Social Promotion in Primary School: Effects on Grade Progression⋆

Margaret Leighton*
Priscila Souza**
Stéphane Straub***

Abstract
This paper evaluates the effect of relaxing promotion criteria in early primary school on grade delay in later years. Exploiting variation in primary school repetition policies across Brazilian municipalities, we find that social promotion in junior primary years reduces grade delay, and that some of this reduction persists through the transition to senior primary school. Twelve-year-old students who have been exposed to the social promotion policy since they were seven have almost 5 percentage points lower chance of being delayed a year or more in their studies than do students who faced the threat of retention every year. We also find that, when the option is available, students sort across schools in response to the policy in a way consistent with negative selection into social promotion.

Keywords: Grade repetition, primary education, remedial education, social promotion, grading thresholds
JEL Codes: I21, I28, I25

1 Introduction

Grade repetition is a costly remedial education policy, with a weak record of efficiency; nevertheless, primary school repetition rates in many developing countries are in the double digits.[1] Such high retention rates often result in

Submitted on 08 March 2019; Reviewed on 30 April 2019

⋆ We thank Matteo Bobba, Liliane Bonnal, Sylvain Chabé-Ferret, Sergio Firpo, Mattia Girotti, Julien Grenet, Mark Schaffer, anonymous referees and seminar participants at the University of Dundee, Heriot–Watt University, the University of St Andrews and the International Workshop on Applied Economics of Education for valuable insights and discussions throughout this project. We also thank Julia Bird for contributions to the dataset. We gratefully acknowledge the *Instituto Nacional de Estudos e Pesquisas Educacionais* for sharing restricted-access data necessary to complete our panel. A version of this paper was previously circulated under the title *Social Promotion in Primary School: Immediate and Cumulated Effects on Attainment*. Stéphane Straub acknowledges funding from ANR under grant ANR-17-EUR-0010 (Investissements d'Avenir Program).

* School of Economics and Finance, University of St Andrews, Castlecliffe, The Scores, KY169AR, St Andrews, UK.

** Climate Policy Initiative, Pontifical Catholic University of Rio de Janeiro (PUC-Rio).

*** Toulouse School of Economics, University of Toulouse Capitole.

✉ mal22@st-andrews.ac.uk

[1] Based on the 2009 PISA questionnaire, Ikeda and García (2013) identify that over 16% of Colombian 15-year-olds, 12% of Indonesian and 25% of similarly-aged students in Trinidad and Tobago self-report repeating at least one grade in primary school.

students repeating the same early primary grade multiple times. But would struggling students in low-resource education systems do any better if they were automatically promoted from grade-to-grade?

This paper evaluates the effect of spells of social promotion on the grade progression of Brazilian primary school students. Using administrative data on student enrolments from Brazil's annual school census, we explore how a policy which periodically dropped the academic promotion threshold affected student progress through primary school. Ultimately, we want to know whether some of those students who would otherwise have been held back, but due to the policy are instead promoted, are able to catch up with their peers and maintain the accelerated grade level conferred mechanically by the policy.

Our analysis relies on a subset of schools, located in municipalities with a uniform promotion policy, whose policy changes at least once between 1999 and 2006. We find that, while easing repetition rates may confer a purely mechanical boost to progress initially, some of this effect does persist through an important gateway year: the transition from junior to senior primary school. By age 12, when all students should have advanced to senior primary school, cohorts exposed to years of social promotion have almost 5 percentage points fewer students still enrolled in junior primary grades. Similar results do not hold in a larger sample including the universe of public schools: in this case, there is evidence of students moving away from schools with social promotion at older ages, potentially confounding estimates of cohort-level academic attainment.

Despite the prevalence of such policies around the world, the evidence on the effect of repetition policies is both limited and mixed. Historically, this literature has been plagued by selection issues; while those are increasingly being addressed, changes in repetition policies are often accompanied by other policy innovations, making it difficult to study in isolation. Much of the modern literature on grade repetition evaluates changes in test scores. Some studies find increases in scores following the introduction of high-stakes testing (Greene & Winters, 2007, Jacob & Lefgren, 2004); however, there is evidence that these effects fade over time (Allen, Chen, Willson, & Hughes, 2009, Schwerdt, West, & Winters, 2017). Other authors find negative net effects on performance (Roderick & Nagaoka, 2005, Fruehwirth, Navarro, & Takahashi, 2016).

Few papers have considered the dynamic effects of grade repetition – either in terms of differential impact by age, or the persistence of a given impact as children grow older. Fruehwirth et al. (2016) are a notable exception, estimating a structural model of repetition from kindergarten to age 11 in the US. The authors find that retained students have significantly lower test scores, but that early repetition allows students more time to catch up. Students retained in kindergarten would have scored up to 27% higher in the following year, had they not been retained; overall, students retained once by age 11 learn 7% less than they would have in the absence of repetition.

A smaller body of work has looked at the effect of grade repetition on dropout. Allensworth (2005) concludes that dropout increases among retained grade eight students in Chicago, but decreases for the larger group of non-retained students. Using data from Uruguay, Manacorda (2012) demonstrates that retention in secondary junior high school increases dropout rates and reduces educational attainment by 0.2–0.8 years. André (2009) and Glick and Sahn (2010) find, in a panel survey of primary school students in Senegal, that repetition increases the risk of dropping out.

A number of studies have exploited Brazil's policy of periodic social promotion to estimate the effect of grade repetition on student outcomes. Koppensteiner (2014) finds that the move to social promotion in second and fourth grades in Minas Gerais decreased test scores of 4th graders in state-run schools. In a study of urban state schools across the country, Menezes-Filho, Vasconcellos, Werlang, and Biondi (2008) find that schools with the policy had lower dropout rates, particularly for eighth graders, and that test scores were lower for eighth graders, but not for fourth graders. Carvalho and Firpo (2014), in contrast, find that the policy had no significant effect on the test scores of students in grades four and eight, across achievement quantiles.

This paper contributes to these literatures in three ways. First, we show that, for one aggregate indicator of student progress, Brazil's high rates of repetition slow down students who have the potential to progress steady through early primary school. While our analysis does not feature test scores—importantly, this means we cannot detect decreases in performance "at the top"—we are able to study a group of students who are often difficult to reach: those who fall behind. We find that a non-trivial number of students are able to maintain the grade advantage conferred by social promotion through the transition to senior primary school.

Second, by accounting for both current and past promotion policies, we are able to demonstrate the effect of social promotion on the intensive margin. We find that exposure to social promotion over several years has a cumulative effect on progression, and that the reduction in grade delay of each year of exposure is of a similar magnitude.

