Grade Repetition and Household Responses in a Low Income Setting

Tahir Andrabi,[*] Ethan Matlin,[†] Gabrielle Vasey[‡]

October 22, 2024

Abstract

Millions of students around the world are retained and repeat a grade each year, yet we may not fully understand the consequences of this choice. Especially in low-income countries, where the majority of repeaters reside and the decision is often at the teacher's discretion, it is important to understand how parents, students, and schools react to this negative information shock. In this paper, we use rich matched household-child-school panel data from Pakistan to study the dynamics of the beliefs, investments, and outcomes of these three agents following a student not being promoted to the next grade. We find that following grade repetition, parents revise downward their expectations, beliefs, and investments. Students are discouraged by the retention, and decrease their beliefs in the value of study effort. Conversely, schools do not play a large role, and teachers have no negative bias towards students who are repeating. Overall, we find negative effects of repetition: repeaters score -0.27 to -0.44 standard deviations worse in math, English, and Urdu and are 7.1 percentage points more likely to drop out than their peers.

Keywords: Education Policy, Achievement, Grade Retention, Parental Investments.

[*]Pomona College
[†]Harvard University
[‡]Concordia University

1 Introduction

Each year, approximately 24 million children around the world repeat at least one grade during primary school, the bulk of this occurring in low- and middle-income countries (UNESCO 2024).[1] These numbers, which are likely an underestimate due to missing data on private schools and multi-grade classrooms, reflect and relate to what the World Bank has called a "learning crisis" (UNESCO [2012]). Holding back poor-performing students in their current grade, a common practice known as grade repetition or retention, is intended to allow students additional time to master basic material before progressing to more difficult material (Pritchett and Beatty [2015], Banerji [2000]). However, research has shown that grade repetition can have a negative effect on students' test scores and longer-term outcomes (Eren, Depew and Barnes [2017], Manacorda [2012], Jacob and Lefgren [2009], Eide and Showalter [2001], Eren, Lovenheim and Mocan [2018])

Despite the disproportionate prevalence of grade repetition in low- and middle-income countries, much of the existing research on grade repetition focuses on (1) students in high income countries, (2) settings where there is a strict cutoff for holding a student back, and (3) total average effects of repetition on academic achievement without considering how parents, students, and teachers respond. Previous literature has shown that grade repetition has negative effects on student learning, however our understanding of *why* that occurs, and the role that parents, students, and teachers play is far more limited (e.g. Eren, Lovenheim and Mocan [2018], Figlio and Özek [2019], Jacob [2005], Jacob and Lefgren [2004], Jacob and Lefgren [2009], Eide and Showalter [2001]). Including the household in the analysis is vital, yet having access to the necessary data is rare.

This paper studies the academic effects of grade repetition and how the ecosystem around a student—her parents, teachers, schools, and the student herself—respond to being held back. Do these actors mitigate or aggravate the initial repetition shock? To answer this question, we use a matched household-child-school panel from the Learning and Education Achievement in Punjab (LEAPS) project. These data allow us to understand household reactions to grade repetition (including potential resource reallocation), something that would be quite challenging in another setting.

[1] Data from 2022 `http://data.uis.unesco.org`

There are minimal guidelines on which students should be retained in primary school in Pakistan.[2] For this reason, we analyze plausibly idiosyncratic variation in repetition decisions, both within and across teachers. This variation may occur because teachers are not trained to make grade promotion decisions and may lack concrete data to make such decisions (Brophy [2006]). We show that conditioning on appropriate measures of student ability (test scores, teacher ratings, parent observations) leaves residual variation which may plausibly stem from judgmental error.[3] Our primary specification controls for two lags of test scores entered in flexible functional form to capture both the level and rate of student learning (Todd and Wolpin [2003] and Andrabi et al. [2011]). Qualitative evidence suggests that this specification is appropriate: 98% of sampled schools cited academic issues as the most common reasons for grade repetition.[4] However, we also show that our results are robust to alternative measures of student ability including ratings of students by their parents and teachers as well as to consideration of a variety of previous educational inputs at the household-, school-, and student-level.

Consistent with prior work, grade repetition significantly decreases student performance. The year following their retention, repeaters score 0.44, 0.28, and 0.37 standard deviations worse in math, English, and Urdu and are 7 percentage points more likely to drop out of school. We find very little differences by gender, both in the probability of being retained and the impact of retention on grades and dropout.

Turning to mechanisms, we find that it is crucial to consider household reactions to understand the full repercussions of repetition. Grade repetition reduces parents' expectations, household investments in education, and children's own beliefs about the benefits of exerting effort in school. Following repetition, parents expect their child to answer 5/100 fewer questions correctly and evaluate their child 0.13 to 0.33 standard deviations lower on measures of school performance, work ethic, and intelligence. They then reduce annual

[2]Until 3rd grade, Pakistan has automatic social promotion in government schools. However, our sample of students moving from 4th to 5th grade are not subject to this policy (Chohan and Qadir [2011]

[3]Our strategy differs from much of the literature, which uses empirical approaches including regression discontinuity based on test score/grade cutoffs (Figlio and Özek [2019], Eren, Lovenheim and Mocan [2018], Schwerdt, West and Winters [2017], Jacob and Lefgren [2009], Jacob and Lefgren [2004], Mariano and Martorell [2013], Manacorda [2012]), IV techniques using day of birth (Eide and Showalter [2001], Mahjoub [2017]), and time-variation in the implementation of repetition policies (Allensworth [2005], Koppensteiner [2014], Jacob [2005]). This is mainly due to institutional setting. These alternative techniques are much more well-suited to well-established thresholds and rules for repetition which exist in most high-income countries, but not in many low- and middle-income countries.

[4]The remaining 2% cited not paying school fees, entering school late in the year, or age.

educational expenditures by 176 PKR. Students who repeat are 26 percentage points less likely to believe that studying hard in school will help them find a better job in the future. These mechanisms help to explain why grade repetition causes test scores to decrease. Schools are less interesting as an explanatory mechanism: teachers at schools seem to treat repetition as an uninformative shock and then get back to business as usual. On the other hand, households and students persistently revise downwards their expectations, beliefs, and investments.

This research contributes to several fields of study. First, we contribute to the literature on grade repetition in low- and middle-income countries (Koppensteiner [2014], Manacorda [2012], Gomes-Neto and Hanushek [1994]). Grade repetition is very common in developing world, much more so than in the U.S. While there is a large literature studying grade repetition in the United States and other high-income countries (e.g. Borgheesan, Reis and Todd [2022], Eren, Lovenheim and Mocan [2018], Figlio and Özek [2019], Jacob [2005], Jacob and Lefgren [2004], Jacob and Lefgren [2009], Eide and Showalter [2001]), high-income countries have more well-established standards for grade repetition, and parents may respond differently in a different setting. The Pakistani educational landscape (with no policy dictating criteria for grade repetition,[5] large intra-classroom age ranges,[6] and a large but unregulated private school sector) more closely resembles other low- and middle-income nations. It is not clear if results from the United States will extend to students in developing countries due to differences in school systems, access to resources, support from schools, and parental attitudes. Our data from Pakistan allow us to consider grade repetition and the case of negative shocks in a setting similar to that faced by many children around the world.

Second, we contribute to the literature on parent beliefs of child ability, and how this interacts with investments and responds to negative shocks. There has long been interest in understanding how parental investments reinforce or compensate differences in child characteristics (Becker and Tomes [1976], Tomes [1981], Griliches [1979], Behrman, Rosenzweig and Taubman [1994], Behrman, Pollak and Taubman [1986], Pitt, Rosenzweig and Hassan [1990], Rosenzweig and Zhang [2009], Kinsler and Pavan [2021]). More recently, several papers have used RCTs to study how parental beliefs respond to information about their

[5] Although Pakistan does have automatic promotion at government schools until 3rd grade, our sample of students who would move from 4th to 5th grade aren't subject to this policy (Chohan and Qadir [2011].

[6] Almost 4 years among third-grade classrooms

children, and how this impacts investments and outcomes (Dizon-Ross [2019], Gan [2021], Barrera-Osorio et al. [2020], Bergman [2021]).[7] Our work is very related to these studies, and brings non-experimental results to complement the experimental studies.

The paper proceeds as follows. Section 2 introduces the LEAPS panel data and institutional details for grade repetition in Pakistan. Section 3 details our empirical strategy. Section 4 presents results. Section 5 concludes.

2 Data and Setting

2.1 Description of Data

We use data from the Learning and Education Achievement in Punjab Schools project, a five-year survey of schools, school-children, and their households in the Attock, Faisalabad, and Rahim Yar Khan districts of Punjab province conducted between 2004-2011. Surveyors randomly sampled 112 villages with at least one private school from the three districts.[8] Within each village, they administered questionnaires to students, teachers, head teachers, households.

Surveyors also proctored tests in three subjects (Urdu, English, and Math) to assess student knowledge and learning. These tests were designed to cover mostly basic topics typically learned in lower grades.[9] The Urdu and English tests covered letters, word-recognition, sentence construction, and reading comprehension. The Math tests covered counting, addition, subtraction, multiplication, division, fractions, and word problems. Tests were scored using item response theory and normalized around zero using the first year score

[7] Note that there are two other literatures to which we are tangentially related. First is the large literature on how parents respond to information about schools (Andrabi, Das and Khwaja [2017], Allende, Gallego and Neilson [2019] Hastings and Weinstein [2008]). Our paper studies how parents respond to information about their *children*. There is also a small literature, which we contribute to, on students beliefs and agency in response to negative information shocks. This literature has mostly focused on psychological interventions for college students in the U.S. to improve mindset and social belonging (Chevalier et al. [2009], Stinebrickner and Stinebrickner [2012], (Walton and Cohen [2011], Yeager et al. [2016], Dobronyi, Oreopoulos and Petronijevic [2019], Oreopoulos and Petronijevic [2019]).

[8] Since, for the most part, only larger villages contain private schools, there exists a potential selection bias, however due to the increasing private school market, this sample is likely to be more representative of the future (Das, Pandey and Zajonc [2012]).

[9] This decision in test construction was made to reflect students' lagged learning (rather than trying to assess where grade-level knowledge).

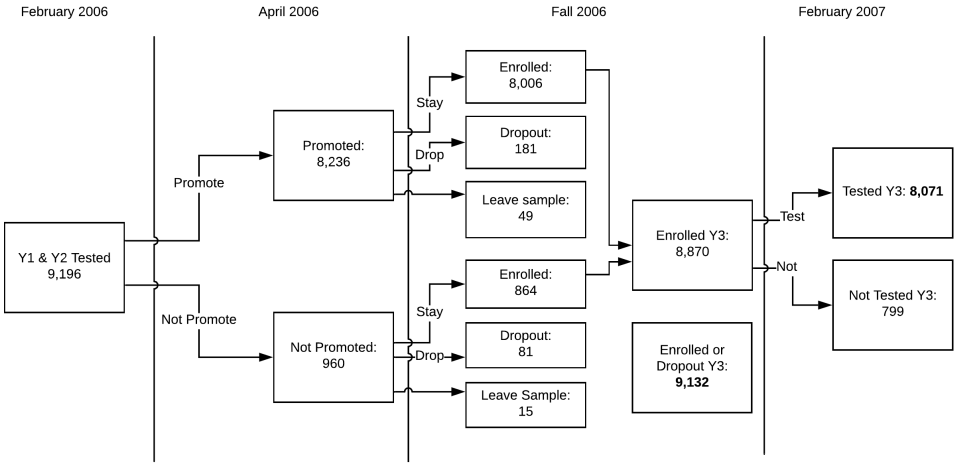
distribution (future years are in units standard deviations of the first year score distribution). Teachers did not receive results of these tests, therefore could not incorporate test scores into their promotion/retention decisions.

In addition to the test score and enrollment data, we have rich demographic and survey data. For all students, we know their sex, age, and mother's education level. For a subsample of students, we also have data on their height and weight, and their feelings and attitudes towards school. From a matched household survey given to a subsample of students, we also have data on parental perceptions of their child, an asset-based wealth index, detailed educational expenditures, assessments of their child's teacher quality, and expectations of their child's ability. Finally, teachers were surveyed and answered the following question about a subsample of students in their class: "On a scale of 1 to 10, how good would you say that this student is in his/her studies?"

In the first year of the survey, 13,735 students in 804 schools which offered third grade instruction were surveyed. Of these schools, 485 (60%) were government schools and 319 private (40%).[10] At sampled schools, every 3rd grade student was sampled and we follow these initial students over time without adding new students to the sample (see figure 1). Since the majority of students in our sample repeat between 4th and 5th grade, we focus our analysis on the 2nd and 3rd years of the survey.

[10]More specifically, there were 303 private schools and 16 NGO/Trust schools, however we group these schools together for our analysis.

Figure 1: Evolution of the Sample



Note: There are 9,196 students between 7.5 and 15.5 years old in year 2 (4th graders are usually 9-10) with non-missing promotion status and test scores for both years 1 and 2. After year 2 (baseline), the teacher/school decided on each student's promotion status: should they be demoted to a lower grade, retained in their current grade, or promoted to the next grade? Promoted means the student advanced to the next grade while not promoted means the student was either retained in the current grade or demoted. 8,236 are promoted of which 8,006 remain enrolled and 181 dropout. 960 are not promoted of which 864 remain enrolled and 81 dropout. Thus, in year 3, there are 8,870 total enrolled students. Of these, 8,071 students are tested in year 3.

