THE COST OF GRADE RETENTION

Marco Manacorda*

Abstract—This paper uses administrative longitudinal microdata on junior high school students in Uruguay to measure the effect of grade failure on students' subsequent school outcomes. Exploiting the discontinuity induced by a rule establishing automatic grade failure for pupils with more than three failed subjects, I show that grade failure leads to substantial dropout and lower educational attainment even four to five years after grade failure occurred.

I. Introduction

GRADE repetition in school is common in less developed countries and often accompanied by low enrollment and high dropout rates, the combination of the two often referred to as "wastage."

Figure 1 plots gross enrollment rates in secondary school over grade repetition rates in primary school in 65 countries.[1] Sub-Saharan Africa shows both the lowest enrollment rate (31%) and the highest repetition rate (around 20%). At the other end of the spectrum, Central Asia, eastern and western Europe, and North America display repetition rates that vary between 1% and 2% and enrollment rates that vary between 86% and 112%. Latin America, North Africa, the Middle East, and Southeast Asia locate somewhere halfway, with repetition rates between 6% and 9% and enrollment rates between 62% and 73% (on Latin America see also Urquiola & Calderon, 2006). Figure 2 also shows that repetition tends to be negatively associated with low levels of income per capita.

Do the hurdles that repetition creates for students' transition through the school system explain why a large fraction of students eventually drop out? Or is the correlation in figure 1 spurious, due, for example, to the circumstance that where the demand for education is low, the efficiency of the system, measured by grade promotion, is also low, perhaps due to low public investment in education? Do poor teachers and schools' quality, teachers' absenteeism, and lack of school infrastructures, often cited as major problems of school systems in developing countries, explain both high repetition rates and students' incentives to abandon the system?[2] Or do students in these countries find it harder to progress through the system due to lack of financial resources, a higher opportunity cost of attending school, or malnutrition, hence leading to both repetition and dropout?

The desirability of grade retention is a controversial issue. This reflects a substantial disagreement on whether grade repetition is beneficial to students and society at large and, more fundamental, the circumstance that there are both costs and benefits associated with this policy.

Although not undisputed (Alexander, Entwisle, & Kabbani, 2003), there is a view among psychologists and part of the pedagogical profession that early grade repetition does not lead to improvements in school achievement (McCoy & Reynolds, 1999), while raising dropout, with negative socio-emotional consequence (Jimerson, Anderson, & Whipple, 2002). Low self-esteem, due to disenfranchisement or stigmatization, low expectations on the part of the environment, or the cost of readjusting to a new class and possibly a new teacher might worsen a student's outcomes and eventually result in dropout.

A different view emphasizes the benefits of grade repetition, which, according to this view, might reinforce a student's knowledge or discipline, with potential beneficial effects on subsequent outcomes. Additional exposure to teaching, especially in early grades, might make a student more apt, and hence presumably more likely, to pursue higher levels of education. Repetition might also improve the quality of the match between the school and the student if his development makes him more apt to attend a certain grade at a later age or if changing peers and teachers leads to an increase in productivity. According to this view, grade repetition is an efficient mechanism to reallocate students to classes.

Possibly the strongest argument in favor of grade repetition is that it acts as a deterrent to poor school performance. By inflicting a high penalty on underperformers, this policy creates an incentive for students to increase their efforts (see Jacob, 2005, on the incentive effect of high-stakes exams on students' outcomes), although this might come at a cost, since students take longer to move through the system. Experiencing the penalty of repeating a grade might also make repeaters less likely to want to experience this again, hence creating an incentive to improve their school

[1] Data come from UNESCO (2002) and refer only to countries that report positive repetition rates. Although the U.S. Department of Education does not provide official figures on repetition, estimates (not in the figure) from the CPS show that around 12% of individuals aged 12 to 15 have repeated at least a grade (Cascio, 2005).

[2] On teachers' absenteeism in Indian schools, see Banerjee and Duflo (2006) and Chaudhury et al. (2005). On the effect of the supply of schools on enrollment, see Duflo (2001).
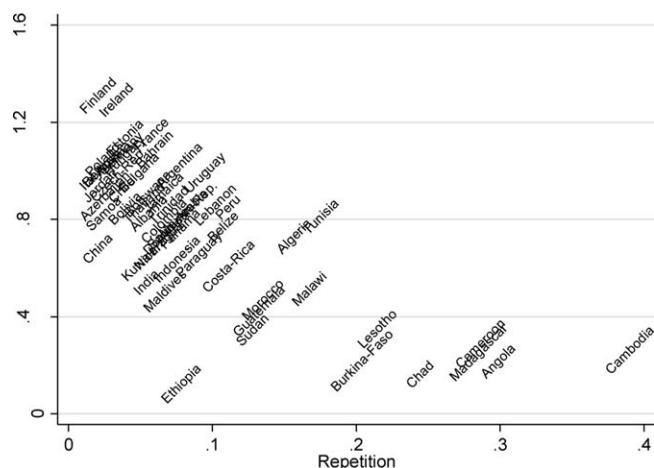
FIGURE 1.—REPETITION RATES IN PRIMARY SCHOOL AND GROSS ENROLLMENT RATE
IN SECONDARY SCHOOL ACROSS COUNTRIES

The graph plots the repetition rate in primary school on the horizontal axis and gross enrollment in secondary school on the vertical axis. The sample refers to countries with income per capita not greater than $28,000. UNESCO (2002).

FIGURE 2.—REPETITION RATES IN PRIMARY SCHOOL AND PPP GDP PER CAPITA
(US$1,000) ACROSS COUNTRIES, 1999

The graph plots PPP GDP per capita on the horizontal axis and the repetition rate in primary school on the vertical axis. See also the note to figure 1.
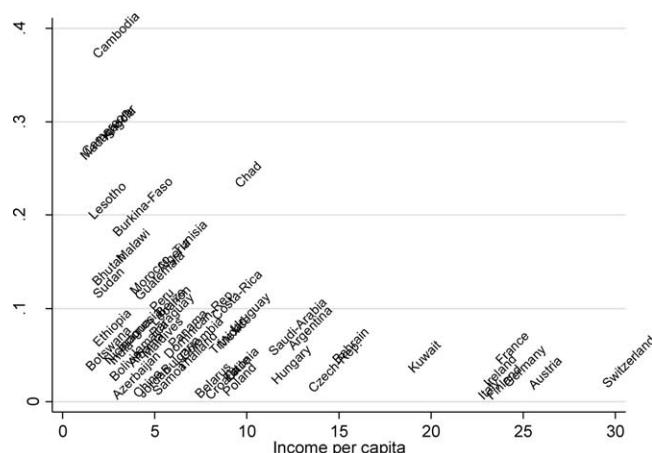
performance, possibly because of learning or the increasing marginal cost of repeating a grade.

Although there is a rather copious body of research on the determinants of grade repetition, convincing quasi-experimental evidence on its effects is scarce, especially for low- and middle-income countries, where repetition is widespread.[3] The main difficulty in identifying the effect of grade failure on subsequent school outcomes is that latent school outcomes—the ones that would be observed in the absence of grade failure—and the propensity to fail a grade are likely to be simultaneously determined. Similar to the spurious cross-country correlation discussed above, characteristics of the pupil, such as her ability or motivation, her teachers, the school, and the environment, are likely to affect simultaneously grade failure, stay-on rates, and attainment. Such correlations will likely overestimate the impact of grade failure on subsequent school outcomes.