Finally, we present evidence of negative selection into social promotion. When we expand our analysis to include schools in municipalities where some public schools use social promotion while others do not, we find both that students sort away from schools with social promotion at older ages, and that those cohorts who remain in social promotion schools show no improvement in age-grade matching. This finding echoes Dong (2010), who finds significant positive selection into schools with repetition in kindergarten in the US. Failing to account for this behaviour leads to strikingly different conclusions.

The remainder of paper proceeds as follows. In section 2 we provide background information on the education context of Brazilian primary schools, and describe the social promotion policy in more detail. Section 3 describes the data, and section 4 presents our empirical strategy. Section 5 presents and discuses our main results, while section 6 concludes. Additional details on the data and the policy, as well as robustness exercises around the primary empirical specifications, are in the Appendices.

2 Education in Brazil and the Cycles Policy

The Brazilian basic education system (*ensino básico* in Portuguese) is mandatory, and is offered both by free public schools and fee-charging private schools. Public school administration is shared between municipalities (responsible for most of primary education), states (primarily responsible for secondary education) and the federal government (responsible for all tertiary education). Prior to 2006, children entered primary school at age 7, completing 4 years of junior primary school and 4 years of senior primary school before progressing to secondary school.[2]

Although historically an important issue in Brazil, the number of school-age children out of school had declined rapidly over the past two decades. While in 1992 13.4% of children between the ages of seven and fourteen were out of school, this number fell to 3.5% in 2001, 2.3% in 2007 and 1.5% in 2013 (PNAD/IBGE).[3]

[2] Since 2006, there are two successive stages, the first one from "ano" 1 to 5 (ensino fundamental 1) and the second one from "ano" 6 to 9 (ensino fundamental 2). This reform extended the first stage from 8 to 9 years by essentially advancing primary school enrolment by one year in stage 1. Previously, for most of our panel years, there were only 4 years, with children starting school at age 7. To avoid inconsistency, throughout this study we use the word 'grade' as an analog to the Brazilian term *serie*, which corresponds to levels 1–8 as they existed before 2006. We abstract from any differences in the two systems beyond their duration, and when necessary convert the *ano* grades (1–9) to their *serie* equivalent, where *ano* 2 = *serie* 1. The grade *ano* 1 is excluded from the analysis.

[3] PNAD is the National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios) from the Brazilian Institute of Geography and Statistics (IBGE – Instituto Brasileiro de

Nationwide a smaller, but still impressive, decline in the out-of-school rate can be seen among older students compared to primary-school-aged children. For those between the ages of fifteen and seventeen, the out-of-school rate fell from 40.3% in 1992 to 15.8% in 2013 (PNAD/IBGE).

Until the early nineties, Brazilian schools followed the practice of allowing the retention of poor-performing students at every grade level: students could not only repeat in every grade, but could also be retained multiple times at the same level. The resulting high rates of repetition and grade delay have attracted interest for decades: using microdata collected by the World Bank from 1981–1985 in Northeast Brazil, Gomes-Neto and Hanushek (1994) find that retained students do in fact increase their knowledge year-on-year, but suggest that repetition is a costly way of achieving these small gains.

Starting in 1997, a number of Brazilian municipalities and states adopted a system of 'learning cycles': groupings of school grades between which promotion is automatic.[4] For example, if the first 8 years of primary school are grouped into 2 cycles of 4 years, then students will pass 1st, 2nd and 3rd grade automatically (subject to a minimum attendance rate), but may repeat the 4th grade. They will also pass the 5th, 6th and 7th grade, but may be retained in the 8th grade. This policy is officially described as one of 'continued progression', (*Progressão Continuada* in Portuguese), but is referred to more popularly as the Cycles policy. Although municipal schools in São Paulo had experimented with learning cycles as early as 1992, the first large-scale adoption of the policy was by the state of São Paulo in 1997.

Adoption of the policy, which was initially unpopular among teachers and parents, took a number of different forms. While in some states and municipalities the adoption of cycles was mandatory, in others the system of cycles was only recommended, and in still others the policy was not taken up at all. Therefore, in some states and municipalities schools could choose whether to adopt the social promotion system. While in many states there was no take-up of the policy, schools in other states adopted it for some years, and then sometimes retracted it.

Geografia e Estatística). The survey collects annual data on the characteristics of the population with a sample size of over 150,000 households.

[4] This policy of learning cycles was nationally recognised in the Law of Guidelines and Foundations for Education (*Lei de Diretrizes e Bases da Educação*, LDB), enacted by the Federal Government in 1996. This law granted additional autonomy to municipalities and states to organise the schooling system.

3 Data

3.1 Overview

The data which provides the starting point of this paper is the *Censo Escolar*, an annual census of schools in Brazil below the tertiary level. The survey, carried out in May, covers both private and public schools, and has been running continuously since 1995. The data are publicly available from the *Instituto Nacional de Estudos e Pesquisas Educacionais* (INEP, the national education research institute). From 1995 to 2006 the *Censo Escolar* measured school-level variables, whereas from 2007 on data is presented at the student level, with associated school and teacher files.[5] The *Censo Escolar* survey varies from year to year; however, the general topics remain fairly consistent over time. The survey sections include basic information, physical and instructional features of the school, teachers and staff, numbers of classes and students, and student flows from the previous year. For the purposes of our study, we merge the data from the *Censo Escolar* from 1999, the first year in which the survey included a question about cycles, to 2006, after which data collection moved from the school level to the individual level.

To give a sense of magnitude, in 2004 the *Censo Escolar* surveyed 248,257 schools. We restrict our attention to students aged 7–12, and to the schools in which they are enrolled. Excluding schools which teach only secondary school or pre-school leaves us with 166,505 schools active in that year. Of these, 116,209 are under municipal jurisdiction, 31,178 are under state jurisdiction, 19,078 are private schools, and the remaining 40 are federally-run. Descriptive statistics on schools can be found in Appendix A. The 1999–2006 school panel is highly unbalanced, due to schools opening, closing or registering as inactive. A total of 216,429 primary schools appear at least once in our 8-year panel; 14,227 are only active in a single year, while 129,942 schools—less than half—are present throughout.[6]

We augment the *Censo Escolar* with census data on municipal population and gross domestic product from the *Instituto de Pesquisa Economica Applicada* (IPEA). Additional data regarding ages of children surveyed in the 2000 and

[5] School identifiers, as well as student-level identifiers, are encrypted in the publicly-available data. This encryption prevents the identification of individual schools, but also the linkage of schools across years. The *Instituto Nacional de Estudos e Pesquisas Educacionais* provided us with necessary codes to link schools year-to-year.

[6] To relate this number to the example given above, this means that 129,942 of the 166,505 schools active in 2004 were open throughout our panel, while the remainder were open for some part of the 8 years under study.