We primarily use four subsamples of the data. For our results on academic outcomes, the subsamples are students who were tested in all three years (N=8,071)[11], students who were tested all three years *and* whose teachers were surveyed about their performance (N=5,529), students who were tested in all three years *and* matched to households given a more detailed household survey (N=698). For the analysis of mechanism and household responses, we use data on students who were tested in the first two years and matched to household given the more detailed household survey (N=741). Table 1 presents baseline characteristics of students. The patterns across each subsample are relatively similar: 7-9% of students repeat their current grade in year 3, just under half are female, and most students are 10 years old. Both teachers and parents slightly inflate student-ratings, however not overly so: teachers rate the average student $6/10$ while parents rate their children $3.3/5$. Finally, there appear to be no overwhelming differences between the subsamples (as we would expect since the subsamples are random).

[11] This number is lower since some students were either new to the sample in the second year of the survey or were absent on the day of the test in year 1

Table 1: Summary Statistics

Variable	Child Sample	Teacher Sample	Parent Sample	Mech Sample
English score (Pre-)	0.306 (0.937)	0.363 (0.923)	0.252 (0.975)	0.238 (0.989)
Math score (Pre-)	0.254 (1.131)	0.304 (1.101)	0.304 (1.044)	0.287 (1.070)
Urdu score (Pre-)	0.380 (0.960)	0.420 (0.948)	0.373 (0.946)	0.352 (0.975)
Not Promoted	0.088 (0.284)	0.077 (0.266)	0.070 (0.256)	0.076 (0.264)
Female	0.467 (0.499)	0.470 (0.499)	0.446 (0.497)	0.441 (0.497)
Age	10.484 (1.394)	10.576 (1.404)	10.517 (1.409)	10.553 (1.488)
Mom Educated	0.277 (0.448)	0.355 (0.478)	0.287 (0.452)	0.265 (0.441)
Rating by teacher	- (-)	6.098 (2.323)	6.255 (2.354)	6.267 (2.377)
Parental Perception of Whether Child is Hard Working	- (-)	- (-)	3.304 (0.622)	3.294 (0.632)
Parental Perception of Child Intelligence	- (-)	- (-)	3.340 (0.621)	3.333 (0.625)
Inteligence Rating	-	-	-	3.333 (0.625)
Hardworking Rating	-	-	-	3.294 (0.632)
School Performance Rating	-	-	-	3.296 (0.656)
Educational Expenditures (short-run)	-	-	-	133.142 (100.163)
Lagged Educational Expenditures (long-run)	-	-	-	916.161 (413.651)
Family Help (any)	-	-	-	0.340 (0.474)
Hours of help with studies from family (median)	-	-	-	2.452 (3.887)
Hours reading or telling stories (median)	-	-	-	0.000 (0.532)
Any tutoring	-	-	-	0.182 (0.386)
Hrs. Tutoring	-	-	-	0.000 (4.749)
Know name of child's teacher	-	-	-	0.565 (0.496)
Met child's teacher	-	-	-	0.591 (0.492)
Look HW	-	-	-	0.560 (0.497)
Observations	8071	5529	698	741

Note: Means and standard deviations of baseline student characteristics. Column (1) includes all students with three years of test scores. Column (2) includes all students with three years of test scores whose teacher also rated their performance. Column (3) includes all students with three years of test scores matched to households that completed a more detailed household survey. Column (4) includes all students with two years of test scores matched to households that completed a more detailed household survey.

2.2 Institutional Setting: Grade Repetition in Pakistan

The setting for this paper is very different than in high-income countries like the U.S. Education levels in our sample are in general low: 72% of mothers in our sample did not complete primary school and 64% received no formal education. Mothers with no formal education spend ~ 0 minutes on a typical day for child educational needs at home. Learning levels for students are also low: by the end of 3rd grade, just over 50 percent of children have mastered the Mathematics curriculum for 1st grade (Andrabi et al. [2007]).

Schools in this setting also look very different than those in high income countries on average. Classes and schools are relatively small: the median class has 13 students and the median school has 119 students (across multiple grades). Classrooms have large age-ranges of students: the average classroom has a three to four year gap between the youngest and oldest student. There is a large low-cost private school sector, making up 40% of the schools in our sample and costing only a dime a day.

Grade repetition after fourth grade is common ($\sim 9\%$ in our sample), and teachers are mainly responsible for deciding who is retained. Although Pakistan does have automatic promotion at government schools until 3rd grade, our sample of students who would move from 4th to 5th grade aren't subject to this policy (Chohan and Qadir [2011]. This differs from high-income countries, which in general have lower repetition rates (Eisemon [1997]).[12]

2.3 How is grade repetition decided?

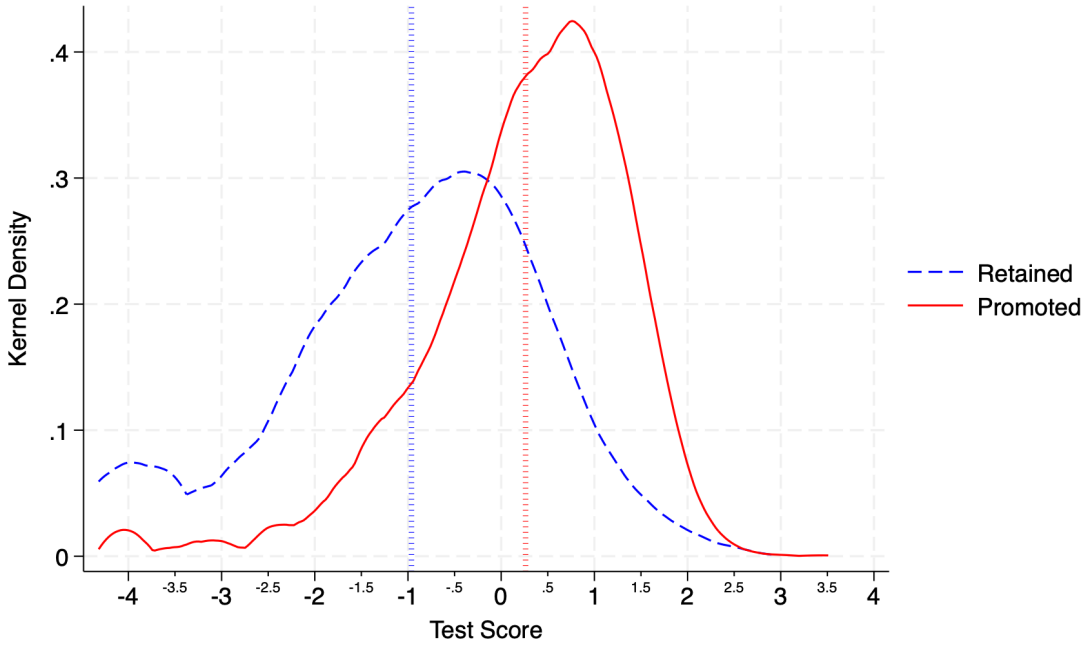
Classroom teachers are central to making judgements for whether students should be retained or promoted to the next grade. In the LEAPS school survey, 72% of schools reported that the classroom teacher played a very important role and 25% an important role. This is in contrast to head teachers (38% very important and 54% important), parents (12% very important and 25% important), and children (2% very important and 2.5% important). This differs from other settings with centrally-determined thresholds for grade repetition, which are more common in higher-income countries. No matter who is responsible for making the decision,

[12]Countries with clearer and properly enforced standards for repetition seem to have lower repetition rates than those with no or loosely-enforced state-mandated criteria for repetition. Repetition rates are highest in Sub-Saharan Africa (approximately 22% of primary students repeat a grade at any given time) and also high in the Middle East and North Africa (12%), however in developed countries only 1-5% ofstudents repeat a grade at any given time ((Eisemon [1997])).

nearly all schools (98%) state academic issues as the most common reasons for repetition.

While on average, retained students have lower pre-retention test scores and teacher ratings (appendix table A.1), there is considerable overlap in the pre-retention test scores distributions for repeaters and non-repeaters (see figure 2). This occurs due to both across-teacher differences in the "standards for repetition" and within-teacher lack of consistency.[13]

Figure 2: Distribution of math test scores for retained and promoted students



Note: Kernel density estimation of the distribution of math test scores for retained students (in blue) and promoted students (in red). The vertical lines represent the mean test score for each group.

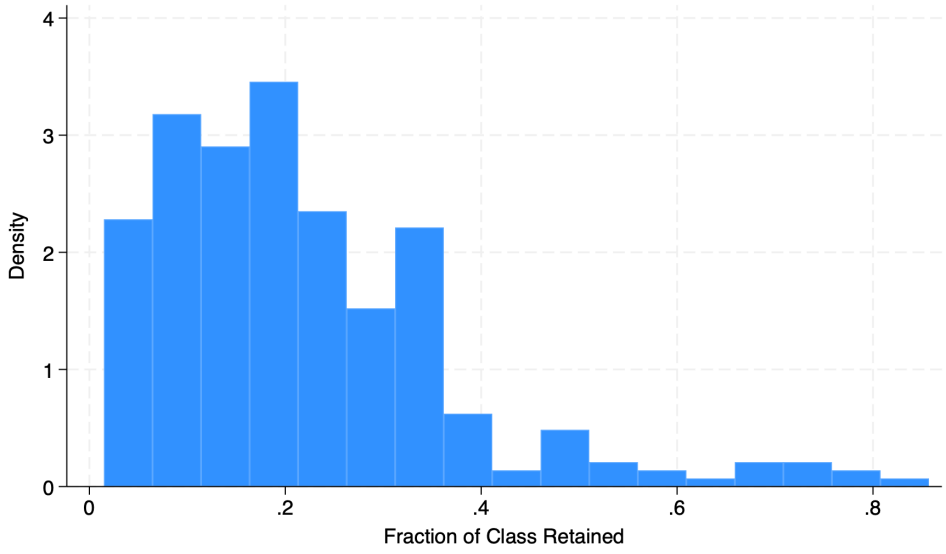
Across teachers, there is substantial variation in the "standards for repetition" which can be measured by looking at the fraction of each classroom which is retained. Figure 3 shows that there is substantial variation in the fraction of each classroom which is retained.[14] This pattern holds even after conditioning on teacher rating. Figure 4 separately considers the subset of students in each class that a student rates poorly,

[13] While the cross-teacher variation may sound reminiscent of a judge IV design, we cannot use that design in this setting for two primary reasons. First, the monotonicity assumption is violated — teachers are unlikely to have a common ranking for all students and differ only due to their cutoff. The lack of consistency within-teacher is evidence for this. Second, the exclusion restriction is violated — while judges affect outcomes only through sentencing, teachers also teacher students and may affect learning trajectory.

[14] Figure 3 is conditional on the teacher retaining any students. Appendix figure A.1 shows the histogram with zeros.

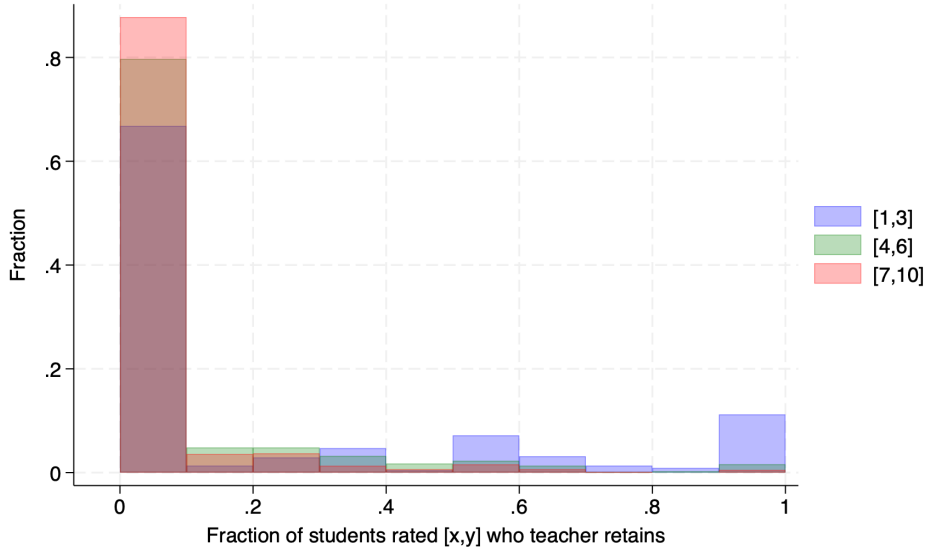
moderately, and highly and plots the histogram of the fraction of each group who is retained. We see that while teachers are generally more likely to retain students who they evaluate poorly, some teachers are more likely to retain a student who they rated poorly than others.

Figure 3: Across-teacher distribution of fraction of class retained



Note: Histogram of the fraction of each teacher's classroom who is retained (conditional on retaining any students).