Three papers, all for the United States, account explicitly for the potential endogeneity of failure rates. Jacob and Lefgren (2004, 2009) use the discontinuous relationship between test scores and promotion to assess the causal impact of grade repetition on achievement among Chicago public school students. Their results show a positive short-term effect of grade retention on third graders' achievement

and no effect on sixth graders'. They also show that grade retention in eighth grade leads to an increase in dropout and a drop in the probability of high school completion. Using the variation in age of entry into kindergarten across states as an instrument for repetition, Eide and Showalter (2001) conclude that for white students, grade repetition leads to lower dropout and higher earnings, although results are not statistically significant.

In order to circumvent the identification problem, in this paper I suggest using a rule in force in secondary junior high school (grades 7 to 9) in Uruguay, a country with remarkably high repetition rates, that establishes automatic grade failure for pupils with more than three failed subjects at the end of the school year. I exploit the discontinuity in grade advancement induced by this rule to assess the causal impact of grade failure on dropout and school attainment later in life. I find that grade failure induces students to drop out at the end of the school year when failure occurs, which has long-lasting effects on their attainment.

One concern with the proposed strategy is that assignment around the threshold of three failed subjects might not be as good as random. This might happen if otherwise better-performing students are able to sort strategically just below the threshold or if teachers are able to manipulate final scores and promote students with better latent outcomes. Indeed, in the empirical analysis, I find an excess bunching of students precisely at three failed subjects that might be taken as an indication of score manipulation. To deal with this issue, I derive worst-case scenario selection estimates assuming that the excess bunching is entirely accounted for by the reclassification of the best-performing pupils among those who would have otherwise (i.e., in the absence of manipulation) failed four subjects. These estimates remain negative, implying that the main conclusion of this paper is not driven by nonrandom sorting around the threshold and that grade failure unequivocally harms subsequent school progression.

[3] Evidence on the determinants of grade repetition illustrates a causal effect of family socioeconomic status: Oreopoulos, Page and Stevens, (2006) for evidence on parental education, Maurin and Goux (2005) for evidence on residential overcrowding, Pischke (2007) for evidence on the length of the school year, and Currie and Neidell (2007) for evidence on Head Start. For less developed countries, there is evidence that family background and school inputs are important determinants of grade failure (Gomes-Neto & Hanushek, 1994, for Brazil and Patrinos & Psacharopoulos, 1996, for Bolivia and Guatemala). Evidence from Colombia based on a randomized school voucher program illustrates the positive effect of increased school choice and the ability to afford private education on promotion rates (Angrist et al., 2002). Conditional and unconditional cash transfers also appear to have a positive effect on grade promotion (Behrman, Sengupta, & Todd, 2005; Schady & Araujo, 2008).

The structure of the paper is as follows. Section II provides background information on the Uruguayan school system. Section III presents the data, section IV discusses the specification and identification of the regression model, section V presents the regression results, and section VI concludes.

II. The School System in Uruguay: Background

Uruguay boasts a long tradition of publicly provided education and social inclusion. Primary school was made compulsory in 1877, universal primary schooling was achieved in the 1950s, and the literacy rate is among the highest in the region (97% for men and 98% for women).

The school system is organized in three basic cycles: primary (grades 1–6), junior high (grades 7–9) and senior high (grades 10–12). Both primary and junior high schooling are compulsory.[4] Junior and senior secondary education are offered in both *Liceos* (nonvocational secondary schools), and in vocational colleges, UTUs (Universidad del Trabajo del Uruguay, literally the Uruguayan Employment University).[5]

Even if Uruguay still ranks high in terms of educational outcomes compared to the rest of Latin America, its education system is not problem free.[6] While enrollment in primary school is almost universal and the majority of children start school at age 6, the system is unable to retain a large share of students in junior high (Da Silveira & Queirolo, 1998; Furtado, 2003; Bucheli & Casacuberta, 2000).[7] One of the hurdles that students face during their progression through the school system is the high probability of failing a grade.[8]

Grade progression in junior high depends on students' performance during the school year, which runs from March to December.[9] For each of the taught subjects (between nine and eleven, depending on the school grade), students are assigned a score on a scale 1 to 12.[10] Students pass a subject if the associated score is no lower than 6. Those who fail a subject must eventually take remedial exam sessions. The first opportunity to retake an exam is just before the beginning of the subsequent school year, in February, and subsequent retake exams take place in July and December each year.

A necessary condition for promotion to the next grade is that the student has no more than three accumulated failed subjects by the beginning of the following school year (after the February retake session). Accumulated fails include both failed subjects in the current school year and subjects failed in previous school years that the student has not in the meantime passed.[11] In the rest of the paper, I use the discontinuity in grade progression between students with three failed subjects (who barely pass) and students with four failed subjects (who barely fail) to identify the effect of grade failure on later school outcomes.

III. Data

The data used in this paper refer to students in junior high (grades 7 to 9) in 1996 and 1997. The data follow these students' progression in both junior and senior high up to 2001 and report information on the institution and grade attended in each year, whether the student passed or failed that grade, number of missed school days, basic demographics (age and gender), and, only for the school years 1996 and 1997, scores for each subject at the end of the school year (in December, that is, ignoring the results from the subsequent February retake session).

The data include almost the universe of public nonvocational schools but exclude UTUs and private institutions.[12] Because of this, I can measure whether a student is retained within the public (whether junior or senior) nonvocational system, but I am unable to distinguish those leaving the educational system *tout court* from those moving to private or vocational institutions. I return to this at the end of the paper. In addition, since there is no information on promotion or failure in 1999 and 2001, I measure school progression as maximum grade attended (as opposed to completed), independent of whether the student actually passed or failed the grade he was last observed attending.

---

[4] Minimum working age over the period of analysis was 14, hence lower than the minimum age required to complete compulsory education.

[5] As of 2001, enrollment in vocational junior high schools accounted for around 11% of overall enrolment in public junior high (Consejo de Educación Técnico Professional, 2005). Not dissimilar from many other Latin American countries, private fee-based education is common, covering about 15% of secondary school enrollment.

[6] Between 1960 and 2000, for example, average education in the population over age 25 in the United States rose by around 4.5 years (from 8.7 to 12.2), while in Uruguay this growth was on the order of 2.1 years (from 5.1 to 7.2) (Barro & Lee, 2001).

[7] Data from a module of the 2001 National Household Survey (Encuesta Continua des Hogares) show, for example, that only around 80% of 24- to 29-year-olds declare having at one point started junior high, and among these, 16% declare not having completed it. Starting from the mid-1990s and in recognition of these problems, a major reform of the educational system has taken place (Administración Nacional de Educación Publica, 2000). This includes, among other things, an additional year of compulsory preprimary education for five-year-olds.

[8] In this paper, I focus specifically on grade failure—the requirement to attend again a certain grade in order to progress. This is different from repetition (or grade retention)—the actual reattendance of a certain school grade following failure.

[9] The regulation mentioned here (Administracíon Nacional de Educación Publica, 1996) refers to old curriculum (Plan 86) and applies only to the school years 1996 and 1997. In 1996 a new curriculum was introduced (Plan 96) that changed both the content and the structure of teaching, the length of the school day (from 3.5 hours to 5 hours), and the rules determining promotion. As Plan 96 was introduced experimentally in a few schools (Liceos Pilotos), the majority of students in 1996 and 1997 were still under the old plan. The analysis in this paper excludes Liceos Pilotos.