2010 censuses were acquired from the *Sistema IBGE de Recuperação Automática* (SIDRA).[7]

3.2 Descriptive statistics

The 1999–2006 panel of schools includes 1,342,592 school-year observations covering 216,429 individual primary schools. In any given year, about 30,000 of these schools (approximately 18.5%) report using the cycles policy. Data on adoption of social promotion cycles are available in the *Censo Escolar* in two forms.[8] In 99, and again from 2003 to 2006, the data include individual schools' reported "total number of cycles and duration of each cycle." While the questionnaires from 2000–2002 include a similar cycles module, the data are absent from the publicly available data files. Cycles data from these years were provided to us on request; however, these supplementary data are only yes/no. For the 1999–2006 panel, therefore, we have a consistent binary measure of cycles adoption, but no consistent details on duration or timing of these cycles: we therefore restrict ourselves to a binary adoption variable. Menezes-Filho et al. (2008) use a similar approach.

While we do not have data on the duration and number of cycles for all years of our panel, we do have this information for some years. Tables 1 and 2 list raw descriptive statistics on the number of cycles and the duration of the first cycle, for all schools giving a valid answer to these questions in the 1999 *Censo Escolar*.

While we present some results using the full panel of schools, our identification strategy relies on municipalities which follow a unified policy: either all schools using a policy of cycles, or no schools using such a policy. Each year, a little over 4,000 schools (4.5%) are in fully-adopting municipalities. While this figure is quite consistent year-to-year, this masks considerable movement in and out of the policy by schools and municipalities.

Figure 1 displays the geographic distribution of municipalities in which public schools adopt the policy universally for at least one year between 1999 and 2006. Consistent with the overall adoption pattern of the policy, none of the

[7] Data come from the section on education, accessed through http://www.sidra.ibge.gov.br/bda/popul

[8] Over the course of our panel, there are approximately 30,000 school-year observations which return a missing policy. A few examples from the data lead us to believe these are either clerical errors or misunderstandings, and are meant to indicate absence of cycles. One such example is the state of Minas Gerais where, between 2002 and 2003, the number of schools responding to the question falls by 61%, while the share of schools adopting cycles rises from 41% to 99%. This situation persists in 2004, before reverting to pattern much more similar to that observed in 2002.

Table 1. Number of cycles.

	Number	Percent
0	99	0
1	14,641	45
2	14,778	46
3	966	3
4	1,819	6
8	2	0
14	1	0
Total	32,306	100

Source: Censo Escolar 1999, raw frequencies. Observations are at the school level.

Table 2. Length of first cycle.

	Number	Percent
0	60	0
1	3,933	13
2	8,385	28
3	4,640	15
4	12,938	43
5	58	0
6	1	0
8	12	0
Total	30,027	100

Source: Censo Escolar 1999, raw frequencies. Observations are at the school level.

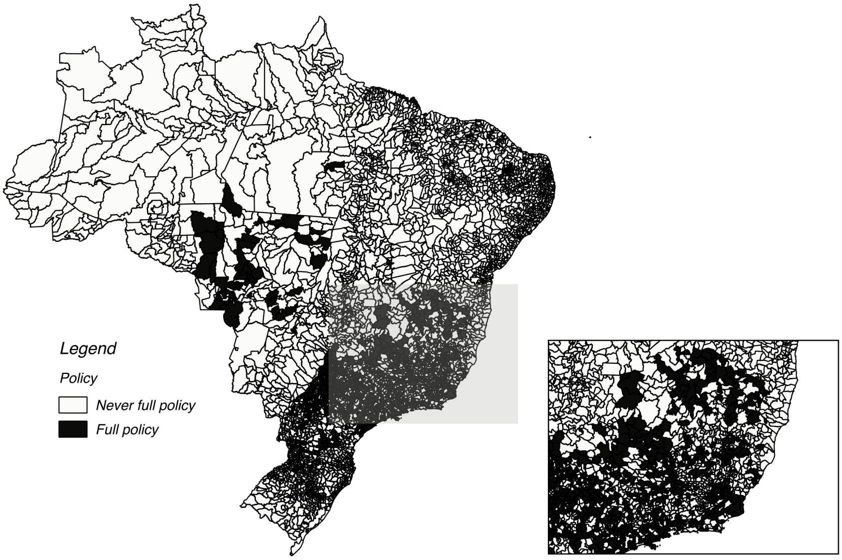
municipalities are located in the North or in the South. While there is some presence in the Northeast, the majority of the policy-changing municipalities are located in the Southeast and the Centre-West.[9]

Our primary interest is in estimating the causal impact of the social promotion cycles on age-grade delay; we will also comment on how the policy affect enrolment and actual repetition rates. Table 3 presents school-level average enrolments for students in grades 1–4, along with repetition and dropout rates.[10] The average school in our sample is fairly small, enrolling a little over 30 students in first grade, a figure which drops to 25 by fourth grade. We calculate student flows, which are reported in the following year's survey, by summing the number of students in each grade who passed, repeated and dropped out during the year, and then calculating the percentage of student in each category. Repetition rates are high, with an average rate of 19% in first grade, declining to 9% by fourth grade. Dropout rates are also substantial, at 11% in first grade and 8% thereafter. It should be noted that dropout rates measured in the school census are of a temporary nature, meaning only that the student did not complete the school year: many of these dropouts will enrol again the following year.

As the school-level data do not allow us to identify individual students, we track student progress through the school system using birth cohort-level grade delay. To generate this measure, we must decide how to define delay. The school census is carried out in May, but the school year runs from February

[9] Further details on the adoption of the policy can be found in Appendix B.

[10] In our empirical work, we follow Glewwe and Kassouf (2012) by using the natural logarithm of student numbers as our primary measure of enrolment. We present summary statistics in levels for ease of interpretation; summary statistics in logs can be found in the data appendix.



Source: Authors' calculations, using municipality shapefile from IBGE and policy data from the *Censo Escolar*.

Figure 1. Location of municipalities (in black) which adopt cycles universally among public schools for at least one year.

Table 3. Average characteristics across schools by grade.

Grade	Enrolment count	Repetition rate	Dropout rate
Grade 1	32.74 (48.32)	19% (21%)	11% (16%)
Grade 2	28.01 (42.42)	15% (18%)	8% (15%)
Grade 3	26.49 (41.18)	11% (16%)	8% (16%)
Grade 4	25.42 (41.35)	9% (15%)	8% (16%)

Notes: School-level means (standard deviation in parenthesis). Sample size varies, as student enrolment data are from the current year, while flow data are collected from previous year—meaning that only schools which are observed in two consecutive survey years are retained.