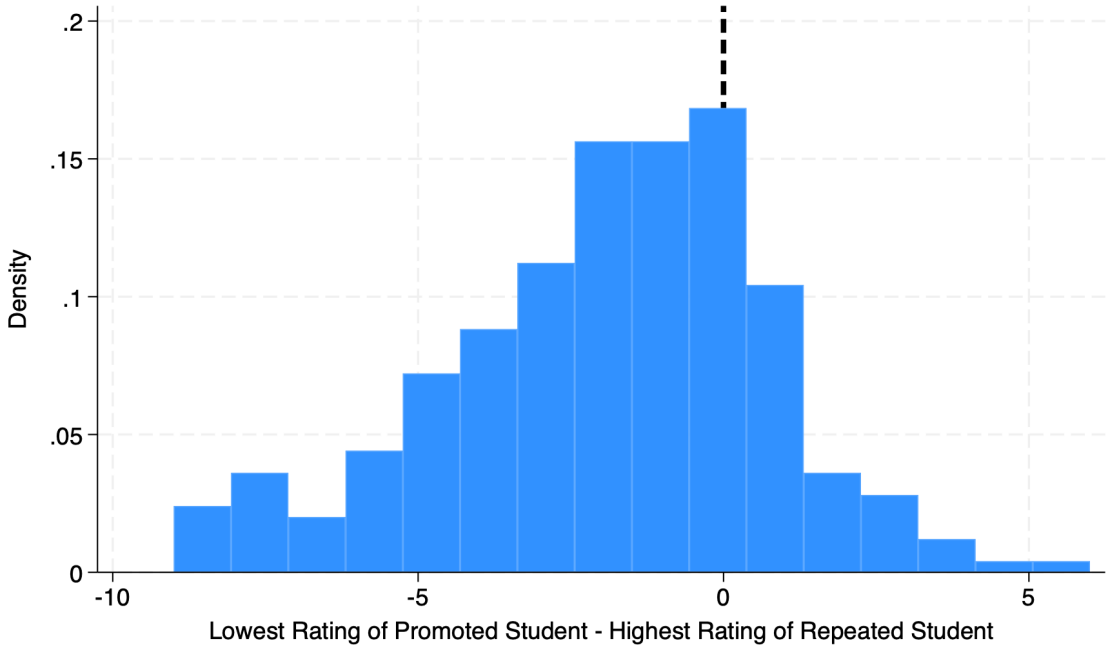
Figure 4: Across-teacher distribution of fraction of class retained, conditional on teacher rating



Note: Conditional on teacher rating, histogram of the fraction of each type of student who is reatined across classrooms (conditional on retaining any students).

There is also within-teacher inconsistency in the types of students who are retained. For example, appendix figure A.2 shows that one teacher retained students who she rated as 5, 6, and 8 but promoted students who she rated as 5, 7, 8, 9, and 10. We can aggregate these types of inconsistencies across teachers by computing the difference between the lowest teacher-rating of a promoted student and the highest teacher-rating of a retained student given by the same teacher. Figure 5 shows that there is a substantial mass of teachers (those with negative values) who retain students who they rated higher than students who they promote.

Figure 5: Across-teacher distribution of difference between lowest-rated promoted student and highest-rated retained student



Note: Across-teacher histogram of the within teacher difference between the teacher rating of that teacher's highest-rated retained student and the teacher rating of that teacher's lowest-rated promoted student. A negative value means that the teacher promoted a student who they rated worse than a student who they retained. A positive value means that the teacher did not retain any students who they rated more highly than a student who they promoted.

In a different setting with a centralized retention cutoff decision, whether a student repeats or not can be predicted with a high level of precision using the test score as an explanatory variable (or any other measures used by the school in the retention decision). In our setting, nearly all schools state that academic issues are the main reason for retention, yet the first row of Table 2 shows a very low R^2 value (0.18) when we estimate

a logit regression with the retained variable as the dependent variable and the child's test scores (two lags), age, sex, and mother's education as independent variables. Even including teacher fixed effects to allow for different standards in repetition still explains only results in an R^2 of 0.35, as shown in the second column of R^2 values.

Although the schools did not report it, it is possible that they are using other information besides student ability to make their retention decisions. Using our rich survey data, we estimate more detailed and flexible specifications in an attempt to explain the variation in the retention decision. Going down the rows in Table 2, we control for the following explanatory variables: "Extended Child" contains information on child height and weight and the wealth of the family, "Teacher" contains the teacher rating of the student, and "Parent" contains parental perceptions of the child. Without teacher fixed effects, even the most flexible specification predicts just over half of the variation in the retention decision. Although the R^2 does get fairly high with teacher fixed effects, it is crucial to also consider the adjusted R^2, as the number of covariates gets large compared to the sample size in several of the specifications.

Table 2: R^2 from regression predicting regression using child and teacher characteristics

		Teacher FEs		
Controls	R^2	R^2	Adjusted R^2	N
Child	.18	.35	.27	8071
Extended Child	.17	.37	.25	5442
Teacher	.2	.38	.27	5529
Teacher $\times$ Test Scores	.21	.39	.27	5529
Parent	.37	.66	.26	698
Parent $\times$ Test Scores	.48	.69	.27	698
Parent $\times$ Test Scores + Parent $\times$ Teacher	.70	.85	.25	487

Note: The table shows R^2 from logit regressions of the indicator variable for repetition on child and teacher characteristics. "Child" includes the child age and sex, as well as lagged test scores. The first column of R^2 values have no fixed effects, whereas the second column of R^2 values have teacher fixed effects.

These regressions show that within- and across-teacher variation exists even after controlling for a rich set of observed characteristics. Because of this, it is possible for there to be comparable students who face different retention decisions simply because they were assigned different teachers or their teacher made a judgement error. The full regression tables are shown in appendix table A.4.

3 Empirical Strategy

The intuition behind our empirical specification is that we need to compare two students who would have scored similarly in year $t+1$ if both progressed normally, however, in the data, one student was retained after year t. By comparing their outcomes in year $t+1$, we then see the impact of retention. The challenge in this estimation strategy is that we cannot see the counterfactual scores in $t+1$ when neither student was retained. Our solution is to estimate a very flexible specification with many controls, to compare similar students to the best of our ability. Below, we explain our strategy in detail and go over potential threats to the identification. For the majority of the explanations we will consider the outcome to be test scores, but the same specification will be used with many different outcome variables in the results.

Consider a student i in school j. Their test score for period $t+1$ is given by $y_{i,j,t+1}$. It is reasonable to assume that $y_{i,j,t+1}$ will depend on the student's ability, their home inputs, and their school inputs, not just in year t but in previous years as well (Todd and Wolpin [2003]). Since we do not have data on the complete history of the student, we assume that three elements are sufficient to capture this information: (1) the ability level of the student in time t, (2) the learning rate of the student between time $t-1$ and time t, and (3) school fixed effects. We also control for some time invariant student characteristics as well as teacher ratings in time t, to capture any observable differences between repeaters and those who are not retained.

More formally, we estimate the following equation,

$$y_{i,j,t+1} = \beta R_i + f(y_{i,j,t}, y_{i,j,t-1}) + X_i\gamma + \mu_j + \varepsilon_{i,j,t} \tag{1}$$

where $y_{i,j,t+1}$ is an outcome for student i in school j in year $t+1$, R_i is an indicator variable equal to one if the student was retained after year t and zero if the student was promoted, $f(\cdot,\cdot)$ is a quartic additive function of student i's lagged and twice-lagged test scores, X_i is a vector of additional student-level controls and their teacher rating, μ_j is a school fixed effect, and $\varepsilon_{i,j,t}$ is an idiosyncratic error.

Our identifying assumption is that conditional on lagged and twice-lagged test scores (as well as school fixed effects and additional student-level controls), repeating and non-repeating students would score similarly if not for one being retained and the other promoted. We assume that (1) the lagged score captures

all permanent and time-varying effects from the prior year and (2) learning speed is approximated by the difference between the lagged and twice-lagged test scores.

This means that we are identifying the effect of grade repetition from plausibly random variation in teachers' promote/retain decisions, i.e. mistakes and cross-teacher differences in standards. This is plausible because teachers are trained to teach, not assess student abilities, and may lack sufficient information to do so (Brophy [2006]). Without clearly defined achievement standards and corresponding mechanisms of assessment (which may not exist in many low-income countries), teachers' decisions should be relatively arbitrary (and we wouldn't expect that these arbitrary judgments systematically point in one direction or another). Therefore, their promotion decisions may be made with idiosyncratic error (this is especially true due to the binary nature of the retention decision; a small change in a teacher's noisy measurement of student ability may push a student into promotion or retention). Section 2.3 presents evidence that there is substantial plausibly random variation, both within and across teachers, in which students are retained and which are promoted.

The main threat to our identification strategy is if teachers are making their retention decision based some variable that is unobserved to us, and this variable also impacts test scores in $t+1$.[15] In this scenario, what we are estimating to be the impact of retention would be the combined effect of retention and the unobserved variable. Because we are controlling for test scores in t and $t-1$, the unobserved variable would have to be a shock that only effects learning in $t+1$ and not in prior periods. It is unlikely that there is such a one-period shock that affects 9% of the sample, and that is not captured in any of the teacher and parent variables that are contained in our data, so we do not think that this scenario is likely.

4 Results

We begin this section by confirming prior research that grade repetition reduces academic outcomes. To understand the mechanisms by which this occurs, we then focus on the reactions of parents, teachers, and students themselves to grade repetition.

[15] We are also assuming that two lags of test scores adequately capture a student's ability, however this assumption is fairly common in the education literature (Chetty, Friedman and Rockoff [2014]).

4.1 How are outcomes impacted by repeating?

Table 3 shows that repetition significantly decreases test scores and increases the probability of dropping out the year following the retention decision. The large and significant repetition coefficient represents the achievement gap between repeaters and non-repeaters of similar initial ability: .44 standard deviations for math, .27 for English, and .37 standard deviations for Urdu. The effect on test scores persists for at least two years, suggesting that this more than just a temporary setback. However, the magnitude of the effect remains quite stable over the two years, suggesting that the main negative impact happens in the first year, and it does not seem to compound and get larger as years go on. All columns include school fixed effects, teacher ratings of the student, and basic student-level controls.[16] Lagged (and double-lagged) test scores enter in 4th order polynomial form to allow flexibility in the functional relationship between previous academic performance and current (as well as what information a student's past performance provides to a teacher deciding whether to promote or retain the student). These results are robust to a variety of specification checks, including controlling for baseline parent ratings (see appendix section A.1.2).[17]

Students who are retained are also 7.1 percentage points more likely to drop out of school in the following year than their promoted peers. In columns 7 and 8, we replace the outcome of the baseline specification with an indicator equal to one if the student dropped out of school and zero otherwise.[18] In this specification, lagged score represents an average of the three subjects.

[16]We control for age because younger students may be less intellectually mature and more likely to be held back (Mahjoub [2017], Eide and Showalter [2001]) or older students may prioritize school less or are held back due to misbehavior. We control for gender because boys may be more disruptive in class, leading to retention for behavioral reasons, however their rowdiness may also distract them from their studies, biasing coefficients upwards. Note however, as shown in tables 4 and 5 we find no difference based on gender, either in terms of the probability of retention or in differential outcomes

[17]These results are robust to including parent ratings of the children (appendix table A.5 shows these results; we do not use those in the main text because the sample size shrinks substantially). We also test whether controlling for two lags of test scores is "enough" by adding time-invariant controls to the specification. As appendix tables A.6-A.8 show, coefficients remain very similar and the coefficient on the time-invariant control doesn't significantly explain student performance, which suggests that the two lags are adequately capturing relevant omitted variables.

[18]By "drop out," we mean whether the student drops out before year $t + 1$. The timeline is as follows. At the end of year t, teachers assign a promotion status to each student (either promoted to next grade or repeat current grade). Between years t and $t + 1$, each student makes a decision to stay in school or dropout (they can also leave the mauza or leave the sample for unspecified reasons, however we ignore those options here). Our dropout variable is equal to one if the child dropped out of school and equal to zero if they are still enrolled in school.

