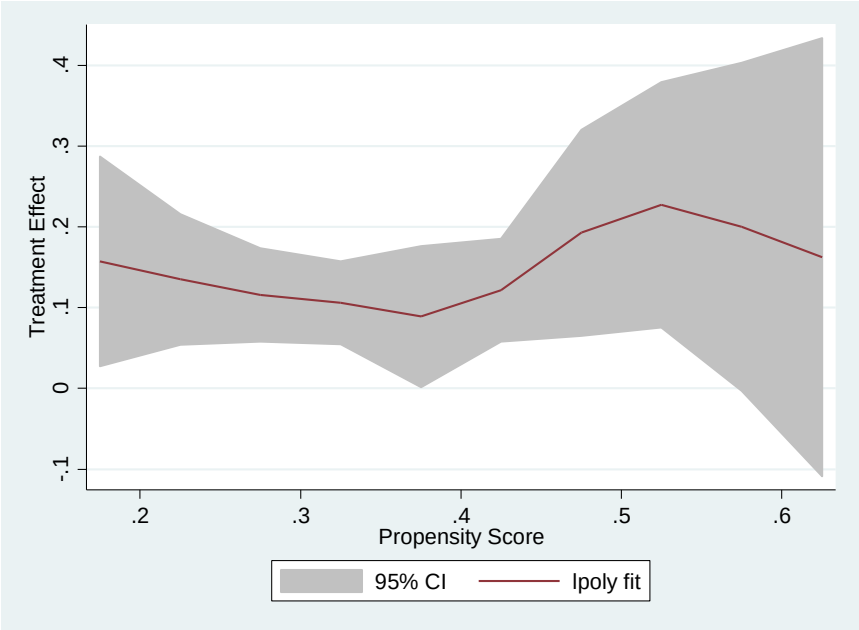


Figure 2. Heterogeneous effect of retention on the 6th grade Mathematics exam score: Heterogeneous ATE = E(score if retained - score if promoted | propensity score) for the sample.



Appendix B

Table B.1: Determinants of school retention rates - Least squares estimation

Dependent Variable: School Retention Rate 4th Grade 06/07		
	Coef.	p-value
Male	0.01	0.46
Nationality: Portuguese-speaking country	0.03	0.65
Mother's nationality: Portuguese-speaking country	0.08	0.12
Mother's education: Primary	0.00	0.82
Mother's Education: Higher	-0.09	0.00
Social support	0.05	0.00
Computer at home	-0.03	0.01
Internet at home	0.00	0.78
Intercept	0.09	0.00
Municipality Fixed Effects	Yes	
No of observations	5800	
R^2	0.0948	
F-statistic	2.03	