[10] I refer to "score" as opposed to subject "grade" as appropriate to avoid confusion with the school grade attended.

[11] The second condition that must be simultaneously fulfilled is that the student has accumulated no more than 25 missed school days during the year. In an earlier version of this paper, I used this rule to identify the effect of grade failure on progression. It turns out, however, that subject scores vary discontinuously at the 25 missed-school-days threshold, casting some doubt on the validity of the identification assumption. For this reason, in this version I use only the discontinuity at three failed subjects.

[12] A few schools are not in the sample, although this problem tends to be less serious at the end of the period: the number of missing institutions is 56 in 1996, 59 in 1997, 13 in 1998, and 4 in 2000 (out of around 250 schools).

TABLE 1.—DESCRIPTIVE STATISTICS: STUDENTS IN JUNIOR HIGH SCHOOL, URUGUAY, 1996–1997

		Failers	Passers	All
1.	Grade failure	1	0	0.259
2.	Additional grades attended (by 2000/2001)	0.599	2.398	1.931
3.	Censored	0.150	0.646	0.517
4.	Failed subjects	6.351	0.928	2.335
5.	Mean subject score	4.701	7.229	6.573
6.	Missed school days	35.167	10.057	16.572
7.	School grade	7.880	7.893	7.889
8.	Age	13.891	13.174	13.360
9.	Grade-age distortion	0.538	−0.065	0.092
10.	Female	0.457	0.575	0.544
	Number of observations	25,878	73,851	99,729

The table reports information on students in junior high School in 1996 and 1997. The first column refers to those who failed a grade, the second column to those who passed, and the last column to the pooled sample. Row 1 reports the proportion failing that grade, row 2 the number of additional grades attended (until 2001 if a student failed a grade or until 2000 if a student did not fail), row 3 the proportion still in school at the end of the period (2000 for nonfailers and 2000 for failers), row 4 the proportion of failed subjects, row 5 the average score across all subjects, row 6 the average number of missed school days, row 7 the average school grade (from 7 to 9), row 8 the average age, row 9 the average grade distortion (age-grade), and row 10 the proportion of girls. Source: *Bases de datos de rendimiento a nivel de estudiantes en educación secundaria, Administración Nacional de Educación Pública*.

Table 1 provides some descriptive statistics. The data refer to individuals with valid subject scores. The first column refers to those who fail, the second column to those who pass, and the last column to the entire sample. Overall, there are 99,729 observations on 73,621 individuals.

Row 1 shows that one in four children fail a grade each year. Row 2 examines the number of additional school grades attended by the end of the period of observation, the main outcome variable. This is the difference between maximum grade attended by 2001 and the grade where the student was observed in junior high in 1996 or 1997. This variable ranges from 0 to 5. Since grade failers have to be exposed to one extra year of schooling in order to potentially make up for the year lost due to repetition, I follow passers only until 2000. The average number of additional grades attended is on the order of 1.9. However, while passers accumulate around 2.4 extra school years, failers accumulate only approximately one extra half of a school year. This clearly suggests that grade failure is associated with worse school outcomes later in life.

Row 3 reports the survival rate—the probability of being in the sample in the last year of observation (respectively, 2000 for passers and 2001 for failers). This is on the order of 15% for failers and almost 50 percentage points higher for passers. Since right censoring is much more pronounced for passers, this implies that the estimated attainment gap between passers and failers in row 2 is likely to be underestimated.

The rest of the table shows that failers are more likely than passers to display characteristics that are associated with poor school performance. Rows 4 to 6 show that grade failers, relative to passers, display higher absenteeism (35 versus 10 missed school days), a higher number of failed subjects (with a score below 6, this number is 6.3 for failers versus 0.9 for passers), and a lower mean subject score (4.7 versus 6.2). The following rows show that grade failure is

equally likely to occur in each of the school grades (row 7) and that failers are clearly older (row 9) than passers, implying that early repeaters are more likely to repeat again.

IV. Specification and Identification

To identify the impact of grade failure on school outcomes, I present regressions based on a (fuzzy) regression discontinuity design derived from the promotion rule. The regression model allows controlling for the observed characteristics of students and their schools, as well for the potential bias in the regression coefficients that stems from differential censoring in school outcomes across individuals originally observed in different grades (7, 8, or 9) or different years (1996 or 1997).

Ignoring for simplicity other covariates, suppose that school outcomes Y depend additively on a continuous function $f(.)$ in the number of failed subjects, S, and on grade failure, F:

$$Y = \beta_0 + \beta_1 F + f(S - 3) + u, \tag{1}$$

where u is an error term and $f(0) = 0$. The error term potentially includes a student's past attainment as well as other unobserved determinants of performance. As already pointed out, the OLS estimate of equation (1) is biased if u is correlated with F due to unobserved heterogeneity of reverse causality.

In order to circumvent this problem, I use the discontinuity in the failure rate at three failed subjects as an instrument for grade failure in equation (1). Consistent with the rule, I assume that grade failure is a continuous function of the number of failed subjects $g(S - 3)$ plus a dummy for more than three failed subjects $P = I(S > 3)$:

$$F = \gamma_0 + \gamma_1 P + g(S - 3) + v, \tag{2}$$

where $g(0) = 0$. Under the assumption that if not for the rule governing grade failure, school progression varies continuously around the three-failed-subjects threshold, an instrumental variable estimate of equation (1), where equation (2) is the first-stage equation, leads to a consistent estimate of β_1, the parameter of interest (Hahn, Todd, & Van der Klaauw, 2001).[13] Consistency of the Instrumental Variable (IV) estimator requires individuals not to sort around the discontinuity point based on unobserved determinants of the outcome variable, that is, that assignment around the discontinuity is as good as random. I return to this point later in the paper.

Figure 3 plots the probability of failing a grade on the number of failed subjects in the current year (see equation [2]). The size of each point is proportional to the number of

---

[13] There are plenty of applications of the RD design to schooling data. Typically procedures and regulations attaining to students', teachers', or schools' behavior lead to discontinuities in treatment. See, for example, Angrist and Lavy (1999), van der Klaauw (2002), Jacob and Lefgren (2004, 2009), and Urquiola and Calderon (2006).

FIGURE 3.—GRADE FAILURE BY NUMBER OF FAILED SUBJECTS IN JUNIOR HIGH SCHOOL, URUGUAY, 1996–1997
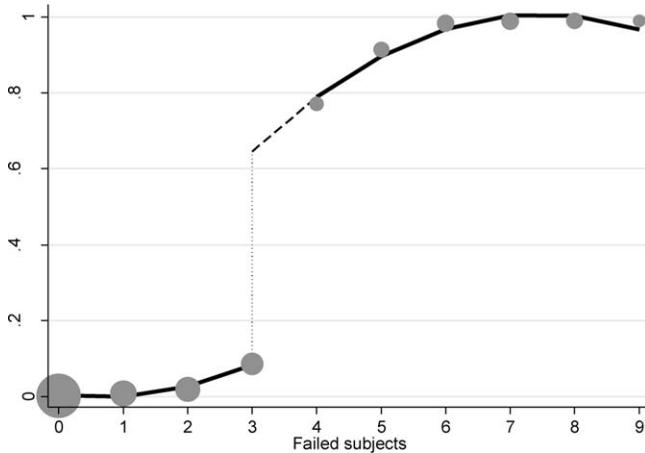
The figure reports the proportion of individuals failing a grade by number of failed subjects in the year. A quadratic spline is superimposed to the data on either side of the threshold. Source: *Bases de datos de rendimiento a nivel de estudiantes en educación secundaria, Administración Nacional de Educación Pública.*