Table 3: The Effect of Repetition on Test Scores and Dropout

	Math		English		Urdu		Dropout	
	(1) 1 yr	(2) 2 yrs	(3) 1 yr	(4) 2 yrs	(5) 1 yr	(6) 2 yrs	(7) 1 yr	(8) 2 yrs
Not Promoted	**-0.440*** **	**-0.322*** **	**-0.278*** **	**-0.338*** **	**-0.371*** **	**-0.364*** **	**0.070*** **	**0.009**
	(0.055)	(0.076)	(0.036)	(0.050)	(0.043)	(0.056)	(0.014)	(0.019)
Lagged Score	**0.467*** **	**0.555*** **	**0.404*** **	**0.364*** **	**0.467*** **	**0.432*** **	**-0.005**	**-0.004**
	(0.027)	(0.037)	(0.023)	(0.033)	(0.024)	(0.028)	(0.007)	(0.011)
2yr Lagged Score	**0.247*** **	**0.240*** **	**0.195*** **	**0.142*** **	**0.226*** **	**0.174*** **	**-0.018** **	**-0.010**
	(0.027)	(0.035)	(0.020)	(0.032)	(0.027)	(0.030)	(0.008)	(0.010)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Rating	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag/Double Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	5529	3992	5529	3992	5529	3992	6222	4633
R^2	0.72	0.70	0.78	0.75	0.75	0.73	0.19	0.25

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. We run separate regressions for each school subject (math, English, and Urdu). Lagged score refers to a student's previous year score on the appropriate school subject (the lagged outcome variable).
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

The previous results all control for gender, yet it is informative to check separately how the effects may differ for girls and boys. First, table 4 shows that girls are not retained at different rates than boys.[19] This result holds in all specifications, and is a fairly precises zero. Table 5 then considers how the impacts of retention differ between girls and boys. The results suggest that girls fare better than boys after being retained, in that their test scores do not decrease by such a large amount. However, the results are noisier than the main effects presented in table 3 and not as large. For the remainder of the analysis, we will control for gender but we will not show the results separately.[20]

[19]The sample sizes are larger in this table than for the main result in table 3 and the heterogeneity by gender in table 5 because this table does not restrict to students with three years of test scores.

[20]We also consider heterogeneity in the effect of repetition between public and private schools. Appendix table A.10 shows that the effects of repetition in private schools are more muted, suggesting potential private school mitigation. This is an interesting finding, however the main specification includes school fixed effects so this does not bias our main results.

Table 4: The Probability of Repeating by Gender

	(1)	(2)	(3)
Female	**0.0064**	**-0.0095**	**0.0003**
	(0.0057)	(0.0107)	(0.0117)
Lagged Score			**-0.0762*****
			(0.0091)
2yr Lagged Score			**-0.0493*****
			(0.0106)
Constant	**0.1035*****	**0.0979*****	**0.0652*****
	(0.0038)	(0.0072)	(0.0117)
AgeControls	No	Yes	Yes
Teacher Rating	No	No	Yes
Fixed Effect	-	School	School
Lag Polynomial	-	-	4th
N	11869	11361	6113
R^2	0.00	0.18	0.34

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. The outcome variable is a dummy variable equal to one if the student repeats the grade. Lagged score refers to a student's previous year score on the appropriate school subject (the lagged outcome variable).
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table 5: The Effect of Repetition on Test Scores and Dropout - By Gender

	Math		English		Urdu		Dropout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted × Female		**0.1644**		**0.1695****		**0.1806****		**0.0036**
		(0.1062)		(0.0701)		(0.0845)		(0.0275)
Not Promoted	**-0.4403*****	**-0.5225*****	**-0.2783*****	**-0.3634*****	**-0.3706*****	**-0.4614*****	**0.0701*****	**0.0684*****
	(0.0553)	(0.0850)	(0.0364)	(0.0575)	(0.0434)	(0.0708)	(0.0145)	(0.0196)
Female	**-0.0341**	**-0.0391**	**0.0404***	**0.0353**	**0.0604****	**0.0550****	**-0.0036**	**-0.0037**
	(0.0279)	(0.0280)	(0.0218)	(0.0218)	(0.0256)	(0.0257)	(0.0080)	(0.0079)
Lagged Score	**0.4673*****	**0.4671*****	**0.4042*****	**0.4046*****	**0.4668*****	**0.4659*****	**-0.0046**	**-0.0046**
	(0.0270)	(0.0269)	(0.0231)	(0.0231)	(0.0238)	(0.0237)	(0.0074)	(0.0074)
2yr Lagged Score	**0.2473*****	**0.2457*****	**0.1951*****	**0.1946*****	**0.2263*****	**0.2256*****	**-0.0177****	**-0.0178****
	(0.0269)	(0.0270)	(0.0199)	(0.0198)	(0.0271)	(0.0271)	(0.0085)	(0.0084)
Constant	**0.7768*****	**0.7805*****	**0.6076*****	**0.6091*****	**0.6420*****	**0.6440*****	**0.0157****	**0.0158****
	(0.0314)	(0.0316)	(0.0261)	(0.0262)	(0.0255)	(0.0255)	(0.0080)	(0.0080)
AgeControls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Rating	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	5529	5529	5529	5529	5529	5529	6222	6222
R^2	0.72	0.72	0.78	0.78	0.75	0.75	0.19	0.19

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. We run separate regressions for each school subject (math, English, and Urdu). Lagged score refers to a student's previous year score on the appropriate school subject (the lagged outcome variable).
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

4.2 Mechanism - Teachers

Considering mechanisms by which grade repetition reduces test scores, we first consider the role of the teacher. To proxy for a teacher's opinion of a given student, we will use the teacher rating. For the contemporaneous teacher—the teacher who holds the student back in our sample—we find that they have a significantly lower rating for the child than we would predict given the student's lagged test scores and other information. This can be seen in column (1) of table 6. If we consider the student's teacher for the next year, we see that they also have a negative view of the student (column (2)), however, when controlling for the students test scores in that year, the fact that they were retained has no significant impact on the teacher rating. In fact, comparing the R^2 of columns (3) and (4), we learn that promotion status has no additional predictive value on the subsequent-year teacher's rating of the student above test scores.

These results tell us two things. First, teachers do not seem to label students as "repeaters" and have a negative bias towards them. If anything, it seems as if each year the teacher views the students as i.i.d draws conditional on their test scores. Second, if we were worried that the contemporaneous teacher retained the student because of something that was observable to them but unobservable to us, it seems as if these traits were also unobservable to the student's teacher the following year. This lends support to our theory that teachers are trained to teach, not to recognize which students should be retained, and that some of the retention that we see in the data is caused by random noise.

Table 6: The Effect of Repetition on Teacher Ratings

	(1) Contemporaneous	(2) Next Year	(3) Next Year	(4) Next Year
Not Promoted	**-0.8575*** **	**-0.3729*** **	**-0.0398**	
	(0.160)	(0.137)	(0.140)	
Lagged Average Score	**1.1570*** **	**1.1220*** **	**0.5880*** **	**0.5891*** **
	(0.095)	(0.088)	(0.093)	(0.093)
2yr lagged Average Score	**0.4842*** **	**0.5342*** **	**0.2781*** **	**0.2786*** **
	(0.106)	(0.099)	(0.093)	(0.093)
Age	**-0.0791*** **			
	(0.026)			
Average Score: Y3			**1.0268*** **	**1.0318*** **
			(0.086)	(0.084)
Constant	**6.0961*** **	**5.4368*** **	**4.7356*** **	**4.7292*** **
	(0.289)	(0.085)	(0.106)	(0.104)
Age/Gender Controls	Yes	Yes	Yes	Yes
Fixed Effect	School	School	School	School
Lag Polynomial	4th Order	4th Order	4th Order	4th Order
N	5529	5446	5446	5446
R^2	0.45	0.49	0.51	0.51

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. The first column refers to the teacher who retained the student, and the remaining three columns refer to the teacher who taught the student the following year. Lagged score refers to a student's previous year average test score. Column (4) illustrates that promotion does not increase the R-squared of the regression.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

4.3 Mechanism - Household

Next, our linked household-student-school panel data allow us to provide new insights into how parents' attitudes and investments in relation to their children change following grade repetition.

We find that parents treat grade repetition as a signal of lower intellectual ability or potential. Table 7 has three parental perceptions as the outcome variables. Parents decrease their ratings of their child's performance in school, work ethic, and intelligence by 0.15 to 0.32 standard deviations after their child is retained. This effect holds even after controlling for lagged parent ratings, demonstrating that there is a clear change in parent evaluations of a child after she has been retained. Appendix table A.13 shows that these results are robust to controlling for birth order/age.

Table 7: The Effect of Repetition on Parent Perceptions

	School Performance		Work Ethic		Inteligence	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted	**-0.1516****	**-0.1321****	**-0.3281****	**-0.3120***	**-0.2552****	**-0.2405***
	(0.062)	(0.063)	(0.158)	(0.160)	(0.129)	(0.124)
Lagged Average Score	**0.1811*****	**0.1668*****	**0.2553*****	**0.2476*****	**0.3758*****	**0.3485*****
	(0.039)	(0.038)	(0.090)	(0.089)	(0.079)	(0.078)
2yr lagged Average Score	**-0.0366**	**-0.0586**	**0.0309**	**0.0225**	**-0.1467**	**-0.1673**
	(0.050)	(0.049)	(0.112)	(0.112)	(0.112)	(0.110)
Lagged School Performance		**0.1532*****				
		(0.027)				
Lagged Hardworking				**0.0830**		
				(0.061)		
Lagged Inteligence						**0.2366*****
						(0.061)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	741	741	741	741	741	741
N	0.13	0.17	0.08	0.08	0.10	0.12

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Outcome questions are listed in full in the text. *** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table 8 examines perceptions two years after the retention, and finds mixed results depending on the question. For school performance, there is no significant effect, perhaps signalling that parents have revised downward their expectations and their child is no longer "disappointing" them. The impacts on work ethic are persistent and similar in size two years after the retention. The impacts on perceived intelligence are no longer significant, but in terms of magnitude are still half the size as the initial estimates. Together, this shows that having their child repeat a grade causes parents to significantly downgrade their perceptions of their child, and some of these shifts are long-lasting.

Table 8: The Effect of Repetition on Parent Perceptions - Two Years After Repeating

	School Performance		Work Ethic		Inteligence	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted	**-0.0467**	**-0.0330**	**-0.3141***	**-0.2730***	**-0.1355**	**-0.1180**
	(0.136)	(0.134)	(0.165)	(0.161)	(0.165)	(0.161)
Lagged Average Score	**0.2392*****	**0.2323*****	**-0.0381**	**-0.0560**	**0.0262**	**0.0056**
	(0.067)	(0.067)	(0.092)	(0.092)	(0.093)	(0.091)
2yr lagged Average Score	**-0.0453**	**-0.0562**	**0.1236**	**0.1057**	**0.0713**	**0.0568**
	(0.080)	(0.080)	(0.108)	(0.106)	(0.123)	(0.122)
Lagged School Performance		**0.0850***				
		(0.044)				
Lagged Hardworking				**0.2042*****		
				(0.062)		
Lagged Inteligence						**0.2039*****
						(0.062)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	710	710	710	710	710	710
R^2	0.09	0.09	0.03	0.05	0.05	0.06

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Outcome questions are listed in full in the text. *** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

While the parental perception questions are insightful, they are also somewhat vague. A more concrete survey questions specifically asked parents what score they believe that their child could achieve on a test (they were asked for a minimum, maximum, and average). Table 9 presents suggestive evidence that parents substantially decreased all three of these measures if their child was retained.[21] This, along with the previous two tables, offers evidence that parents significantly lower their expectations of their child once they are retained.

[21]The sample size shrinks in this table because the question was asked in a follow-up survey for which not all households could be found and matched.

Table 9: The Effect of Repetition on Parent Beliefs

	(1) Average	(2) Maximum	(3) Minimum
Not Promoted	**-4.9701***	**-5.9614***	**-3.9929**
	(2.761)	(3.084)	(2.550)
Lagged Average Score	**1.1569**	**1.6451**	**0.3857**
	(2.246)	(2.510)	(2.120)
2yr lagged Average Score	**3.0027**	**2.4747**	**2.8306**
	(2.658)	(2.820)	(2.484)
Constant	**50.2885*****	**61.3707*****	**39.6159*****
	(1.944)	(2.097)	(1.803)
Age/Gender Controls	Yes	Yes	Yes
Lag Polynomial	4th	4th	4th
N	513	513	514
R^2	0.43	0.43	0.40

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Outcome refers to how many questions parents expect a student to answer correctly on a test.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

In addition to revising their beliefs and expectations downwards, parents also reallocate resources away from repeating students. Table 10 shows annual expenditure on a repeater's education (which includes longer-term educational expenses including annual school fees, school uniforms, textbooks, and school supplies) decreases by approximately 176 PKR, or 20% of mean annual educational expenditure (column 4).[22] While there appears to be less of a decrease in monthly educational expenditures (which includes shorter-term expenses such as transportation costs, private tutoring, and pocket money for school), the point estimate is negative.[23] The magnitude of the decrease in expenditure gets even larger two years after the retention. Appendix table A.14 shows that results are robust to controlling for birth order.

[22]This drop in educational expenditure is not due to dropouts–all children in this regression's sample remained in school before and after grade repetition.

[23]We are lacking statistical power for this regression as it is identified off of only 56 repeaters in the smaller subsample of students whose household was included in the household survey.