Source: Censo Escolar, authors' calculations.

to December: classifying all students who are a year ahead of their target age for grade as delayed would be an overestimate of grade delay.[11] We choose to be conservative, and classify only those students who are two years behind the target age as delayed (e.g. a nine year old who is in first grade). Table 4 presents rates of grade delay by age, averaged at the school level. Delay is substantial and begins early: 26% of nine year olds are already delayed (meaning they are enrolled in grade 1), and by age 12 this has risen to 72%. The numbers are even more striking when considering the group that is possibly delayed: e.g. one or more years older than the target.

Table 4. Average share of age-group delayed in their studies.

Age	Possibly delayed	Delayed	N
Age 8	49% (34%)	–	1,208,865
Age 9	61% (33%)	26% (32%)	1,221,617
Age 10	66% (34%)	37% (36%)	1,253,106
Age 11	80% (34%)	49% (38%)	1,251,465
Age 12	85% (26%)	72% (40%)	1,179,187

Notes: School-level means (standard deviation in parenthesis). Sample size varies, as student enrolment data are from the current year, while flow data are collected from previous year—meaning that only schools which are observed in two consecutive survey years are retained.

Source: Censo Escolar, authors' calculations.

4 Estimation Strategy

4.1 Sample

As described in section 2 above, the policy was not randomly assigned to schools: in some areas, schools faced little choice in the matter, while in others the policy was discretionary. A naive regression of policy status on student outcomes would therefore face two selection issues: first, that schools select into the policy based on characteristics related to our outcomes of interest, and second, that individual schools' policy status affects parents' decisions on where to send their children, in ways that are correlated with student characteristics.

Our approach to minimising these two sources of bias is to restrict our analysis to schools in municipalities with a uniform promotion policy across public schools. In such municipalities, it is less likely that schools were able to

[11] An eight-year-old student in grade one, where the target age is seven, would not be a year behind in her studies if she recently turned eight - she would if she is about to turn nine.

opt in or out of the policy at will, and more likely that the policy was imposed top-down. This prevents individual schools from opting into social promotion in response to a particular cohort of students: for example, a large cohort which puts strain on grade one resources. This does not protect against entire municipalities reacting to features of the student body. To mitigate this we control for time-variant municipality characteristics such as population and GDP, as well as estimated number of children in each age group.[12]

This sample restriction also addresses the second, and arguably more severe, selection issue: in municipalities with a uniform promotion policy, parents have no incentive to sort children across schools in response to the policy, unless they move them into private schools.[13] Anecdotal evidence that parents felt that social promotion reduced children's incentives for learning make this a real concern. In US data, Dong (2010) finds significant positive selection into schools with repetition in kindergarten: failure to control for this can lead strict repetition policies to be spuriously correlated with better academic outcomes.

The cost of this restriction comes both in a reduction in the sample, and in the non-representativeness of the remaining schools. Table B.1 in the Appendix B compares average school characteristics overall and for our sample. Compared to the average school, those in our sample are more likely to be in a rural location and be municipality-run, and are located in municipalities with lower GDP per capita and a higher ratio of schools to population.

4.2 Treatment variable and identifying assumptions

4.2.1 Treatment

Throughout our analysis, we define treatment as a binary variable, equal to one when all state or municipality-run schools teaching grades one-four report they they use cycles of social promotion. This variable is equal to zero when no schools in the municipality report using cycles, and is left as missing when some intermediate fraction of schools use social promotion.

Analysis, in contrast, is done at the school level: the treatment variable is carried over from the municipality to all the municipal- or state-run public schools within it. For schools in municipalities with a uniform policy, this treatment variable is the same as reported in the school census: whether or not

[12] Number of children in each municipality are reported, in tranches, in the 2000 and 2010 censuses. We use these data to predict cohort sizes by age at each intermediate year.

[13] We replicate out main results in a more restricted sample that excludes municipalities in which private schools have a different promotion policy from public schools. See Table C.5 in Appendix C.

social promotion cycles are used. For schools in municipalities with a mix of cycle-using and cycle not-using schools, treatment is coded as missing: these schools are not included in the analysis.[14]

Data on the number and duration of social promotion cycles confirms that the policy was implemented with great heterogeneity (see Tables 1 and 2). While it would be desirable to calculate the timing of cycles in each school, the data is only available in some survey years, and appears quite noisy. Our treatment variable, therefore, should not be interpreted as implying social promotion in any particular year, but rather lowering the risk of repetition for children in early primary school.[15] We refer to this as "exposure to the policy," acknowledging that the impact of such exposure will be more significant in some grades than in others.

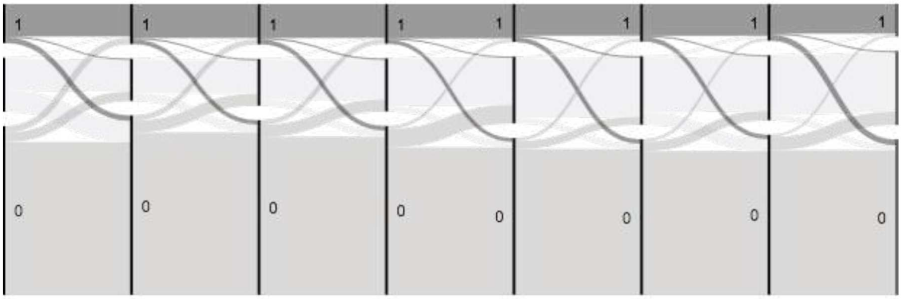
4.2.2 Identification

Identification in our regressions comes from variation in promotion policies over time within the same school. Outcomes are averaged across children of a given age (or at a given grade level) within each school, and compared over time. For example, when regressing our treatment variable on the share of ten-year-old children delayed for their age, this statistic is compared across different years, within a school. Treatment is equal to one in those years in which the municipality has uniformly adopted the policy and equal to zero when the municipality has uniformly maintained a traditional repetition policy (years in which the municipality does not have a uniform policy are dropped). Identification comes from comparing treated cohorts with "untreated" cohorts: schools who never vary their repetition policy are therefore not included in the analysis.

Figure 2 shows the flow of 216,429 schools in (1) and out (0) of treatment over the course of our panel. The majority of schools are in municipalities which never adopt social promotion cycles; nevertheless, there is movement in and out of the policy in every period. In total, 44,255 schools have at least one treated and one untreated observation: these are the schools with the necessary variation to be retained in our analysis. Note however that, in those regressions where a number of lagged values of the policy are included, the sample restrictions is more severe: schools with missing policy variables, as well as those that opened

[14] This is done in a year-by-year basis: schools in municipalities with a uniform policy for four years will be in our panel for those four years, even if they are later dropped due to policy heterogeneity.

[15] The exception to this is first grade, which is included in (almost) all cycles (Table 2).