FIGURE 4.—ADDITIONAL GRADES ATTENDED BY 2001 (2000 FOR NONFAILERS) BY NUMBER OF FAILED SUBJECTS IN JUNIOR HIGH SCHOOL, URUGUAY, 1996–1997

The figure reports the number of additional grades attended by 2001 among those in junior high in 1996 or 1997 as a function of the number of failed subjects. The series is censored to the year 2000 for those who did not fail. A quadratic spline is superimposed to the data on either side of the threshold. See also the note to figure 3.

observations. Because I have no information on subjects failed in previous school years that the student has not in the meantime passed, I use only information on failed subjects in the current year, hence abstracting from the circumstance that students with three or fewer failed subjects in the current school year might eventually fail. Similarly, since I have no information on the result from the February retake session, just before the beginning of the following school year, I use information on failed subjects at the end of the school year (in December), hence ignoring that some students with four or more failed subjects in December might still eventually pass. Although this induces some fuzziness in the relationship between school progression and number of failed subjects, for the purposes of the identification, it is sufficient that some discontinuity is present at three failed subjects.

Figure 3 shows that the probability of failing a grade is 0.2% for those with no failed subjects, and it rises to around 9% at three failed subjects.[14] One can see that at four failed subjects, the probability of grade failure jumps to around 77%, and it continues to grow. This is consistent with the rule establishing grade failure with more than three failed subjects, although for the reasons mentioned above, compliance is imperfect.

Since grade failure appears to be a discontinuous function of the number of failed subjects, for failure rates to have an effect on school outcomes, one would expect outcomes to vary discontinuously at the threshold of three failed subjects. Figure 4 analyzes the correlation between additional grades attended (up to 2001 for failers and up to 2000 for passers) and number of failed subjects for those in junior high between 1996 and 1997.
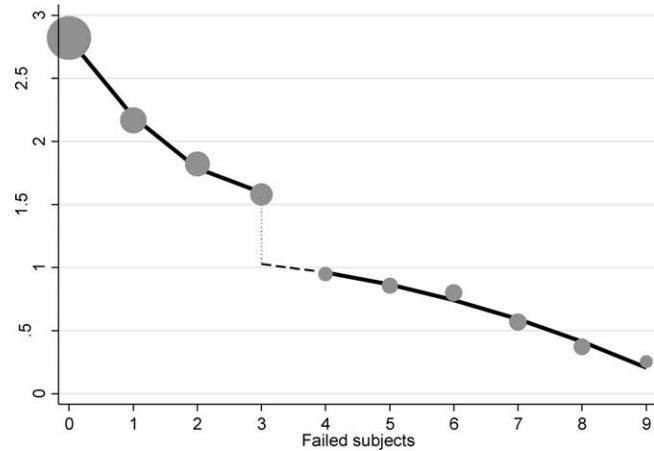
The number of additional grades attended falls monotonically with the number of failed subjects in junior high. More interesting is the large jump in the outcome variable that is apparent between three and four failed subjects. This is a fall of around half a year from around 1.5 years at three failed subjects to 1 year at four failed subjects. Because this is the mirror image of the effect of failed subjects on grade failure in figure 3, this suggests that grade failure has a negative effect on subsequent school outcomes.

In the following section, I present estimates of model (1) and test for the sensitivity of the results to the inclusion of a number of observable controls (see Lee, 2008, for an interesting application; see also McEwan & Urquiola, 2005) and to the possibility that the results are driven by nonrandom selection around the threshold.

V. Estimates

A. RD Estimates: Failed Subjects, Grade Failure, and Final School Outcomes

In this section I present IV estimates of model (1). I measure the effect of the rule on failure rate in equation (2) as the estimated difference between the actual and the counterfactual failure rates at three failed subjects. I pool observations for 1996 and 1997, treating individuals who appear in the sample more than once as two separate observations. I follow Lee and Card's (2008) suggestion in the context of a regression discontinuity design with a discrete running variable by modeling $f(.)$ as a parametric spline, whose shape and intercept I allow to vary on either side of the discontinuity point. In practice, I model $f(.)$ as a parametric polynomial in its argument $(S - 3)$ interacted with the dummy for more than three failed subjects $P = I(S > 3)$. The coefficient γ_1 in equation (2) hence measures the predicted

---

[14] Note that students with fewer than three failed subjects can still fail due to the additional rule that prescribes a maximum number of 25 missed school days for promotion.

TABLE 2.—FAILED SUBJECTS, GRADE FAILURE, AND SUBSEQUENT SCHOOL ATTAINMENT IN JUNIOR HIGH SCHOOL, URUGUAY, 1996–1997

	(1)	(2)	(3)	(4)	(5)
	First Stage (Dependent Variable: Grade Failure $t = 0$)				
I(Failed subjects$_0 > 3$)	0.562***	0.428***	0.417***	0.438***	0.426***
	(0.034)	(0.008)	(0.024)	(0.087)	(0.085)
	Reduced Form (Dependent Variable: Additional Grades)				
I(Failed subjects$_0 > 3$)	−0.565***	−0.759***	−0.381**	−0.380**	−0.324**
	(0.049)	(0.092)	(0.134)	(0.173)	(0.153)
	Instrumental Variables (Dependent Variable: Additional Grades)				
Grade failure $t = 0$	−1.005***	−1.773***	−0.912**	−0.869**	−0.760**
	(0.136)	(0.236)	(0.373)	(0.358)	(0.327)
Polynomial of degree	2	3	4	4	4
School fixed effects	No	No	No	Yes	Yes
Grade fixed effects	No	No	No	Yes	Yes
Year effects	No	No	No	Yes	Yes
Additional controls	No	No	No	No	Yes

The top panel reports the OLS coefficients from a regression of a dummy equal to 1 for grade failure on a dummy equal to 1 for more than three failed subjects in junior high (first-stage equation). The middle panel reports the coefficient of the number of additional grades attended by the end of the period (censored to the year 2000 for nonfailers) on the dummy for more than three failed subjects (reduced-form equation). The bottom panel reports the IV estimates of the effect of grade failure on the number of additional grades attended, where grade failure is instrumented by a dummy for more than three failed subjects. Each column refers to a different specification. Specifications in columns 1 to 3 include, respectively, a parametric function in the number of failed subjects of degree 2, 3, and 4, interacted with a dummy for more than three failed subjects. Columns 4 and 5 also control for school, year, and grade fixed effects. Additional controls include dummies for number of missed school days, gender, age-grade distortion, and score in each subject. Standard errors in parentheses are clustered by number of failed subjects by school, grade, and year. Number of observations: 99,729. ***, **, * denote, respectively, significant at 1%, 5%, and 10% level. See also the note to table 1.

discontinuous increase in grade failure at the three-failed-subjects threshold that I attribute to the grade retention rule.

Row 1 of table 2 presents the OLS estimates of equation (2), where the dependent variable is 1 for individuals failing a grade and 0 otherwise (the first-stage equation). Column 1 includes no controls if not a quadratic spline in failed subjects, as in figure 3. Columns 2 and 3 include polynomials in the number of failed subjects of third and fourth order, respectively. Column 4 presents a similar specification as in column 3, with a fourth-order polynomial in the number of failed subjects, with the addition of school, year, and grade (7, 8, 9) fixed effects. Standard errors in all specifications are clustered by the number of failed subjects in each class-school-year group.