Table 10: The Effect of Repetition on Parent Investments

	Monthly \$ $(t+1)$		Annual \$ $(t+1)$		Monthly \$ $(t+2)$		Annual \$ $((t+2))$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-17.7139**	**-15.0460**	**-176.1899*** **	**-111.4625***	**-47.2153** **	**-46.2378** **	**-209.9473***	**-199.9287***
	(14.908)	(18.070)	(59.553)	(57.743)	(22.513)	(21.416)	(115.382)	(114.009)
Lagged Average Score	**12.3195**	**-0.0491**	**-15.5160**	**-20.4853**	**13.0877**	**8.5559**	**122.1426**	**121.3734**
	(15.489)	(13.301)	(47.249)	(45.470)	(17.020)	(16.707)	(77.506)	(77.480)
2yr lagged Average Score	**33.6261**	**10.5395**	**102.3851***	**65.1619**	**20.5382**	**12.0793**	**32.8975**	**27.1361**
	(22.248)	(15.056)	(59.911)	(57.953)	(21.311)	(21.183)	(102.077)	(101.723)
Lagged Monthly \$		**0.8893*** **				**0.3258*** **		
		(0.149)				(0.091)		
Lagged Anual \$				**0.4742*** **				**0.0734**
				(0.055)				(0.085)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	741	741	741	741	741	741	741	741
R^2	0.05	0.33	0.05	0.16	0.04	0.06	0.05	0.05

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Full descriptions of short-term and long-term investments are in the text.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

To confirm that these effects represent a resource reallocation away from the repeating student specifically (rather than from some confounding household-level shock), we estimate the same regression with outcome variable average expenditure on all other siblings in the household. Table 11 shows that household expenditure on the siblings of repeaters may even rise, though the effect is not statistically distinguishable from zero (columns 1-4).[24] Correspondingly, parent perceptions of the repeater's siblings also rise or stay constant (table 11 columns 5-10). Table A.15 shows this is robust to controlling for birth order.

[24]The sample sizes in this table are smaller than in tables 9 and 10 due to students who have no siblings.

Table 11: Effect of repetition on investment in and perceptions of siblings

	Monthly \$ $(t+1)$		Annual \$ $(t+1)$		Performance (t+1)		Work Ethic (t+1)		Inteligence (t+1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Not Promoted	**0.5588**	**5.3413**	**29.2969**	**74.7752**	**0.1453**	**0.0234**	**0.2836***	**0.2352**	**0.0218**	**0.0073**
	(29.768)	(27.152)	(78.805)	(69.771)	(0.136)	(0.116)	(0.168)	(0.170)	(0.158)	(0.156)
Lagged Average Score	**26.8716**	**13.8832**	**92.1976**	**61.4781**	**0.1592***	**0.1126***	**0.1884****	**0.1850****	**0.0803**	**0.0649**
	(16.515)	(14.155)	(60.140)	(56.643)	(0.086)	(0.066)	(0.083)	(0.080)	(0.079)	(0.073)
2yr lagged Average Score	**19.2244**	**14.8686**	**18.2714**	**32.7411**	**-0.0202**	**-0.0223**	**-0.2928*****	**-0.2922*****	**-0.0812**	**-0.0163**
	(21.413)	(17.907)	(74.761)	(69.083)	(0.106)	(0.082)	(0.107)	(0.104)	(0.105)	(0.100)
Lagged Sibling Monthly \$		**1.9400*****								
		(0.267)								
Lagged Sibling Annual \$				**1.1220*****						
				(0.223)						
Lagged Sibling Overall						**0.4903*****				
						(0.036)				
Lagged Sibling Hardworking								**0.2441*****		
								(0.089)		
Lagged Sibling Inteligence										**0.3269*****
										(0.079)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	719	719	719	719	719	702	719	702	719	702
R^2	0.05	0.20	0.07	0.22	0.07	0.34	0.04	0.05	0.02	0.05

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. The table shows results from a student-level regression of average household expenditure on all siblings (leaving out the student) on whether the student repeated or not (columns 1-4) and of average parent evaluation of siblings (leaving out the student) on whether the student repeated or not. Lagged score refers to a student's previous year average test score. Full descriptions of short-term and long-term investments are in the text. *** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

The previous evidence indicates that parents place the blame of retention on their child, and the following tables suggest that they do not put any on their child's teachers. In the parent surveys, parents were asked questions about how they view their child's teachers. Three of these questions were "How regular is your child's class-teacher overall?", "How good would you say that your child's class-teacher is in his/her teaching skills?", and "How good would you say that your child's class-teacher is overall?". Parents were allowed to respond that they "do not know", which is why the sample size decreases from the previous tables. Table 12 shows that parents of repeating students do not hold worse views of their child's teacher than parents of promoted students (conditional on expressing a view of their child's teacher). All coefficients are fairly precisely estimated zeros. Table 13 shows the same analysis, but for teachers the following year (the year during which the student is repeating the grade). The sample size is small, but if anything parents have a more positive view of next-year teachers if their child is repeating. The opposite of blaming a school and teachers for their child's retention, it seems as if parents are thankful that teachers are still willing to put effort towards their child after such a negative signal.

Table 12: The Effect of Repetition on Parental Ratings of Teachers (current year)

	How Regular		Good Teacher		Overall	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted	**-0.0538**	**0.0158**	**0.0130**	**0.0381**	**-0.0408**	**-0.0548**
	(0.054)	(0.062)	(0.091)	(0.108)	(0.092)	(0.109)
Lagged Average Score		**0.0062**		**-0.0268**		**-0.0785**
		(0.033)		(0.059)		(0.060)
2yr lagged Average Score		**0.0773***		**0.1743****		**0.0616**
		(0.045)		(0.075)		(0.074)
Age/Gender Controls	No	Yes	No	Yes	No	Yes
Lag Polynomial	-	4th	-	4th	-	4th
N	700	631	699	630	686	620
R^2	0.00	0.07	0.00	0.04	0.00	0.04

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table 13: The Effect of Repetition on Parental Ratings of Teachers (next year)

	How Regular		Good Teacher		Overall	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted	**0.1783**	**0.1777**	**0.2072**	**0.2642**	**0.3333****	**0.3270***
	(0.119)	(0.136)	(0.150)	(0.167)	(0.157)	(0.174)
Lagged Average Score		**0.1151**		**0.2245****		**0.2931****
		(0.088)		(0.108)		(0.113)
2yr lagged Average Score		**0.0209**		**0.0967**		**-0.0864**
		(0.113)		(0.138)		(0.146)
Age/Gender Controls	No	Yes	No	Yes	No	Yes
Lag Polynomial	-	4th	-	4th	-	4th
N	375	332	375	331	369	325
R^2	0.01	0.06	0.01	0.09	0.01	0.08

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

4.4 Mechanisms - Student

Being retained significantly lowers a child's test scores the next year, however, it is possible that there are other impacts as well, which could be harder to measure. Especially if retention has a random component (after controlling for test scores), it is possible that children would feel quite discouraged if they feel like they have been retained at random. To analyze this further, we consider several survey questions as outcomes in the same specification. Table 14, column (1) shows that retention has a large negative impact on the student

agreeing with the statement "If I study hard at school I will be rewarded with a better job in the future."[25] The sample size is smaller because this question was asked in a follow-up survey for which not every student as found. We take this as evidence that retention discourages students, and makes them doubt their own agency in their studies.

Columns (2) through (5) show that retention seems to have a specific impact on a student's relationship with school, and does not discourage them in other areas of their life. There are no significant impacts of retention on any of the four other questions. The questions for each column are the following: (2) "If I try hard, I can improve my situation in life"; (3) "Other people in my family make all the decisions about how I spend my time"; (4) "I like to make plans for my future studies and work"; (5) "I have no choice about the work I do - I must work". This specific decrease in belief that effort in school will benefit their life goals could be one reason why students who are retained see a decrease in their test scores.

Table 14: The Effect of Repetition on Student Beliefs

	(1) Study Hard	(2) Try Hard	(3) Others Decisions	(4) Make Plans	(5) Must Work
Not Promoted	**-0.2602****	**-0.0375**	**0.0110**	**-0.0584**	**0.0772**
	(0.102)	(0.032)	(0.084)	(0.087)	(0.091)
Lagged Average Score	**0.0210**	**-0.0226**	**-0.0378**	**0.0323**	**-0.0243**
	(0.046)	(0.015)	(0.061)	(0.043)	(0.057)
2yr lagged Average Score	**-0.0017**	**0.0277**	**-0.0312**	**0.0200**	**-0.0800**
	(0.062)	(0.021)	(0.078)	(0.056)	(0.068)
Constant	**0.8329*****	**0.9918*****	**0.5317*****	**0.8900*****	**0.1568*****
	(0.052)	(0.006)	(0.054)	(0.042)	(0.051)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order
N	578	578	578	578	578
R^2	0.35	0.27	0.30	0.32	0.20

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Outcome questions are listed in full in the text. *** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

5 Conclusion

This paper provides new evidence on the effects of grade repetition on academic outcomes and household reactions. Using rich panel data, we provide a collage of evidence that sheds light on the dynamics of *how* grade repetition reduces future learning. The teacher who chooses to repeat the student may be making a

[25] To obtain the outcome variable for this regression, we recode a Likert scale response into a binary for agree/disagree.

quasi-random judgement error, as their teacher next year shows no negative bias towards repeaters conditional on test scores. Students who are retained shows signs of being discouraged, specifically downgrading their beliefs in the usefulness of putting effort towards their studies. Finally, parents play arguably the largest role, with significant and persistent decreases in the beliefs, expectations, and investments towards their children who repeat a grade. In terms of outcomes, children who are repeating score lower in tests the next year and have an increased probability of dropping out. Together, this evidence shows that retaining students has long-running repercussions, many of which may begin at home.

References

Allende, Claudia, Francisco Gallego, and Christopher Neilson. 2019. "Approximating the equilibrium effects of informed school choice."

Allensworth, Elaine M. 2005. "Dropout rates after high-stakes testing in elementary school: A study of the contradictory effects of Chicago's efforts to end social promotion." *Educational Evaluation and Policy Analysis*, 27(4): 341–364.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2017. "Report cards: The impact of providing school and child test scores on educational markets." *American Economic Review*, 107(6): 1535–1563.

Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, and Tristan Zajonc. 2011. "Do value-added estimates add value? Accounting for learning dynamics." *American Economic Journal: Applied Economics*, 3(3): 29–54.

Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, Tara Vishwanath, and Tristan Zajonc. 2007. "Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to inform the education policy debate." *World Bank, Washington, DC*.

Banerji, Rukmini. 2000. "Poverty and primary schooling: field studies from Mumbai and Delhi." *Economic and Political Weekly*, 795–802.

Barrera-Osorio, Felipe, Kathryn Gonzalez, Francisco Lagos, and David J Deming. 2020. "Providing performance information in education: An experimental evaluation in Colombia." *Journal of Public Economics*, 186: 104185.

Becker, Gary S, and Nigel Tomes. 1976. "Child endowments and the quantity and quality of children." *Journal of political Economy*, 84(4, Part 2): S143–S162.

Behrman, Jere R, Mark R Rosenzweig, and Paul Taubman. 1994. "Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment." *Journal of political economy*, 102(6): 1131–1174.

Behrman, Jere R, Robert A Pollak, and Paul Taubman. 1986. "Do parents favor boys?" *International Economic Review*, 33–54.

Bergman, Peter. 2021. "Parent-child information frictions and human capital investment: Evidence from a field experiment." *Journal of political economy*, 129(1): 286–322.

Borghesan, Emilio, Hugo Reis, and Petra E Todd. 2022. "Learning through repetition? A dynamic evaluation of grade retention in Portugal." *PIER Working Paper*, , (22-030): 1–80.

Brophy, Jere. 2006. *Grade repetition.* International Institute for Educational Planning.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American economic review*, 104(9): 2593–2632.

Chevalier, Arnaud, Steve Gibbons, Andy Thorpe, Martin Snell, and Sherria Hoskins. 2009. "Students' academic self-perception." *Economics of Education Review*, 28(6): 716–727.

Chohan, Bushra I, and Samina A Qadir. 2011. "Automatic Promotion Policy at Primary Level and MDG-2." *Journal of Research & Reflections in Education (JRRE)*, 5(1).

Das, Jishnu, Priyanka Pandey, and Tristan Zajonc. 2012. "Learning levels and gaps in Pakistan: A comparison with Uttar Pradesh and Madhya Pradesh." *Economic and Political Weekly*, 228–240.

Dizon-Ross, Rebecca. 2019. "Parents' beliefs about their children's academic ability: Implications for educational investments." *American Economic Review*, 109(8): 2728–2765.

Dobronyi, Christopher R, Philip Oreopoulos, and Uros Petronijevic. 2019. "Goal setting, academic reminders, and college success: A large-scale field experiment." *Journal of Research on Educational Effectiveness*, 12(1): 38–66.

Eide, Eric R, and Mark H Showalter. 2001. "The effect of grade retention on educational and labor market outcomes." *Economics of Education review*, 20(6): 563–576.

Eisemon, Thomas Owen. 1997. "Reducing Repetition: Issues and Strategies."

Eren, Ozkan, Briggs Depew, and Stephen Barnes. 2017. "Test-based promotion policies, dropping out, and juvenile crime." *Journal of Public Economics*, 153: 9–31.

Eren, Ozkan, Michael F Lovenheim, and Naci H Mocan. 2018. "The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy." National Bureau of Economic Research.

Figlio, David N, and Umut Özek. 2019. "An Extra Year to Learn English? Early Grade Retention and the Human Capital Development of English Learners." National Bureau of Economic Research.

Gan, Tianqi. 2021. "How Parental Educational Investments Respond to Changes in Ability Belief: An Application of Big Data Techniques in Education." *Working Paper*.