Source: Authors' calculations, using policy data from the *Censo Escolar*, plotted using RAWGraphs, http://app.rawgraphs.io

Figure 2. Flow of schools in and out of treatment.
Flow of schools into (1; dark grey) and out of (0; grey) treatment from 1999 (first panel) to 2006 (last panel). Schools with missing treatment status, or who are closed for part of the panel, fall into the centre (pale grey). Data from 216,429 schools which offer junior primary grades.

or closed during the panel, are more likely to drop out as they do not have the necessary panel length.

Our intention with these sample restrictions is to minimise the risk that schools' decision to adopt the policy is systematically related to the outcome of interest in ways not accounted for in our empirical model. Our focus on schools who follow the municipal policy should protect us to some degree against reverse-causality from school-level characteristics. For example, supposing that social promotion is indeed effective at reducing age-grade gaps, an effective principal wishing to target this outcome might change the policy, while simultaneously implementing other beneficial reforms.

Identification of a causal parameter also requires that similar factors are not at play at the municipal level: a concern that our sampling strategy does not address. We argue, however, that such reverse-causality is likely to be minimal. The policy itself was quite controversial when it was first implemented, and it remains far from obvious what impacts it has had on educational outcomes. At the time, it would have been difficult or impossible to predict these effects accurately. Furthermore, the initial intent of the policy seems to have been primarily to move students out of first grade, where numbers were large, rather than to achieve any particular learning gains. Finally, we are able to control for some time-varying municipal characteristics which could be related to the policy, including population size at different ages and municipal GDP.

Identification of the effect of the policy requires several further assumptions.

We do not estimate separately the adoption of social policy from the adoption of annual repetition: in other words, we assume that the effect of adopting social policy cycles is symmetric to moving away from social promotion. For repetition rates themselves, this is likely to be quite incorrect: reverting from social promotion ought to induce a spike in repetition in those cohorts previously exposed to it. For our main outcome of interest, however—age-grade delays through the gateway year—the question is whether automatic promotion of struggling students at *any point* helps or hinders their progress through a high stakes exam.

Policy parameters estimated at different ages can only be interpreted as the causal effect of reducing retention rates at that age if we are willing to assume the policy affected all ages equally. This is unlikely to be the case: most schools with social promotion include first grade in a cycle, but many schools allow retention in second grade. Second grade students at that school are nonetheless "treated" with social promotion, according to our definition of treatment. We therefore maintain a more general interpretation of treatment, and do not draw conclusions on the differential impact of retention at younger or older ages.

4.3 Estimating Equation

If the policy affected only current-year outcomes, we would estimate regressions of the following form:

$$Y_{it}^{\tau} = \alpha + \beta^{\tau} D_{it} + \mathbf{X}_{it}^{\prime}\gamma^{\tau} + \mathbf{Z}_{t}^{\prime}\delta^{\tau} + \theta_i + \epsilon_{it}, \qquad (1)$$

where Y_{it}^{τ} is the outcome of interest in school i at time t for students τ (with τ representing either the grade-level or age group, as appropriate); D_{it} is a dummy for the policy; $\mathbf{X}_{it}$ is a vector of time-varying characteristics at the school level; $\mathbf{Z}_t$ a vector of time-varying municipal characteristics; θ_i is a school fixed effect and ϵ_{it} is the unobserved error term. Y_{it}^{τ} will be, for instance, the repetition rate of 3rd grade students in school i. In each case we are interested in estimating the coefficient β: the effect of exposure to the policy in the current year on the outcome under consideration.

While there is movement in and out of the policy every year, the adoption of cycles is highly persistent. Students who have been enrolled in the same school for several years therefore are likely to have been treated with the same policy for several years in a row. If we fail to control for this past exposure we run the risk of overestimating the effect of the policy in the current year: students exposed today were likely also exposed last year.

We therefore estimate an analog of equation (1) which includes lagged values of the policy. This approach can be summarised in the following equation, where $\mathbf{D}_i$ is a vector containing both current and past policy dummies:

$$Y_{it}^{\tau} = \alpha + \mathbf{D}_i'\beta^{\tau} + \mathbf{X}_{it}'\gamma^{\tau} + \mathbf{Z}_t'\delta^{\tau} + \theta_i + \epsilon_{it}. \tag{2}$$

More explicitly, taking as example an outcome which we examine by age, we estimate the series of equations given in (3) below:

$$\begin{aligned}
Y_{it}^{7} &= \alpha + \beta_0^7 D_{it} + \beta_1^7 D_{it-1} + \mathbf{X}_{it}'\gamma + \mathbf{Z}_t'\delta + \theta_i + \epsilon_{it},\\
Y_{it}^{8} &= \alpha + \beta_0^8 D_{it} + \beta_1^8 D_{it-1} + \beta_2^8 D_{it-2} + \mathbf{X}_{it}'\gamma + \mathbf{Z}_t'\delta + \theta_i + \epsilon_{it},\\
Y_{it}^{9} &= \alpha + \beta_0^9 D_{it} + \beta_1^9 D_{it-1} + \beta_2^9 D_{it-2} + \beta_3^9 D_{it-3} + \mathbf{X}_{it}'\gamma + \mathbf{Z}_t'\delta\\
&\qquad + \theta_i + \epsilon_{it},\\
Y_{it}^{10} &= \alpha + \beta_0^{10} D_{it} + \beta_1^{10} D_{it-1} + \beta_2^{10} D_{it-2} + \beta_3^{10} D_{it-3} + \beta_4^{10} D_{it-4}\\
&\qquad + \mathbf{X}_{it}'\gamma + \mathbf{Z}_t'\delta + \theta_i + \epsilon_{it},\\
Y_{it}^{11} &= \alpha + \beta_0^{11} D_{it} + \beta_1^{11} D_{it-1} + \beta_2^{11} D_{it-2} + \beta_3^{11} D_{it-3}\\
&\qquad + \beta_4^{11} D_{it-4} + \beta_5^{11} D_{it-5} + \mathbf{X}_{it}'\gamma + \mathbf{Z}_t'\delta + \theta_i + \epsilon_{it},\\
Y_{it}^{12} &= \alpha + \beta_0^{12} D_{it} + \beta_1^{12} D_{it-1} + \beta_2^{12} D_{it-2} + \beta_3^{12} D_{it-3}\\
&\qquad + \beta_4^{12} D_{it-4} + \beta_5^{12} D_{it-5} + \mathbf{X}_{it}'\gamma + \mathbf{Z}_t'\delta + \theta_i + \epsilon_{it}.
\end{aligned} \tag{3}$$

The coefficients of interest, the $\hat{\beta}_0^{\tau} \dots \hat{\beta}_5^{\tau}$, provide estimates of the effect of current and past exposure to the policy.