The estimated jump in the failure rate at the discontinuity point in column 2 is 56 percentage points. One can visualize this jump in figure 3, where I have superimposed on the data the estimated quadratic splines on either side of the threshold using the estimates in column 1, row 1 of table 2. The estimated gap at three failed subjects is the difference between the two curves fitted on either side of the threshold (0 to 3 and 4 to 9 failed subjects). The point estimate falls by around 20% (to, respectively, 43 or 42 percentage points) when cubic or quartic polynomials are included. The inclusion of school, grade, and time fixed effects makes virtually no difference to the estimated coefficient: this is 44 percentage points (column 4 of table 2).

In the middle part of table 2, I report reduced-form estimates of the model, where the dependent variable is now additional grades attended up to 2001 (censored to 2000 for nonfailers). In formulas,

$$Y = \delta_0 + \delta_1 P + h(S - 3) + e \qquad (3)$$

where we let $h(0) = 0$. Predicted values from column 1 are also reported alongside the actual data in figure 4: one can clearly see the predicted jump at three failed subjects on the

order of half a school year. Results are somewhat sensitive to the inclusion of higher-order polynomials: when a quartic in the number of failed subjects is included, the point estimate is −0.38 and significant at conventional levels. Estimates that additionally control for grade, school, and time fixed effects, in column 4, are essentially unchanged.

Instrumental variable estimates are presented in the bottom part of the table. In column 1, with a quadratic spline, point estimates are on the order of −1, implying that grade failure leads to a loss of around one year in education. The specification with a quartic polynomial in column 3 leads to a point estimate of −0.9, a reduction in schooling of just less than one school year. Results are similar when year, school, and grade fixed effects are included.[15]

B. Potential Threats to the Consistency of the IV Estimates

One major concern with the previous estimates is that assignment around the discontinuity point might not be as good as random, invalidating the conditions required for consistency of the RD estimator (Lee, 2008; McCrary, 2008). If pupils with better latent school outcomes are able to sort strategically precisely at three failed subjects or if teachers manipulate individual subject scores so that students with higher chances of progressing are made to pass,[16] then the IV estimate will be downward biased, exaggerating the negative effect of grade failure on subsequent school outcomes.

As a first check for nonrandom assignment, I examine the discontinuity in the observed covariates at the thresh-

---

[15] The OLS estimate with the same controls as those in column 5 is, for comparison, −1.746 (SE 0.017), around twice the RD estimate. As expected, OLS tends to exaggerate the negative effect of grade failure on subsequent school outcomes.

[16] See Jacob and Levitt (2003) for evidence on teachers' cheating in response to incentives.

old (Lee, 2008). If sorting around the discontinuity point is a serious concern, one would expect observed covariates that are known to affect progression to vary discontinuously at the threshold of three failed subjects. Table 3 reports reduced-form estimates where the dependent variable is in turn a separate variable. I report specifications with quartic polynomials in number of failed subjects, plus school, year, and grade fixed effects as in column 4 of table 2. Observed covariates are missed school days, gender, age-grade distortion, scores (from 1 to 12) in each subject, and mean score across all subjects. There is some slight evidence that those to the right of the threshold are worse performers—both accumulated delay and absenteeism are higher—but the coefficients are not statistically significant at conventional levels. For nine of the eleven subjects considered (with the exception of physical education and music/drawing), there is no evidence of a discontinuous change in test scores. The same is true for the mean test score.

This is further confirmed in column 5 of table 2, which reports the same specifications as in column 4 with the further addition of the controls in table 3 (dummies for missed school days, a gender dummy, age-grade distortion dummies, and dummies for the score in each subject). Relative to the specification with no controls in column 4, the IV estimate falls from around -0.9 to -0.8 but the difference is not significant at conventional levels.

As a second check for nonrandom assignment, figure 5 reports the distribution of failed subjects. The figure shows an almost monotonic fall in the density as the number of failed subjects increases. There is evidence, however, of some discontinuity in the density of the running variable at the threshold, with a bunching at three failed subjects and what appears to be some missing mass at four failed subjects. A discontinuity in the pdf of the running variable is sometimes taken as an indication of a failure of the random assignment hypothesis, since this is suggestive of students' or teachers' ability to manipulate the running variable (McCrary, 2008).

Although this test is uninformative on the direction of selection (whether positive, negative, or random), some progress can still be made by assuming worst-case scenario positive selection, that is, assuming that the excess mass at three failed subjects is entirely accounted for by students with positive latent outcomes who would have otherwise (in the absence of the progression rule) failed more than three subjects. This allows me to estimate an upper bound for the coefficient of interest.

Keeping with the evidence in figure 5 and consistent with the notion that reclassification must be costly for schools or teachers, I assume that only students at the margin of failing ($S^* = 4$) are reclassified, and only by a sufficient margin to make them barely pass ($S = 3$), where the asterisk denotes "true" unobserved variables. Worst-case-scenario positive selection implies that reclassified students are the best among those who would have otherwise failed four subjects.

TABLE 3.—FAILED SUBJECTS AND OBSERVED COVARIATES IN JUNIOR HIGH SCHOOL, URUGUAY, 1996–1997: REDUCED-FORM ESTIMATES

|	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
										Subject Scores					
	Missed School Days	Female	Age-Grade Distortion	Spanish	Math	Language	Geography	Biology	History	Physics	Drawing/Music	Physical Education	Chemistry	Social Education	Mean
	7.872	0.016	0.173	−0.091	0.133	0.181	−0.117	−0.240	0.010	0.029	−0.799**	0.753***	0.592	−0.135	0.003
	(4.905)	(0.068)	(0.143)	(0.256)	(0.291)	(0.277)	(0.235)	(0.250)	(0.254)	(0.254)	(0.359)	(0.273)	(0.415)	(0.363)	(0.102)

Each entry in the table refers to a separate regression of each dependent variable on a dummy for more than three failed subjects. All specifications include a parametric function in the number of failed subjects of degree 4, school, year and grade fixed effects. Number of observations is 28,259. See also the note to table 2. The dependent variable is available only for those in ninth grade and the number of observations is 99,729 except in columns 13 and 14, where the dependent variable is available only for those in ninth grade and the number of observations is 99,729.

FIGURE 5.—PROPORTION OF FAILED SUBJECTS IN JUNIOR HIGH SCHOOL,
URUGUAY, 1996–1997
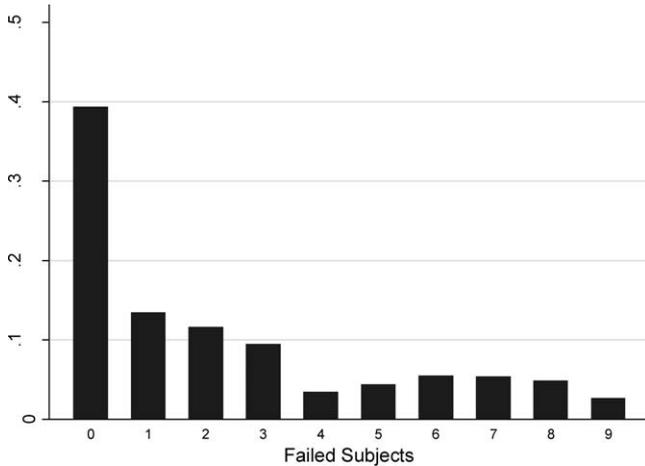
The figure reports the distribution of failed subjects. See also the note to figure 3.