Gomes-Neto, Joao Batista, and Eric A Hanushek. 1994. "Causes and consequences of grade repetition: Evidence from Brazil." *Economic Development and Cultural Change*, 43(1): 117–148.

Griliches, Zvi. 1979. "Sibling models and data in economics: Beginnings of a survey." *Journal of political Economy*, 87(5, Part 2): S37–S64.

Hastings, Justine S, and Jeffrey M Weinstein. 2008. "Information, school choice, and academic achievement: Evidence from two experiments." *The Quarterly journal of economics*, 123(4): 1373–1414.

Jacob, Brian A. 2005. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of public Economics*, 89(5): 761–796.

Jacob, Brian A, and Lars Lefgren. 2004. "Remedial education and student achievement: A regression-discontinuity analysis." *Review of economics and statistics*, 86(1): 226–244.

Jacob, Brian A, and Lars Lefgren. 2009. "The effect of grade retention on high school completion." *American Economic Journal: Applied Economics*, 1(3): 33–58.

Kinsler, Josh, and Ronni Pavan. 2021. "Local distortions in parental beliefs over child skill." *Journal of Political Economy*, 129(1): 81–100.

Koppensteiner, Martin Foureaux. 2014. "Automatic grade promotion and student performance: Evidence from Brazil." *Journal of Development Economics*, 107: 277–290.

Mahjoub, Mohamed-Badrane. 2017. "The treatment effect of grade repetitions." *Education Economics*, 25(4): 418–432.

Manacorda, Marco. 2012. "The cost of grade retention." *Review of Economics and Statistics*, 94(2): 596–606.

Mariano, Louis T, and Paco Martorell. 2013. "The academic effects of summer instruction and retention in New York City." *Educational Evaluation and Policy Analysis*, 35(1): 96–117.

Oreopoulos, Philip, and Uros Petronijevic. 2019. "The remarkable unresponsiveness of college students to nudging and what we can learn from it." National Bureau of Economic Research.

Pitt, Mark M, Mark R Rosenzweig, and Md Nazmul Hassan. 1990. "Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries." *The American Economic Review*, 1139–1156.

Pritchett, Lant, and Amanda Beatty. 2015. "Slow down, you're going too fast: Matching curricula to student skill levels." *International Journal of Educational Development*, 40: 276–288.

Rosenzweig, Mark R, and Junsen Zhang. 2009. "Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy." *The Review of Economic Studies*, 76(3): 1149–1174.

Schwerdt, Guido, Martin R West, and Marcus A Winters. 2017. "The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida." *Journal of Public Economics*, 152: 154–169.

Stinebrickner, Todd, and Ralph Stinebrickner. 2012. "Learning about academic ability and the college dropout decision." *Journal of Labor Economics*, 30(4): 707–748.

Todd, Petra E, and Kenneth I Wolpin. 2003. "On the specification and estimation of the production function for cognitive achievement." *The Economic Journal*, 113(485).

Tomes, Nigel. 1981. "The family, inheritance, and the intergenerational transmission of inequality." *Journal of Political Economy*, 89(5): 928–958.

UNESCO. 2012. "Opportunities lost: The impact of grade repetition and early school leaving."

Walton, Gregory M, and Geoffrey L Cohen. 2011. "A brief social-belonging intervention improves academic and health outcomes of minority students." *Science*, 331(6023): 1447–1451.

Yeager, David S, Gregory M Walton, Shannon T Brady, Ezgi N Akcinar, David Paunesku, Laura Keane, Donald Kamentz, Gretchen Ritter, Angela Lee Duckworth, Robert Urstein, et al. 2016. "Teaching a lay theory before college narrows achievement gaps at scale." *Proceedings of the National Academy of Sciences*, 113(24): E3341–E3348.

A.1 Appendix

A.1.1 Additional descriptive tables and figures

Table A.1: Average test scores and teacher ratings for retained and promoted students

Table A.2: Average Test Scores, Year 3 (Post-)

	Promoted	Not Promoted	Difference
English score (Pre-)	.3	-.563	-.864***
	(.921)	(1.092)	
Math score (Pre-)	.26	-.961	-1.221***
	(1.092)	(1.392)	
Urdu score (Pre-)	.367	-.637	-1.005***
	(.946)	(1.156)	
N	10463	1249	11712

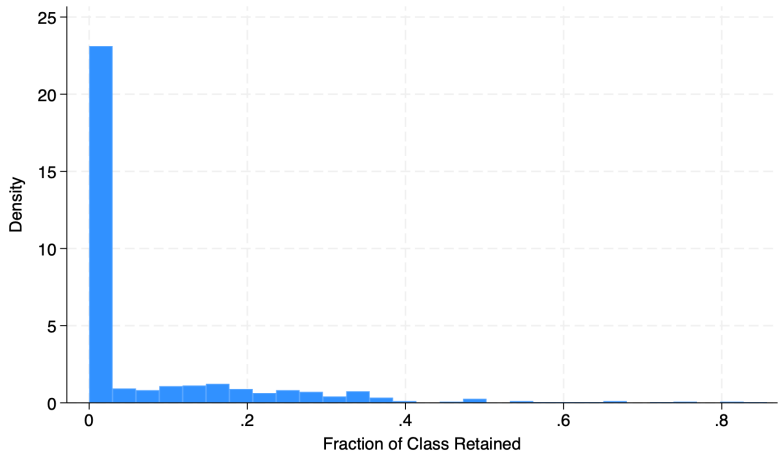
Note:
*** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level

Table A.3: Average Test Scores, Year 3 (Post-)

	Promoted	Not Promoted	Difference
Rating by teacher	6.148	4.296	-1.852***
	(2.27)	(2.421)	
N	6458	594	7052

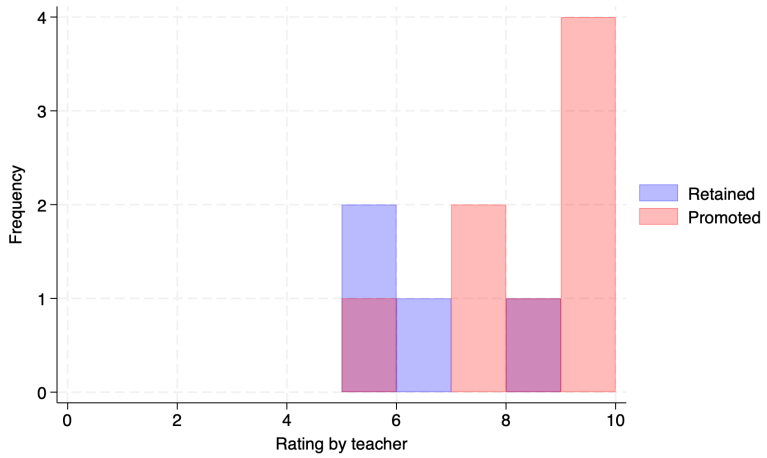
Note:
*** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level

Figure A.1: Across-teacher distribution of the fraction of class retained



Note: Histogram of the fraction of each teacher's classroom who is retained (including classrooms without any retained students).

Figure A.2: Example teacher: frequency histogram of number of promoted and retained students with each rating



Note: Plot shows the number of retained and promoted students (y-axis) with different teacher ratings (shown on x-axis)

Table A.4: Regression Predicting Repetition

	(1)	(2)	(3)	(4)	(5)	(6)	(7) int	(8) int
Lagged Average Score	**-0.1099*** (0.0079)**	**-0.1037*** (0.0094)**	**-2.3451*** (0.4743)**	**-0.0872*** (0.0095)**	**-0.1226*** (0.0324)**	**-0.1304*** (0.0337)**	**-0.0389 (0.1111)**	**-0.0085 (0.1931)**
2yr lagged Average Score	**-0.0463*** (0.0095)**	**-0.0444*** (0.0111)**	**0.4857 (0.5080)**	**-0.0372*** (0.0110)**	**-0.0328 (0.0303)**	**-0.0199 (0.0450)**	**-0.1719 (0.1433)**	**-0.2791 (0.2704)**
Mom Educated		**0.0010 (0.0078)**	**0.0117 (0.1393)**		**0.0011 (0.0076)**			**-0.0771 (0.0594)**
Child Wealth		**-0.0038 (0.0022)**	**-0.0458 (0.0389)**					**-0.0036 (0.0214)**
Height		**-0.0020 (0.0046)**	**-0.0786 (0.0778)**					**-0.0001 (0.0396)**
Weight		**-0.0002 (0.0052)**	**-0.0136 (0.0875)**					**0.0071 (0.0444)**
Child Talent								
$= 1$				**0.1831*** (0.0206)**	**0.1756*** (0.0208)**			**0.1585 (0.1536)**
$= 2$				**0.1180*** (0.0186)**	**0.1114*** (0.0190)**			**-0.1605 (0.1339)**
$= 3$				**0.0713*** (0.0161)**	**0.0687*** (0.0162)**			**-0.1470 (0.1230)**
$= 4$				**0.0231 (0.0131)**	**0.0224 (0.0133)**			**0.1044 (0.1299)**
$= 6$				**-0.0208 (0.0117)**	**-0.0253** (0.0125)**			**-0.1889 (0.1063)**
$= 7$				**-0.0095 (0.0123)**	**-0.0216 (0.0135)**			**-0.1121 (0.1191)**
$= 8$				**-0.0070 (0.0126)**	**0.0052 (0.0155)**			**-0.0778 (0.1150)**
$= 9$				**0.0048 (0.0142)**	**-0.0022 (0.0184)**			**-0.0678 (0.1257)**
$= 10$				**-0.0019 (0.0162)**	**0.0314 (0.0227)**			**0.1208 (0.2219)**
Hardworking								
Below Average						**0.1650*** (0.0629)**	**0.1945*** (0.0656)**	**0.3568*** (0.1211)**
Above Average						**-0.0330 (0.0470)**	**-0.0600 (0.0521)**	**-0.0636 (0.0949)**
Intelligent								
Below Average						**-0.2381*** (0.0746)**	**-0.2619*** (0.0790)**	**-0.6469** (0.2744)**
Above Average						**0.0080 (0.0457)**	**0.0331 (0.0508)**	**-0.0598 (0.1267)**
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial 4th Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Test Score Components	No	No	Yes	No	Yes	No	Yes	Yes
Interact Scores, Child Chars	No	No	Yes	No	No	No	No	Yes
Interact Scores, TeachRatings	No	No	No	No	Yes	No	No	Yes
Interact Scores, ParRatings	No	No	No	No	No	No	Yes	Yes
Interact Par Ratings, TeachRatings	No	No	No	No	No	No	No	Yes
N	8071	5442	5434	5529	5529	698	698	487
Pseudo R^2	0.35	0.37		0.38	0.39	0.66	0.69	0.85

Note: This table shows coefficients from the regressions used to predict repetition used for computing the R^2's shown in table 2.

A.1.2 Robustness of main results

A.1.2.1 Controlling for parent ratings

We add baseline parent ratings separately to the specification from teacher ratings since only very small subsample received ratings from *both* their teachers and their parents. We have two separate ratings from parents: one of a child's intelligence and one of a child's work ethic. Since the two are highly correlated, we introduce them to the model one at a time.

While inclusion of parent rating in the regression substantially reduces power,[26] the effect of repetition remains large and significant. Table A.5 (columns 2, 3, 5, 6, 8, and 9) shows that, accounting for lagged parental ratings, grade repetition reduces test scores by between .29 and .44 standard deviations for the three subjects.[27] Comparing the second and third columns for each subject to the first, we see that the repetition coefficient decreases by at most only 0.024 standard deviations when including the parent ratings, suggesting that our baseline model isn't missing any important dimension of latent student ability. Additionally, none of the lagged parent ratings significantly explain test score variation, controlling for lagged scores.[28]

[26]These variables come from a household survey which was administered only to ten randomly selected households per village. The sample is all tested students matched with a surveyed household. Due to the reduced sample and lack of intra-school variation, we switch from school- to mauza-level fixed effects.

[27]Note that due to the smaller sample size, we switch from school- to mauza-level fixed effects here– otherwise, identification comes from only 36 repeaters and 58 non-repeaters in classrooms for which we observe both repeaters and non-repeaters. However, even when we include the school fixed effect in this specification, point estimates remain negative, though with larger standard errors; see appendix table **??**

[28]Columns 10-12 show the specification with dropping out as the outcome variable. While we lack sufficient power to claim an effect, it is reassuring that the coefficient barely changes between column 16 (the baseline specification on the smaller subsample) and columns 14 and 15 (the specification including parent ratings). Hence, the lack of significance seems to be due to small sample size rather than because the effect disappears when controlling for parent ratings.