5 Results

5.1 Main Results

The adoption of policies of social promotion did reduce repetition rates (see Tables C.1 and C.2 in the Appendix C): but did this translate into lasting improvements in age-grade correspondence? Table 5 presents our main results. Each column of the table lists coefficients on the treatment variables from a separate regression: for example, the regression in Column (1) has as dependant variable the share of nine year olds delayed in their studies, while Column (2) reports a regression on the same share for ten year olds. Lagged policy variables are included back to when the children in question would have been seven years old, and therefore expected to be entering first grade: two lags for nine year old, three lags for ten year olds, and so forth.

Table 5. Share of age group delayed.

	(1) Age 9	(2) Age 10	(3) Age 11	(4) Age 12
Municipal cycle policy	−0.0335***	−0.0267***	−0.0268**	−0.0164**
	(0.00514)	(0.00684)	(0.0117)	(0.00742)
L.Municipal cycle policy	−0.0419***	−0.0226***	−0.0367***	−0.0253***
	(0.00540)	(0.00782)	(0.0130)	(0.00859)
L2.Municipal cycle policy	−0.0232***	−0.00126	−0.0129	−0.00807
	(0.00459)	(0.00838)	(0.0140)	(0.00671)
L3.Municipal cycle policy		−0.00927	−0.0300*	−0.0112**
		(0.00695)	(0.0157)	(0.00480)
L4.Municipal cycle policy			−0.00195	0.00196
			(0.0113)	(0.00720)
L5.Municipal cycle policy				−0.00568
				(0.00503)
Joint F-test cycle variables	74.71	7.71	4.47	2.85
Prob > F	0.0000	0.0000	0.0004	0.0090
Observations	393,046	289,285	201,457	126,462

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Joint F-test performed over coefficients on current and lagged policy. Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of this age per school.

Table 5 shows that exposure to the social promotion policy, both in the current year and in past years, is consistently associated with reductions in age-grade delay. Exposure to the policy in the current or previous year is associated with a decrease of 1.1 pp–4.2 pp in the share of children delayed for their age.

This finding is not surprising: by promoting students year-by-year regardless of their academic level, the social promotion policy will mechanically reduce the share of children who are delayed for their age. For this reason, we focus our interest on Column (4): grade delay of 12 year-olds. Twelve year-old children should be in sixth grade: by our definition, they are classified as “delayed” if they are two or more years behind the grade they should be in. In this case, delayed 12 year-olds would be in grade four or less: in other words, they have not completed the transition from junior to senior primary school.

Strikingly, the contemporary and lagged effect of exposure to social promotion remains negative, and in many cases significant, into the twelfth year. While the coefficients on several of the lags are close to zero, others are meaningful and statistically significant. Summing these coefficients suggests that a cohort of 12 year-olds who have been exposed to social promotion since they were seven have, on average, 4.9 pp fewer members who are still in junior primary school, as compared to a similar cohort with no exposure to the policy.

While Table 5 indicates that cohorts exposed to social promotion have less grade delay than those not exposed to the policy, it raises an important question: could this be due to changing composition of the cohort? If students face grade repetition for the first time when they are 11 or 12, perhaps they are dropping out in larger numbers, leaving only the strongest students in their place.

To explore this, we estimate the effect of the policy on enrolment. If indeed students are dropping out after being exposed to social promotion, we should see reductions in enrolment at ages 11 and 12 for these cohorts. Table 6 shows that the policy had no effect on enrolment at older ages (in contrast, the policy did change enrolment by grade significantly, see Table C.3 in Appendix C).

Table 6. Schools: log total enrolments.

	(1) Age 7	(2) Age 8	(3) Age 9	(4) Age 10	(5) Age 11	(6) Age 12
Municipal cycle policy	0.0188 (0.0151)	−0.00809 (0.0138)	−0.00410 (0.0164)	0.00428 (0.0195)	−0.0256 (0.0296)	0.0133 (0.0344)
L.Municipal cycle policy	−0.0377*** (0.0133)	0.0143 (0.0146)	0.00754 (0.0178)	0.00531 (0.0207)	−0.00448 (0.0331)	−0.0322 (0.0364)
L2.Municipal cycle policy		−0.0151 (0.0138)	0.0215 (0.0198)	0.000644 (0.0218)	0.0209 (0.0319)	−0.0431 (0.0373)
L3.Municipal cycle policy			−0.0214 (0.0157)	−0.0172 (0.0243)	−0.0285 (0.0454)	−0.0590 (0.0523)
L4.Municipal cycle policy				−0.0282 (0.0202)	−0.0232 (0.0390)	−0.0511 (0.0547)
L5.Municipal cycle policy					−0.0481 (0.0297)	−0.00480 (0.0359)
Joint F-test cycle variables	4.04	0.59	0.74	0.72	0.76	0.75
Prob > F	0.0175	0.6206	0.5676	0.6108	0.6033	0.6057
Observations	484,043	388,606	290,130	209,005	135,836	126,462

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of this age per school.

5.2 Extension to Full Sample

The restrictions imposed by our sampling strategy are stringent: but are they necessary? Leaving aside selection of schools into the policy, is there any evidence for the negative selection into social promotion that Dong (2010) finds among US kindergarteners? Specifically, we are concerned that parents who are very invested in their child's education demand that schools impose strict promotion criteria, and oppose social promotion. We explore this by repeating the analysis presented above using the universe of Brazilian public primary schools.

Table 7 repeats the analysis presented in Table 5. Clearly, in this case the decrease in age grade mis-match due to social promotion *does not* withstand the transition from junior to senior primary school (Column (4)). While students exposed to the policy are not falling behind at faster rates than their peers who face repetition in every year, the effect of the policy appears to be purely mechanical: students automatically promoted through junior primary school remain there just as long as they would have without the policy.

Table 7. All public schools: share of age group delayed.

	(1) Age 9	(2) Age 10	(3) Age 11	(4) Age 12
School cycle policy	-0.0173^{***}	-0.00236	-0.00497^{**}	-0.00105
	(0.00141)	(0.00169)	(0.00215)	(0.00181)
L.School cycle policy	-0.0317^{***}	-0.0126^{***}	-0.0151^{***}	0.00100
	(0.00140)	(0.00164)	(0.00199)	(0.00155)
L2.School cycle policy	-0.0212^{***}	-0.0138^{***}	-0.0122^{***}	0.00155
	(0.00130)	(0.00160)	(0.00201)	(0.00151)
L3.School cycle policy		-0.00745^{***}	-0.0121^{***}	-0.00141
		(0.00157)	(0.00206)	(0.00152)
L4.School cycle policy			-0.00511^{***}	-0.000786
			(0.00192)	(0.00138)
L5.School cycle policy				-0.000491
				(0.00130)
Joint F-test cycle variables	351.46	41.19	21.99	0.72
Prob > F	0.0000	0.0000	0.0000	0.6342
Observations	716,684	591,214	452,822	308,139

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Joint F-test performed over coefficients on current and lagged policy. Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of this age per school.