FIGURE 6.—PROPORTION OF LATENT FAILED SUBJECTS IN JUNIOR HIGH SCHOOL,
URUGUAY, 1996–1997, ASSUMING WORST-CASE-SCENARIO SELECTION
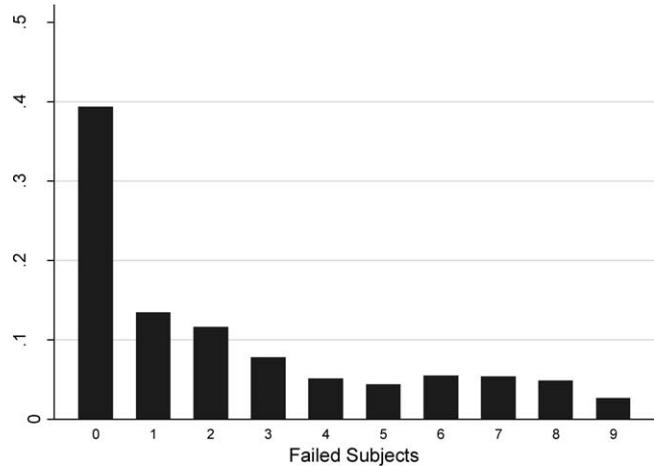
The figure reports the distribution of latent failed subjects assuming worst-case-scenario positive selection. See the text for details and the note to figure 3.

If the maximum outcome among those observed failing four subjects in each class-school-year group (G) is $Y_G = \max(Y|G, S = 4)$, a simple model for positive selection assumes that everyone with an outcome above this level has been reclassified. In formulas, this is equivalent to assuming the following misclassification rule:

$$S = 3 \quad \text{if } Y > Y_G \text{ and } S^* = 4$$
$$S = S^* \quad \text{otherwise,} \tag{4}$$

where the value of Y_G can be directly observed in the data and one can recover the value of S^* for all misclassified individuals in the sample. Under the assumptions in equation (4), the mean classification error in the data is 32%, implying that one in three students with four failed subjects ($S^* = 4$) would be misclassified ($S = 3$). Figure 6 presents the histogram of the selection-adjusted failure rate: one can see that this is roughly smooth at the three-failed-subject threshold.

One can also derive estimates of the coefficients of interest under worst-case-scenario selection. To do so, however, it is not sufficient to assign to the best students with three observed failed subjects a number of failed subjects equal to four and examine average outcomes on the two sides of the threshold. To the extent that grade failure negatively affects outcomes, misclassified students' outcomes will be inflated as a result of having been spared the penalty of grade failure, so their outcomes will need to be adjusted to take this circumstance into account.

If the true unobserved (as opposed to the actual) variable Y^* depends on $P^* = I(S^*>3)$ and S^* according to equation (3), $Y^* = \delta_0 + \delta_1 P^* + h(S^* - 3) + e$, under the assumptions in equation (4), the observed outcome variable Y can be expressed as a function of the true number of failed subjects as follows:

$$Y = \delta_0 + \delta_1[P^* - p\, d_4] + h(S^* - 3) + e, \tag{5}$$

FIGURE 7.—ADDITIONAL GRADES ATTENDED BY 2001 (2000 FOR NONFAILERS) BY
NUMBER OF FAILED SUBJECTS IN JUNIOR HIGH SCHOOL, URUGUAY, 1996–1997,
WORST-CASE-SCENARIO SELECTION ESTIMATES
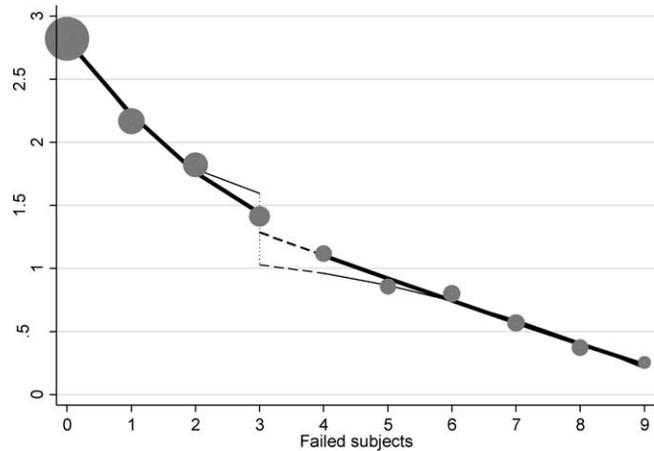
The solid line in the figure reports the estimated number of additional grades attended by 2001 among those in junior high in 1996 or 1997 as a function of the number of latent failed subjects. The thin line reports the same data as in figure 4. See also the note to figure 3 and the text for details.

where p is the probability of misclassification [$Pr(S = 3| S^* = 4)$] that can be estimated on the data and d_4 is an indicator variable for individuals with a latent number of failed subjects equal to 4. The second term in square brackets accounts precisely for the inflated outcomes among reclassified students. Model (5) can be estimated based on an OLS regression of Y on [$P^* - p\, d_4$]. With an estimate of δ_1, one can recover the value of the latent variable $Y^* = Y$ for each misclassified individual. This can be used to derive reduced-form and IV estimates that correct for worst-case-scenario selection.[17]

Figure 7 presents the selection-adjusted outcome variable Y^*, using the procedure just described, as a function of the

[17] Obviously this involves assigning to reclassified students a value for failure $F^* = 1$.

TABLE 4.—GRADE FAILURE AND SUBSEQUENT SCHOOL OUTCOMES IN JUNIOR HIGH SCHOOL, URUGUAY: 1996–1997—IV ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Survival at Time $t =$				Duration	Intermittent Attendance	Additional Grades at Time $t =$			
	1	2	3	4			1	2	3	4
Grade failure $t = 0$	−0.495***	−0.280*	0.064	−0.096	−0.886*	0.028	−0.955***	−0.809***	−0.796***	−0.903***
	(0.154)	(0.158)	(0.150)	(0.129)	(0.473)	(0.212)	(0.053)	(0.167)	(0.244)	(0.307)

Entries are IV estimates of the effect of grade failure at time $t = 0$ on the dependent variable (in the top row) where grade failure is instrumented by the discontinuity at three failed subjects. Columns 1 to 4 refer to the probability of being in the sample in any given year (1, . . ., 4) following grade failure. Column 5 to duration. Column 6 to intermittent attendance. Columns 7 to 10 to the number of additional grades attended by the individuals in the sample in any given year (1,. . .,4) following $t = 0$. All regressions include a fourth-order polynomial in number of failed subjects interacted with a dummy for more than three failed subjects plus dummies for number of missed school days, gender, age-grade distortion, the score in each subject, school, grade, and year fixed effects. Number of observations 99,729 except in column 6 where number of observations is 86,422. See also the note to table 2.

number of true failed subjects S^*, alongside the observed relationship between actual outcomes Y and actual failed subjects S (as in figure 4). Some positively selected students originally at three failed subjects are moved to four failed subjects so the estimated gap becomes smaller. However, even under worst-case-scenario selection, the estimate of the change in attainment at the discontinuity is negative. The point estimate in figure 7, where a quadratic polynomial is fit to the data as in figure 4, is −0.16 (SE 0.05) and the implied IV estimate is −0.24 (SE 0.07) around 75% less than the corresponding estimate in table 2, column 1, but still negative and significant. Results that use higher polynomials lead to similar results, with generally negative and significant point estimates.