	Math			English			Urdu			Dropout		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Not Promoted	**-0.433****	**-0.409****	**-0.436****	**-0.305****	**-0.291****	**-0.311****	**-0.393*****	**-0.369*****	**-0.389*****	**0.045**	**0.045**	**0.044**
	(0.179)	(0.181)	(0.179)	(0.143)	(0.143)	(0.142)	(0.143)	(0.137)	(0.141)	(0.036)	(0.036)	(0.036)
Lagged Score	**0.554*****	**0.549*****	**0.538*****	**0.481*****	**0.472*****	**0.484*****	**0.491*****	**0.478*****	**0.483*****	**0.001**	**0.001**	**0.002**
	(0.058)	(0.059)	(0.058)	(0.047)	(0.049)	(0.048)	(0.057)	(0.056)	(0.057)	(0.009)	(0.009)	(0.009)
2yr Lagged Score	**0.321*****	**0.317*****	**0.324*****	**0.317*****	**0.322*****	**0.317*****	**0.199*****	**0.196*****	**0.197*****	**-0.007**	**-0.007**	**-0.007**
	(0.067)	(0.067)	(0.065)	(0.053)	(0.053)	(0.052)	(0.064)	(0.062)	(0.063)	(0.007)	(0.006)	(0.007)
Hardworking (Lag)												
Below Average		**-0.196**			**-0.130**			**-0.146**			**-0.007**	
		(0.140)			(0.104)			(0.109)			(0.006)	
Above Average		**0.058**			**0.027**			**0.102****			**-0.008**	
		(0.061)			(0.055)			(0.050)			(0.006)	
Intelligent (Lag)												
Below Average			**-0.303**			**-0.112**			**-0.053**			**-0.001**
			(0.173)			(0.134)			(0.127)			(0.005)
Above Average			**0.054**			**-0.049**			**0.069**			**-0.009**
			(0.058)			(0.049)			(0.047)			(0.006)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mauza Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	698	698	698	698	698	698	698	698	698	760	760	760
R^2	0.68	0.69	0.69	0.73	0.73	0.73	0.74	0.74	0.74	0.25	0.26	0.26

Table A.5: Effect of Repetition on Test Scores – Considering Characteristics Observable to Parents

A.1.2.2 Specification checks on controlling for two lags of test scores

The purpose of the value-added model is to "difference out" time-invariant factors. One way to check whether it succeeds is by testing if results are sensitive to inclusion of covariates that (1) should bias results in the absence of the value-added model and (2) we expect to remain constant year to year. If coefficients remain largely unchanged whether we include the variable or not and the time-invariant covariate doesn't significantly explain student performance, it should be taken as evidence that the value-added model is effectively accounting for time-invariant factors. We consider three different time-invariant confounding factors at the household-, child-, and parent-level. Two of these (household wealth and child health) may be subject to small shocks year to year. In each case, absent the value-added specification, lack of inclusion would bias results. However, using the value-added model, results remain largely unchanged, indicating that the model is successful.

The first test variable is household wealth. It affects both achievement and repetition since wealthier households may spend more resources to help their children, including pressuring schools to promote their children to the next grade. Household wealth should also be relatively time-invariant (affecting past test scores as much as present test scores) so it should have little predictive power using the value-added model. We compute a principal components child wealth index based on 20 survey questions asking whether the child's household owned a variety of products.[29] Table A.6, column 2 shows results from our baseline specification

[29]These include beds, radio, television, refrigerator, bicycle, plough, small agricultural tools, tables, fans,

plus the wealth index. The repetition coefficients remains similar in magnitude (differing by at most .002 standard deviations). Furthermore, as expected, the wealth index is a poor predictor of test scores, conditional on lagged test scores.[30]

	Math		English		Urdu		Dropout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-0.468*** **	**-0.468*** **	**-0.313*** **	**-0.313*** **	**-0.400*** **	**-0.402*** **	**0.074*** **	**0.074*** **
	(0.054)	(0.054)	(0.037)	(0.037)	(0.042)	(0.041)	(0.015)	(0.015)
Lagged Score	**0.501*** **	**0.501*** **	**0.440*** **	**0.440*** **	**0.493*** **	**0.492*** **	**-0.008**	**-0.008**
	(0.026)	(0.027)	(0.023)	(0.023)	(0.024)	(0.024)	(0.007)	(0.007)
2yr Lagged Score	**0.262*** **	**0.262*** **	**0.213*** **	**0.213*** **	**0.249*** **	**0.251*** **	**-0.017** **	**-0.017** **
	(0.027)	(0.027)	(0.020)	(0.020)	(0.027)	(0.027)	(0.008)	(0.008)
Child Wealth		**-0.002**		**-0.001**		**-0.009** **		**-0.000**
		(0.005)		(0.004)		(0.004)		(0.002)
Constant	**0.775*** **	**0.774*** **	**0.641*** **	**0.641*** **	**0.661*** **	**0.659*** **	**0.019*** **	**0.019*** **
	(0.026)	(0.026)	(0.021)	(0.021)	(0.021)	(0.021)	(0.007)	(0.007)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	5543	5543	5543	5543	5543	5543	6239	6239
R^2	0.72	0.72	0.78	0.78	0.75	0.75	0.19	0.19

Table A.6: Robustness–Checking the Value-Added Model Differences out Household wealth

Second, child health affects both achievement and promotion decisions (malnourished children may be held back due to perceived immaturity and also perform worse due to lower energy/cognitive capacity), however it also should be relatively time-invariant (aside from short-term shocks). Again, table A.7 shows that the value-added model is successful: after inclusion of the health variables, coefficients change by at most 0.015 standard deviations and become *more negative* for all three subjects. Additionally, height/weight explain little of the variation in test scores.

tractor, cattle, goats, chicken, watches, motor rickshaw, motorcycle/scooter, car/taxi/van/pickup, telephone, and tubewell

[30]The exception is for English which could be because wealthier households are more likely to contain parents who speak English. Nonetheless, the limited change in the repetition coefficient indicates that even if wealth is related to English ability, it is mostly unrelated to repetition.

	Math		English		Urdu		Dropout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-0.463*** **	**-0.478*** **	**-0.309*** **	**-0.313*** **	**-0.389*** **	**-0.396*** **	**0.075*** **	**0.078*** **
	(0.055)	(0.056)	(0.037)	(0.037)	(0.042)	(0.042)	(0.015)	(0.015)
Lagged Score	**0.503*** **	**0.500*** **	**0.440*** **	**0.441*** **	**0.496*** **	**0.497*** **	**-0.007**	**-0.005**
	(0.026)	(0.027)	(0.023)	(0.023)	(0.024)	(0.024)	(0.007)	(0.007)
2yr Lagged Score	**0.261*** **	**0.266*** **	**0.213*** **	**0.212*** **	**0.249*** **	**0.250*** **	**-0.017** **	**-0.018** **
	(0.027)	(0.027)	(0.020)	(0.020)	(0.027)	(0.027)	(0.008)	(0.008)
Height		**-0.015**		**-0.012**		**-0.003**		**0.005**
		(0.012)		(0.009)		(0.009)		(0.003)
Weight		**-0.001**		**0.005**		**-0.015**		**0.008**
		(0.013)		(0.010)		(0.010)		(0.004)
Constant	**0.773*** **	**0.770*** **	**0.641*** **	**0.642*** **	**0.658*** **	**0.651*** **	**0.019*** **	**0.025*** **
	(0.026)	(0.027)	(0.020)	(0.021)	(0.021)	(0.022)	(0.007)	(0.007)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	5493	5442	5493	5442	5493	5442	6178	6123
R^2	0.72	0.72	0.78	0.78	0.75	0.75	0.19	0.20

Table A.7: Robustness–Checking the Value-Added Model Differences out Child Health

Finally, a mother's education level—which is also time-invariant—may influence both promotion decisions (through pressure on the school) and children's learning. Table A.8 shows that, as expected, the model successfully differences this effect out. The repetition coefficient is identical up to three decimal places and an educated mother doesn't significantly predict test scores once we hold lagged scores constant. For the dropout outcome specification, the repetition coefficient also differs only in the third decimal place.[31]

	Math		English		Urdu		Dropout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-0.480*** **	**-0.480*** **	**-0.374*** **	**-0.374*** **	**-0.409*** **	**-0.409*** **	**0.055*** **	**0.054*** **
	(0.047)	(0.047)	(0.032)	(0.032)	(0.032)	(0.032)	(0.011)	(0.011)
Lagged Score	**0.544*** **	**0.544*** **	**0.455*** **	**0.454*** **	**0.532*** **	**0.532*** **	**-0.013** **	**-0.011**
	(0.025)	(0.025)	(0.019)	(0.019)	(0.020)	(0.020)	(0.006)	(0.006)
2yr Lagged Score	**0.226*** **	**0.226*** **	**0.211*** **	**0.210*** **	**0.228*** **	**0.228*** **	**-0.017** **	**-0.016**
	(0.023)	(0.023)	(0.016)	(0.016)	(0.022)	(0.022)	(0.008)	(0.008)
Mom Educated		**0.009**		**0.023**		**0.010**		**-0.037*** **
		(0.016)		(0.014)		(0.014)		(0.004)
Constant	**0.746*** **	**0.743*** **	**0.615*** **	**0.609*** **	**0.639*** **	**0.636*** **	**0.023*** **	**0.031*** **
	(0.022)	(0.023)	(0.017)	(0.017)	(0.017)	(0.017)	(0.006)	(0.006)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	8071	8071	8071	8071	8071	8071	9132	9132
R^2	0.71	0.71	0.77	0.77	0.74	0.74	0.14	0.14

Table A.8: Robustness–Checking the Value-Added Model Differences out Mother's Education

[31] Interestingly, no children of educated mothers dropped out of school, indicating that beyond simply improving her children's academic performance, an educated mother is also highly unlikely to allow her child to drop out. However, given that there appears to be no relationship between mother's education and repetition (after accounting for her children's academic performance), this does not appear to bias results in a meaningful way.

We also show results are identical when including teacher rating as a categorical rather than continuous variable (only showed it as continuous in main table for convenience of exposition)

Table A.9: Main specification with factorized teacher rating

	Math		English		Urdu		Dropout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-0.472*****	**-0.440*****	**-0.313*****	**-0.278*****	**-0.402*****	**-0.371*****	**0.073*****	**0.070*****
	(0.055)	(0.055)	(0.037)	(0.036)	(0.043)	(0.043)	(0.014)	(0.014)
Lagged Score	**0.502*****	**0.467*****	**0.439*****	**0.404*****	**0.496*****	**0.467*****	**-0.007**	**-0.005**
	(0.027)	(0.027)	(0.023)	(0.023)	(0.024)	(0.024)	(0.007)	(0.007)
2yr Lagged Score	**0.258*****	**0.247*****	**0.213*****	**0.195*****	**0.245*****	**0.226*****	**-0.018****	**-0.018****
	(0.027)	(0.027)	(0.020)	(0.020)	(0.027)	(0.027)	(0.009)	(0.008)
Child Talent								
$= 1$		**-0.122**		**-0.060**		**-0.114****		**0.027**
		(0.066)		(0.050)		(0.056)		(0.019)
$= 2$		**-0.096**		**-0.098****		**-0.066**		**-0.001**
		(0.054)		(0.040)		(0.043)		(0.013)
$= 3$		**-0.120*****		**-0.052**		**-0.094****		**0.021**
		(0.045)		(0.036)		(0.037)		(0.014)
$= 4$		**-0.071**		**-0.019**		**-0.036**		**0.003**
		(0.039)		(0.028)		(0.031)		(0.009)
$= 6$		**0.026**		**0.052****		**0.052****		**0.004**
		(0.030)		(0.024)		(0.024)		(0.007)
$= 7$		**0.033**		**0.081*****		**0.049****		**0.001**
		(0.030)		(0.025)		(0.024)		(0.007)
$= 8$		**0.093*****		**0.133*****		**0.102*****		**0.000**
		(0.033)		(0.025)		(0.026)		(0.008)
$= 9$		**0.123*****		**0.166*****		**0.150*****		**-0.007**
		(0.038)		(0.030)		(0.029)		(0.009)
$= 10$		**0.176*****		**0.234*****		**0.166*****		**-0.009**
		(0.040)		(0.032)		(0.031)		(0.010)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	5529	5529	5529	5529	5529	5529	6222	6222
R^2	0.72	0.72	0.78	0.78	0.75	0.75	0.19	0.19