What is happening to enrolment? Table 8, which replicates Table 6, suggests that, indeed, in the presence of other options, older students are moving away from schools with social promotion towards those without. It is worth re-iterating our identification strategy at this point: these regressions present *within school, across cohort* analyses, with identification coming from year-to-year policy variation at the school level. Within a school, those cohorts exposed to social promotion are smaller, at age 12, than those cohorts that are not.

There are two strong hypotheses for why this is the case: either parents are shifting successful students away from schools with social promotion policies, or junior primary schools are more likely to have social promotion cycles, and in years that they do have cycles, they are more efficient at moving 12 year-olds out. The latter, however, should hold as well in municipalities where all junior primary schools adopt social promotion—yet we fail to see this decrease in Table 6. This suggests, therefore that negative selection may be substantial: parents with high aspirations for their children shift them to better schools as they grow older, leaving worse-performing 12 year-olds in socially promoting schools.

Table 8. All public schools: log total enrolments.

	(1) Age 7	(2) Age 8	(3) Age 9	(4) Age 10	(5) Age 11	(6) Age 12
School cycle policy	−0.00663	−0.00229	0.00688	0.0213***	0.00463	−0.0132**
	(0.00437)	(0.00397)	(0.00450)	(0.00558)	(0.00671)	(0.00653)
L.School cycle policy	0.00322	0.00583	0.00490	0.00769	−0.00911	−0.0117*
	(0.00397)	(0.00379)	(0.00433)	(0.00549)	(0.00623)	(0.00603)
L2.School cycle policy		0.00867**	0.00974**	0.0136**	0.000545	−0.0156**
		(0.00359)	(0.00423)	(0.00542)	(0.00644)	(0.00634)
L3.School cycle policy			0.0177***	0.0135***	−0.00791	−0.0210***
			(0.00401)	(0.00523)	(0.00657)	(0.00660)
L4.School cycle policy				0.0124***	−0.00146	−0.0151**
				(0.00478)	(0.00598)	(0.00614)
L5.School cycle policy					0.00118	0.00680
					(0.00568)	(0.00585)
Joint F-test cycle variables	1.36	2.99	6.45	5.38	0.70	3.75
Prob > F	0.2571	0.0296	0.0000	0.0001	0.6520	0.0010
Observations	817,538	704,847	568,690	450,995	326,013	308,139

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of this age per school.

5.3 Private School Enrolment

If there is negative selection into social promotion when public schools have different policies, do private schools see an increase in enrolment when all public schools in the municipality adopt the policy? We test this by exploring how the social promotion policy adopted by public schools affects private school enrolment.

We find no evidence for significant movements towards private schools as a result of the policy. Table 9 presents the results of regressions on private school enrolment in first to fourth fourth grade. The sample is restricted to private schools in municipalities where public schools have a uniform repetition policy. Because this sample is much smaller, only current and lagged policy variables—for both the school itself and the public schools in the municipality—are included.

Although several individual coefficients are marginally significant, joint significance tests of the four policy variables together fail to reject the null hypothesis that the coefficients are equal to zero. It is worth noting that the negative impact of the policy on grade 1 enrolment suggested in Column (1) is consistent with the effect across all schools (see Table C.3 in the Appendix C). The increase in enrolment in fourth grade, just prior to the transition to senior

Table 9. Total private school enrolments.

	(1) Grade 1	(2) Grade 2	(3) Grade 3	(4) Grade 4
Municipal cycle policy	0.0447	0.0141	0.0159	0.0888*
	(0.0404)	(0.0423)	(0.0389)	(0.0461)
L.Municipal cycle policy	−0.0721*	0.00438	0.0154	0.00517
	(0.0376)	(0.0491)	(0.0410)	(0.0498)
School cycle policy	0.00172	−0.0128	−0.00449	−0.0375
	(0.0311)	(0.0313)	(0.0318)	(0.0317)
L.School cycle policy	−0.0111	−0.00127	−0.0152	−0.00230
	(0.0274)	(0.0248)	(0.0265)	(0.0253)
Joint F-test cycle variables	1.05	0.09	0.22	1.62
Prob > F	0.3817	0.9846	0.9259	0.1669
Observations	34,700	34,427	33,550	31,987

Notes: Each column presents results from a fixed effect school-level regression, restricted to private schools. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Joint F-test performed over coefficients on current and lagged policy. Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school.

primary school, would be consistent with parents shifting children away from public schools in key years.

6 Conclusion

This paper adds to the growing body of evidence highlighting the learning gaps endemic to public schools systems in many developing countries, and evaluating efforts to bridge them. Brazil is not alone in struggling with high repetition rates in primary school; however, the country's large-scale experiment with social promotion is unique.

We find that exposure to social promotion led to modest reductions in grade delay, which persist through the transition from junior to senior primary school. Our results are only meaningful to the extent that this transition represents a genuine hurdle. While the vast majority of schools with cycles have a first cycle that ends within junior primary school (see Table 2), there will be heterogeneity in the extent to which fourth grade selects students. Given the magnitude of our results, however, we argue that such heterogeneity is unlikely to full explain the persistent reduction in grade delay: a reduction which is only slightly smaller at age 12 than it is at younger ages where promotion is more likely to be mechanical.

It is worth emphasising that, *a priori*, the measured effect of social promotion on progression from junior to senior primary school could have been either positive or negative. While social promotion does mechanically improve age-grade correspondence, if such a policy is pushing children through primary school too quickly, it could reduce learning outcomes. In this case, we would expect the improvements in grade delay to be entirely undone at the first opportunity for repetition, and even increase grade delay at older ages if fewer children are able to make the grade.

Our analysis focusses on a particular group of students: those students who would have been held back under the traditional repetition policy, but are instead promoted. We have no information on an alternate, but also important group: students who would have been promoted anyway, but who may invest less effort under the social promotion regime. Previous studies have found evidence that the Cycles policy reduced test scores (Koppensteiner, 2014, Menezes-Filho et al., 2008). Our analysis contributes new, relatively positive, evidence to understanding the effects of this policy—but is only one part of the story.