In conclusion, although selection might be a source of bias in the estimates presented, this appears to be unable to fully account for the negative effect of grade failure on school progression found above. Some caution is in order here because the assumptions on the selection process in equation (4) are somewhat arbitrary and ultimately untestable.

C. Dynamics

Having ascertained that selection is not a major source of concern for the estimates in table 2, I now investigate in further detail the dynamics of school progression. For this I revert to the data used in table 2, ignoring the selection adjustment and keeping in mind that such estimates might somewhat exaggerate the negative effect of grade failure.

Nothing so far allows us to understand why failers appear to lag behind nonfailers. Is this due to dropout or subsequent grade failure? And if dropout contributes to explain this result, where does this occur? Is this just following grade failure, or instead do grade failers tend to drop out of the system at a higher rate than nonfailers even after a certain number of years? Or is it instead the case that lower educational attainment four to five years down the line is due neither to these students' failing again nor to their dropping out earlier but to the circumstance that grade failers are more likely to temporarily exit the system and then reenter, so that the estimated gap in educational attainment masks a higher probability of intermittent attendance among failers?

In table 4 I report IV estimates for a number of additional outcome variables. Similar to column 5 of table 2, all of the

regressions in the table include the whole set of controls plus grade, year, and school fixed effects. Columns 1 to 4 show the survival probability at time t ($t = 1, 4$). One can see that grade failure is followed by a high dropout rate. Grade failers are on average 50 percentage points less likely to be in school after one year compared to nonfailers. Note that this includes the probability of being in any (nonvocational) school in the public system, not just the school where the student was observed at time $t = 0$. Over time, as passers drop out or end their school cycle, the two distributions tend to converge and, after three years, failers effectively catch up with nonfailers. By year 4, the difference is on the order of −10 percentage points but statistically insignificant.

Column 5 reports the overall duration in the sample. This is a variable that ranges from 0 to 5. On average failers spend about 0.89 fewer years in the sample than nonfailers, suggesting that early dropout, rather than the compounded effect of grade failure (among failers who repeat), explains why grade failers end up with lower educational attainment than nonfailers.

As an additional outcome variable, in column 6 I analyze the effect of grade failure at time $t = 0$ on intermittent attendance. I measure this as the probability of being in the sample at any time between two and five periods after failure conditional on not being in the sample 1 year after. I find no significant evidence of failer's being more likely to attend intermittently than nonfailers; the estimated effect is 0.028 but not statistically significant.

The following columns of the table report information on the number of additional grades attended by failers and nonfailers, whether still in school or not, at any time t ($=1,. . ., 4$) following grade failure ($t = 0$). Because attendance is measured in terms of the highest grade attended (rather than successfully completed), it does not make any difference to the result for the first period if failers drop out or not following grade failure. In either case, maximum grade attended will be the one they failed. Some nonfailers also drop out, though, so the difference in maximum grade attended the year after grade failure will be strictly less than 1. As expected, grade failers have just below a one-year gap compared to the nonfailers at time $t = 1$ (−0.95). After two years, failers partially catch up to nonfailers. The estimated gap is −0.81. This is possibly the combined effect of lower

dropout rates and lower failure rates among those who originally failed at time $t = 0$ and stayed compared to those who passed. After four years, the difference in maximum grades attended is -0.90. This is close to, but slightly lower than, the effect on the censored distribution reported in table 2 (-0.76), implying some additional gain among failers in the last year of observation.

In sum, the data show a persistent disadvantage for grade failers in terms of additional grades that is largely explained by early dropout.

D. Endogenous Mobility

Because of the nature of the data, which refer only to students in public nonvocational secondary schools, I might have erroneously classified students who move to vocational or private schools as dropouts and assigned them zero additional years of education while these students in fact pursue their studies elsewhere. This might potentially lead to downward-biased estimates of the effect of grade failure if failers are more likely than passers to leave *Liceos* for schools outside the system. In practice, though, two pieces of evidence suggest that this is unlikely to be a major source of concern.

First, evidence from a follow-up phone survey of 660 individuals who dropped out of the first year of nonvocational junior high school in 1997 reported in Administración Nacional de Educación Publica (2000) shows that of these, only around 1.6% had moved to a private institution and 15% had moved to a vocational school in 1998. The largest majority of dropouts from junior high effectively fail to enroll in other schools.

As a second check, I have used microdata from the 1999 Uruguayan National Learning Census (Evaluation censal de aprendizajes en terceros años del ciclo medio), which collected information on all students (in both vocational and nonvocational schools and in both private and public institutions) in ninth grade. I have linked these data to the administrative records on students in public nonvocational junior high in 1998 using a unique student identifier. This allows analyzing the destination state of students enrolled in public nonvocational schools in 1998. I restrict to those who either failed ninth grade in 1998 or passed eighth grade in 1998, that is, individuals potentially in ninth grade in 1999. The data show that among dropouts, the largest majority (95%) exits the school system completely. More important, among those who drop out of the public nonvocational system, there is little appreciable difference in the probability of moving to a private or a vocational school between failers (6%) and nonfailers (3%). This margin of endogenous selection is hence unlikely to affect my conclusions.

VI Conclusion

This paper uses administrative longitudinal microdata on students enrolled in public nonvocational junior high school in Uruguay between 1996 and 1997 to assess the cost of grade failure as measured by its effect on students' subsequent school outcomes. Exploiting the discontinuity in promotion induced by a rule that establishes that a pupil failing more than three subjects will automatically fail that grade, I show that grade failure leads to dropout and lower educational attainment four to five years after failure on the order of -0.8 school years. When I account for the potential nonrandom selection of students around the discontinuity threshold using reasonable (although untestable) assumptions, I find estimates for the effect of grade failure that are negative and on the order of -0.2 school years. If anything, these are conservative estimates of the effect of grade failure on outcomes as these are derived under the assumption of perfect positive selection.

Although this paper concentrates on the costs of grade failure, it must be emphasized that the benefits of this policy due to its deterrence effect on students' underperformance might be nonnegligible. This is probably the ultimate reason that repetition policies are in place. Simple back-of-the-envelope calculations show that for such an incentive effect to offset the cost of repetition, one would expect 22% of individuals who do not incur the penalty to accumulate one extra year of schooling due to the threat effect of the rule.[18] In practice, it appears that this policy should have high incentive effects to compensate for what I estimate being its high costs.

Precisely because of this trade-off, in the United States, the emphasis now seems to have shifted toward policies that combine grade retention—to preserve the incentive effect—with remedial interventions—to attenuate the negative consequences of repetition and potentially make failure less likely in the future (for the experience of the Chicago Public School, a front-runner in implementing these policies, see Roderick et al., 1999). Recent experimental evidence shows substantial gains from informal inexpensive remedial education among more disadvantaged children in India (Banerjee et al., 2007), suggesting that even in developing countries, this might be a viable alternative to repetition.