A.1.3 Additional results tables

Table A.10: The Differential Effects of Repetition in Public and Private Schools

	Math		English		Urdu	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted × Govt	**-0.5367*** **	**-0.4861*** **	**-0.1573** **	**-0.1009**	**-0.4054*** **	**-0.3661*** **
	(0.136)	(0.141)	(0.077)	(0.075)	(0.105)	(0.105)
Not Promoted	**0.0405**	**0.0151**	**-0.1635** **	**-0.1962*** **	**-0.0434**	**-0.0573**
	(0.115)	(0.117)	(0.074)	(0.069)	(0.100)	(0.099)
Govt	**-0.1445*** **	**-0.1393*** **	**-0.1722*** **	**-0.1790*** **	**-0.1354*** **	**-0.1314*** **
	(0.032)	(0.032)	(0.030)	(0.030)	(0.032)	(0.032)
Lagged Score	**0.5403*** **	**0.5201*** **	**0.4962*** **	**0.4726*** **	**0.5122*** **	**0.4915*** **
	(0.025)	(0.027)	(0.025)	(0.026)	(0.023)	(0.024)
2yr Lagged Score	**0.2465*** **	**0.2407*** **	**0.2201*** **	**0.2124*** **	**0.2441*** **	**0.2319*** **
	(0.025)	(0.025)	(0.023)	(0.023)	(0.024)	(0.024)
PCA Wealth Index	**0.0170** **	**0.0168** **	**0.0158** **	**0.0158** **	**0.0111**	**0.0103**
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Lagged PCA Wealth Index	**-0.0102**	**-0.0099**	**-0.0045**	**-0.0044**	**-0.0104**	**-0.0099**
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Teacher Rating						
$= 1$		**-0.0782**		**-0.0251**		**-0.0792**
		(0.074)		(0.046)		(0.051)
$= 2$		**-0.0752**		**-0.0678**		**-0.0411**
		(0.058)		(0.038)		(0.037)
$= 3$		**-0.1046** **		**0.0164**		**-0.0609**
		(0.050)		(0.040)		(0.034)
$= 4$		**-0.0587**		**-0.0021**		**-0.0151**
		(0.043)		(0.031)		(0.034)
$= 6$		**0.0023**		**0.0485**		**0.0517** **
		(0.032)		(0.026)		(0.021)
$= 7$		**0.0391**		**0.0763*** **		**0.0502** **
		(0.030)		(0.027)		(0.023)
$= 8$		**0.0741** **		**0.0888*** **		**0.0858*** **
		(0.032)		(0.028)		(0.025)
$= 9$		**0.1015** **		**0.1539*** **		**0.1372*** **
		(0.042)		(0.029)		(0.029)
$= 10$		**0.1050*** **		**0.1853*** **		**0.1199*** **
		(0.040)		(0.032)		(0.029)
Constant	**0.8745*** **	**0.8644*** **	**0.7203*** **	**0.6783*** **	**0.7296*** **	**0.6990*** **
	(0.036)	(0.042)	(0.034)	(0.038)	(0.032)	(0.031)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mauza Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	5500	5434	5500	5434	5500	5434
R^2	0.62	0.63	0.67	0.68	0.67	0.67

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below.
Lagged score refers to a student's previous year average test score. Outcome refers to how many questions parents expect a student to answer correctly on a test.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table A.11: The Effect of Repetition on Parent Beliefs, including mauza fixed effects

	School Performance		Work Ethic		Inteligence	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted	**-0.1525****	**-0.1214***	**-0.3752****	**-0.3345***	**-0.2055**	**-0.1851**
	(0.066)	(0.068)	(0.172)	(0.178)	(0.157)	(0.157)
Lagged Average Score	**0.1406*****	**0.1134****	**0.1971***	**0.1704***	**0.3106*****	**0.2610****
	(0.050)	(0.045)	(0.104)	(0.102)	(0.109)	(0.107)
2yr lagged Average Score	**-0.0335**	**-0.0497**	**0.0496**	**0.0429**	**-0.0917**	**-0.1078**
	(0.053)	(0.052)	(0.121)	(0.121)	(0.130)	(0.129)
Lagged School Performance		**0.2004*****				
		(0.028)				
Lagged Hardworking				**0.1793*****		
				(0.066)		
Lagged Inteligence						**0.3012*****
						(0.067)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
FixedEffect	Mauza	Mauza	Mauza	Mauza	Mauza	Mauza
N	741	741	741	741	741	741
R^2	0.31	0.36	0.28	0.29	0.28	0.31

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Outcome refers to how many questions parents expect a student to answer correctly on a test.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table A.12: The Effect of Repetition on Parent Investments, including mauza fixed effects

	Monthly \$ $(t+1)$		Annual \$ $(t+1)$		Monthly \$ $(t+2)$		Annual \$ $((t+2))$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-21.6038**	**-12.2776**	**-168.7079****	**-126.5743***	**-34.7777**	**-31.2394**	**-252.7615***	**-246.0382***
	(22.763)	(20.516)	(74.742)	(72.986)	(26.330)	(26.206)	(135.423)	(133.723)
Lagged Average Score	**22.9649**	**14.8470**	**30.8385**	**24.6540**	**34.0989**	**31.0190**	**123.1262**	**122.1393**
	(15.310)	(13.177)	(51.434)	(49.954)	(22.477)	(22.823)	(88.797)	(88.922)
2yr lagged Average Score	**33.3709**	**10.6115**	**19.3269**	**4.2053**	**-10.1118**	**-18.7466**	**-17.4110**	**-19.8240**
	(26.416)	(20.072)	(61.363)	(59.183)	(30.279)	(30.616)	(117.132)	(117.977)
Lagged Monthly \$		**0.9255*****				**0.3511*****		
		(0.154)				(0.095)		
Lagged Anual \$				**0.4308*****				**0.0687**
				(0.087)				(0.117)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
FixedEffect	Mauza	Mauza	Mauza	Mauza	Mauza	Mauza	Mauza	Mauza
N	741	741	741	741	741	741	741	741
R^2	0.24	0.46	0.35	0.40	0.22	0.24	0.21	0.21

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Lagged score refers to a student's previous year average test score. Full descriptions of short-term and long-term investments are in the text.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table A.13: The Effect of Repetition on Parent Perceptions, controlling for birth order

	School Performance		Work Ethic		Inteligence	
	(1)	(2)	(3)	(4)	(5)	(6)
Not Promoted	**-0.1492****	**-0.1296****	**-0.3212****	**-0.3056***	**-0.2507***	**-0.2356***
	(0.062)	(0.063)	(0.159)	(0.160)	(0.130)	(0.125)
Oldest Child	**0.0421**	**0.0419**	**0.1166**	**0.1137**	**0.0763**	**0.0832**
	(0.056)	(0.055)	(0.124)	(0.124)	(0.103)	(0.104)
Lagged Average Score	**0.1818*****	**0.1675*****	**0.2572*****	**0.2496*****	**0.3771*****	**0.3499*****
	(0.039)	(0.038)	(0.090)	(0.089)	(0.079)	(0.079)
2yr lagged Average Score	**-0.0379**	**-0.0600**	**0.0273**	**0.0191**	**-0.1491**	**-0.1700**
	(0.050)	(0.049)	(0.112)	(0.112)	(0.112)	(0.110)
Lagged School Performance		**0.1531*****				
		(0.027)				
Lagged Hardworking				**0.0818**		
				(0.061)		
Lagged Inteligence						**0.2373*****
						(0.062)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	741	741	741	741	741	741
N	0.13	0.17	0.08	0.08	0.10	0.12

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Oldest sibling is a dummy variable equal to one if the student is the oldest in his/her family and zero otherwise. Lagged score refers to a student's previous year average test score. Outcome questions are listed in full in the text.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table A.14: The Effect of Repetition on Parent Investments, controlling for birth order

	Monthly \$ $(t+1)$		Annual \$ $(t+1)$		Monthly \$ $(t+2)$		Annual \$ $((t+2))$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-18.5447**	**-14.9605**	**-182.6158*****	**-117.7518****	**-46.4322****	**-45.1035****	**-225.9871***	**-216.1331***
	(15.012)	(18.069)	(59.080)	(57.530)	(22.536)	(21.393)	(115.493)	(114.227)
Oldest Child	**-14.1284**	**1.4378**	**-109.2693***	**-105.6470***	**13.3169**	**19.0874**	**-272.7512****	**-272.2009****
	(16.614)	(14.351)	(65.051)	(62.786)	(39.538)	(39.701)	(115.424)	(115.183)
Lagged Average Score	**12.0783**	**-0.0286**	**-17.3809**	**-22.2825**	**13.3150**	**8.8268**	**117.4875**	**116.7429**
	(15.614)	(13.356)	(47.465)	(45.464)	(17.335)	(17.040)	(76.016)	(75.939)
2yr lagged Average Score	**34.0652**	**10.4871**	**105.7818***	**68.4899**	**20.1242**	**11.3837**	**41.3761**	**35.7108**
	(22.508)	(15.231)	(60.350)	(58.244)	(21.598)	(21.562)	(100.894)	(100.496)
Lagged Monthly \$		**0.8895*****				**0.3298*****		
		(0.150)				(0.090)		
Lagged Anual \$				**0.4736*****				**0.0720**
				(0.055)				(0.086)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	741	741	741	741	741	741	741	741
R^2	0.05	0.33	0.05	0.17	0.04	0.06	0.06	0.06

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Oldest sibling is a dummy variable equal to one if the student is the oldest in his/her family and zero otherwise. Lagged score refers to a student's previous year average test score. Full descriptions of short-term and long-term investments are in the text.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table A.15: Effect of repetition on siblings, controlling for birth order

	Monthly \$ $(t+1)$		Annual \$ $(t+1)$		Performance (t+1)		Work Ethic (t+1)		Inteligence (t+1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Not Promoted	**-1.3319**	**4.0869**	**29.4357**	**76.2711**	**0.1443**	**0.0292**	**0.2954***	**0.2476**	**0.0236**	**0.0076**
	(29.907)	(27.235)	(79.599)	(70.498)	(0.136)	(0.117)	(0.167)	(0.168)	(0.158)	(0.156)
Oldest Child	**-29.0792**	**-19.0896**	**2.1355**	**22.5209**	**-0.0154**	**0.1019**	**0.1819**	**0.2170**	**0.0281**	**0.0045**
	(18.405)	(17.608)	(76.703)	(75.262)	(0.101)	(0.090)	(0.134)	(0.139)	(0.105)	(0.107)
Lagged Average Score	**26.2313**	**13.4989**	**92.2446**	**61.9526**	**0.1588***	**0.1139***	**0.1924****	**0.1885****	**0.0810**	**0.0649**
	(16.501)	(14.158)	(60.420)	(56.814)	(0.086)	(0.066)	(0.082)	(0.080)	(0.079)	(0.074)
2yr lagged Average Score	**20.2750**	**15.5704**	**18.1942**	**31.9374**	**-0.0197**	**-0.0262**	**-0.2994*****	**-0.3009*****	**-0.0822**	**-0.0165**
	(21.524)	(18.031)	(75.259)	(69.530)	(0.106)	(0.082)	(0.106)	(0.103)	(0.106)	(0.100)
Lagged Sibling Monthly \$		**1.9346*****								
		(0.267)								
Lagged Sibling Annual \$				**1.1228*****						
				(0.223)						
Lagged Sibling Overall						**0.4921*****				
						(0.035)				
Lagged Sibling Hardworking								**0.2495*****		
								(0.089)		
Lagged Sibling Inteligence										**0.3271*****
										(0.079)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	719	719	719	719	719	702	719	702	719	702
R^2	0.05	0.20	0.07	0.22	0.07	0.34	0.04	0.06	0.02	0.05

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. The table shows results from a student-level regression of average household expenditure on all siblings (leaving out the student) on whether the student repeated or not (columns 1-4) and of average parent evaluation of siblings (leaving out the student) on whether the student repeated or not. Oldest sibling is a dummy variable equal to one if the student is the oldest in his/her family and zero otherwise. Lagged score refers to a student's previous year average test score. Full descriptions of short-term and long-term investments are in the text.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level

Table A.16: Effect of repetition on sibling educational expenditures, controlling for birth order and outcomes 2 years later

	Monthly \$ $(t+1)$		Annual \$ $(t+1)$		Monthly \$ $(t+2)$		Annual \$ $((t+2))$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not Promoted	**-1.3319**	**4.0869**	**29.4357**	**76.2711**	**19.1249**	**29.3617**	**37.8810**	**98.9245**
	(29.907)	(27.235)	(79.599)	(70.498)	(55.013)	(50.289)	(226.394)	(219.577)
Oldest Child	**-29.0792**	**-19.0896**	**2.1355**	**22.5209**	**22.0088**	**45.1024**	**-121.6904**	**-79.7053**
	(18.405)	(17.608)	(76.703)	(75.262)	(34.860)	(31.060)	(105.804)	(92.857)
Lagged Average Score	**26.2313**	**13.4989**	**92.2446**	**61.9526**	**23.8705**	**10.8207**	**123.7902**	**86.8411**
	(16.501)	(14.158)	(60.420)	(56.814)	(21.108)	(18.382)	(82.328)	(80.022)
2yr lagged Average Score	**20.2750**	**15.5704**	**18.1942**	**31.9374**	**7.3671**	**3.1839**	**25.7512**	**55.4716**
	(21.524)	(18.031)	(75.259)	(69.530)	(27.075)	(22.042)	(102.548)	(95.254)
Lagged Sibling Monthly \$		**1.9346*****				**2.5869*****		
		(0.267)				(0.382)		
Lagged Sibling Annual \$				**1.1228*****				**1.3520*****
				(0.223)				(0.311)
Age/Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag Polynomial	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order	4th Order
N	719	719	719	719	684	684	684	684
R^2	0.05	0.20	0.07	0.22	0.05	0.24	0.04	0.15

Note: Coefficient estimates bolded, standard errors (clustered at school-level) in parentheses below. Oldest sibling is a dummy variable equal to one if the student is the oldest in his/her family and zero otherwise. Lagged score refers to a student's previous year average test score. Full descriptions of short-term and long-term investments are in the text.
*** Significant at 1% level, ** Significant at 5% level, * Significant at 1% level