Because of this, many Latin American countries, including Uruguay, have, especially in the past decade, introduced compensatory education policies. These include a variety of measures: from conditional cash transfer programs—such as Progresa/Oportunidades in Mexico or Bolsa Escola/Bolsa Familia in Brazil, (for all, see Fiszbein, Schady, & Ferreira, 2009), to resource equalization across schools, such as the P-900 program in Chile (see Chay, McEwan, & Urquiola, 2005), FUNDEF in Brazil (see Menezes-Filho & Pazello, 2004), and CONAFE in Mexico (see Shapiro & Moreno-Trevino, 2004) and early childhood interventions. For Uruguay, in particular, there is evidence suggesting that the recent universalization of pre-

---

[18] This is the effect of grade failure on additional grades, -0.76 (from table 2, column 5, bottom row) times the proportion of failers, 26%, from table 1.

schooling might have led to large gains in terms of lower grade repetition and lower school dropout (Berlinski, Galiani, & Manacorda, 2008).

REFERENCES

Alexander, Karl L., Doris Entwisle, and Nader S. Kabbani, "Grade Retention, Social Promotion, and 'Third Way' Alternatives," (pp. 197–238), in Arthur J. Reynolds, Margaret C. Wang, and Herbert J. Walberg (Eds.), *Early Childhood Programs for a New Century* (Washington, DC: Child Welfare League of America Press, 2003).
Administración Nacional de Educación Publica, *Circular 2306*, Montevideo (1996).
——— *Una visión integral del Proceso de Reforma Educativa en Uruguay, 1995–1999*, Montevideo (2000).
Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer, "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review* 2 (2002), 1535–1558.
Angrist, Joshua, and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114 (1999), 533–575.
Banerjee, Abhijit V., and Esther Duflo, "Addressing Absence," *Journal of Economic Perspectives* 20 (2006), 117–132.
Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden, "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics* 122 (2007), 1235–1264.
Barro, Robert J., and Jong-Wha Lee, "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers* 3 (2001), 541–563.
Behrman, Jere, Piyali Sengupta, and Petra Todd, "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment," *Economic Development and Cultural Change* 54 (2005), 237–275.
Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda, "Giving Children a Better Start: Preschool Attendance and School-Age Profiles," *Journal of Public Economics* 92 (2006), 1416–1440.
Bucheli, Marisa, and Carlos Casacuberta, "Asistencia escolar y participación en el mercado de trabajo de los adolescentes en Uruguay," *El Trimestre Económico* 67 (2000), 395–420.
Cascio, Elizabeth U., "School Progression and the Grade Distribution of Students: Evidence from the Current Population Survey," IZA discussion papers no. 1747 (2005).
Chaudhury, Nazmul, Michael Kremer, Jeffrey Hammer, Karthik Muralidharan, and F. Halsey Rogers, "Teacher Absence in India: A Snapshot," *Journal of the European Economic Association* 3 (2005), 658–667.
Chay, Kenneth Y., Patrick J. McEwan, and Miguel Urquiola, "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools," *American Economic Review* 95 (2005), 1237–1258.
Consejo de Educación Técnico Professional, *Ciclo Básico Tecnológico* (Montevideo, 2005).
Currie, Janet, and Matthew Neidell, "Getting inside the 'Black Box' of Head Start Quality: What Matters and What Doesn't," *Economics of Education Review* 26:1 (2007), 83–99.
Da Silveira, Pablo, and Rosario Queirolo, "Son nuestras escuelas y Liceos capaces de enseñar?" CERES working papers no. 7 (1998).
Duflo, Esther, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review* 91 (2001), 795–813.
Eide, Eric R., and Mark H. Showalter, "The Effect of Grade Retention on Educational and Labor Market Outcomes," *Economics of Education Review* 20 (2001), 563–576.
Fiszbein, Ariel, Norbert Schady, and Francisco H. G. Ferreira, *Conditional Cash Transfers* (Washington, DC: World Bank, 2009).
Furtado, Magdalena, "Trayectorias Educativas de los Jóvenes: el problema de la deserción," Cuaderno de trabajo TEMS no. 22 (2003).

Gomes-Neto, Joao Batista, and Eric A. Hanushek, "Causes and Consequences of Grade Repetition: Evidence from Brazil," *Economic Development and Cultural Change* 43:1 (1994), 117–148.
Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69:1 (2001), 201–209.
Jacob, Brian A., "Accountability, Incentives and Behavior: Evidence from School Reform in Chicago," *Journal of Public Economics* 89 (2005), 761–796.
Jacob, Brian A., and Lars Lefgren, "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis," this REVIEW 86 (2004), 226–244.
——— "The Effect of Grade Retention on High School Completion," *American Economic Journal: Applied Economics* 1 (2009), 33–58.
Jacob, Brian A., and Steven D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics* 118 (2003), 843–878.
Jimerson, Shane, Gabrielle E. Anderson, and Angela D. Whipple, "Winning the Battle and Losing the War: Examining the Relation between Grade Retention and Dropping Out of High School," *Psychology in the Schools* 39 (2002), 441–457.
Lee, David, "Randomized Experiments from Non-Random Selection in U.S.A. House Elections," *Journal of Econometrics* 142 (2008), 675–697.
Lee, David, and David Card, "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics* 142 (2008), 655–674.
Maurin, Eric, and Dominique Goux, "The Impact of Overcrowded Housing on Children's Performance at School," *Journal of Public Economics* 89 (2005), 797–819.
McCoy, Ann R., and Arthur J. Reynolds, "Grade Retention and School Performance: An Extended Investigation," *Journal of School Psychology* 37 (1999), 273–298.
McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design," *Journal of Econometrics* 142 (2008), 698–714.
McEwan, Patrick, and Miguel Urquiola, Precise Sorting around Cutoffs in the Regression-Discontinuity Design: Evidence from Class Size Reduction, mimeograph, Columbia University (2005).
Menezes-Filho, Naércio, and Elaine Pazello, "Evaluating the Effects of FUNDEF on Wages and Test Scores in Brazil," mimeograph, University of São Paulo, Brazil (2004).
Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens, "Does Human Capital Transfer from Parent to Child? The Intergenerational Effects of Compulsory Schooling," *Journal of Labor Economics* 24 (2006), 729–760.
Patrinos, Harry Anthony, and George Psacharopoulos, "Socioeconomic and Ethnic Determinants of Age-Grade Distortion in Bolivian and Guatemalan Primary Schools," *International Journal of Educational Development* 16 (1996), 3–14.
Pischke, Jorn-Steffen, "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years," *Economic Journal* 117 (2007), 1216–1242.
Roderick, Melissa, Anthony S. Bryk, Brian A. Jacob, John O. Easton, and Elaine Allensworth, *Interpretive Summary: Ending Social Promotion, Results from the First Two Years* (Chicago: Consortium on Chicago School Research, Chicago, 1999).
Schady, Norbert, and Maria Caridad Araujo, "Cash Transfers, Conditions, and School Enrollment in Ecuador," *Economia* 8:2 (2008), 53–77.
Shapiro Joseph, and Jorge O. Moreno-Trevino, "Compensatory Education for Disadvantaged Mexican Students: An Impact Evaluation Using Propensity Score Matching," World Bank Policy Research working papers no. 3334 (2004).
UNESCO, *EFA Global Monitoring Report* (2002), at http://portal.unesco.org.
Urquiola, Miguel, "Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia," this REVIEW 88 (2006), 171–177.
Urquiola, Miguel, and Valentina Calderon, 2006, "Apples and Oranges: Educational Enrolment and Attainment across Countries in Latin America and the Caribbean," *International Journal of Educational Development* 26 (2006), 572–590.
Van der Klaauw, W., "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach?" *International Economic Review* 43 (2002), 1249–1287.