References

Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis*, *31*(4), 480–499. http://dx.doi.org/10.3102/0162373709352239

Allensworth, E. 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. http://dx.doi.org/10.3102/01623737027004341

André, P. (2009, November). *Is grade repetition one of the causes of early school dropout? Evidence from Senegalese primary schools.* MPRA. https://mpra.ub.uni-muenchen.de/id/eprint/25665

Carvalho, S., & Firpo, S. (2014). O regime de ciclos de aprendizagem e a heterogeneidade de seus efeitos sobre a proficiência dos alunos. *Economia Aplicada*, *18*(2), 199–214. http://dx.doi.org/10.1590/1413-8050/ea374

Dong, Y. (2010). Kept back to get ahead? Kindergarten retention and academic performance. *European Economic Review*, *54*(2), 219–236. http://dx.doi.org/10.1016/j.euroecorev.2009.06.004

Fruehwirth, J. C., Navarro, S., & Takahashi, Y. (2016). How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. *Journal of Labor Economics*, *34*(4), 979–1021. http://dx.doi.org/10.1086/686262

Glewwe, P., & Kassouf, A. L. (2012). The impact of the *bolsa escola/familia* conditional cash transfer program on enrollment, dropout rates and grade promotion in Brazil. *Journal of Development Economics*, *97*(2), 505–517. http://dx.doi.org/10.1016/j.jdeveco.2011.05.008

Glick, P., & Sahn, D. E. (2010). Early academic performance, grade repetition, and school attainment in Senegal: A panel data analysis. *The World Bank Economic Review*, *24*(1), 93–120. http://dx.doi.org/10.1093/wber/lhp023

Gomes-Neto, J. B., & Hanushek, E. A. (1994). Causes and consequences of grade repetition: Evidence from Brazil. *Economic Development and Cultural Change*, *43*(1), 117–148. http://dx.doi.org/10.1086/452138

Greene, J. P., & Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, *2*(4), 319–340. http://dx.doi.org/10.1162/edfp.2007.2.4.319

Ikeda, M., & García, E. (2013). Grade repetition: A comparative study of academic and non-academic consequences. *OECD Journal Economic Studies*, 269–315. http://dx.doi.org/10.1787/eco_studies-2013-5k3w65mx3hnx

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, *86*(1), 226–244. http://dx.doi.org/10.1162/003465304323023778

Koppensteiner, M. F. (2014). Automatic grade promotion and student performance: Evidence from Brazil. *Journal of Development Economics*, *107*(C), 277–290. http://dx.doi.org/10.1016/j.jdeveco.2013.12.007

Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics*, *94*(2), 596–606. http://dx.doi.org/10.1162/REST_a_00165

Menezes-Filho, N., Vasconcellos, L., Werlang, S. R. d. C., & Biondi, R. L. (2008, November). Evaluating the impact of the Progressão Continuada Program on student flow rates and performance in Brazil. In *XIII Latin American and Caribbean Economic Association (LACEA) Annual Meeting*, Rio de Janeiro.

Roderick, M., & Nagaoka, J. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *Educational Evaluation and Policy Analysis*, *27*(4), 309–340. http://dx.doi.org/10.3102/01623737027004309

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

Appendix A Data

A.1 Schools

Table A.1 gives the mean number of schools per municipality across the panel, both overall and by administrative jurisdiction.

Table A.1. Mean number of primary schools per municipality.

Year	Total	Municipal	State	Private	Federal
1999	33.192	23.750	6.208	3.224	0.009
2000	32.838	23.547	6.027	3.256	0.008
2001	31.854	22.711	5.837	3.297	0.008
2002	30.901	21.828	5.724	3.342	0.008
2003	30.282	21.247	5.635	3.393	0.007
2004	29.814	20.905	5.517	3.386	0.007
2005	29.115	20.484	5.239	3.384	0.007
2006	28.449	19.915	5.128	3.399	0.007

Notes: Mean values across municipalities. Municipal, State, Private and Federal refer to administrative authority of the schools from 1999–2006.

Source: Censo Escolar, authors' calculations.

Table A.2 lists the mean number of students enrolled in each grade, conditional on enrolment in that grade being positive. Class sizes jump in 5th grade as there are approximately three times more primary schools offering junior grades (1–4) than senior grades (5–8). Table A.3 presents summary statistics of school-level age-specific enrolments, in levels and in natural logs.

Table A.2. Mean number of students enrolled per school by grade.

Variable	Mean	Std. Dev.	N
Grade 1	36.045	49.51	1,233,454
Grade 2	31.226	43.653	1,214,054
Grade 3	30.159	42.665	1,180,755
Grade 4	30.31	43.487	1,114,851
Grade 5	89.472	90.005	418,137
Grade 6	81.211	76.407	389,640
Grade 7	76.745	74.320	368,378
Grade 8	73.184	71.726	347,244

Note: Mean values across municipalities from 1999–2006.
Source: *Censo Escolar*, authors' calculations.

Table A.3. School-level mean enrolments: levels and natural logarith.

	Level	Std. Dev.	Ln	Std. Dev.	N
Age 6	6.645	12.833	1.19	1.066	474,989
Age 7	18.628	28.59	2.108	1.285	1,147,125
Age 8	22.546	34.779	2.262	1.318	1,208,863
Age 9	22.651	34.84	2.258	1.327	1,221,614
Age 10	22.29	34.585	2.217	1.346	1,253,105
Age 11	22.394	32.961	2.213	1.383	1,251,465
Age 12	23.948	38.664	2.142	1.47	1017822

Note: Mean values across schools.
Source: Censo escolar, authors' calculations.

Table A.4 describes the annual grade attainment of a single birth-cohort of students, illustrating the early onset and large extent of grade delay. The table shows municipal mean enrolment numbers and grade level of children born in 1994, from 2000 (when the children were six years old) to 2006 (when they were twelve).[16] Children are deemed to be "on time" with their studies if they are in grade one when they are seven years old, in grade two when they are eight, and so on. The "on time" students are highlighted in bold in Table A.4; those to the left of the bold figure are delayed, while those to the right are advanced for their age.

[16] We restrict our analysis to students six to twelve years old who are enrolled in grades that are no more than two years ahead of the age-appropriate level.

Table A.4. Annual enrolment and grade of 1994-born cohort.

Year	Gr. 1	Gr. 2	Gr. 3	Gr. 4	Gr. 5	Gr. 6	Gr. 7	Gr. 8
2000	82.38	1.78	–	–	–	–	–	–
2001	**456.16**	34.49	1.61	–	–	–	–	–
2002	241.65	**345.12**	29.45	1.63	–	–	–	–
2003	73.68	219.17	**308.06**	28.21	1.45	–	–	–
2004	32.29	83.73	196.34	**289.20**	26.43	1.40	–	–
2005	17.38	42.56	80.24	187.87	**273.99**	25.10	1.35	–
2006	8.59	23.43	45.30	83.15	193.66	**247.70**	21.53	1.41

Note: Mean values across municipalities.
Source: *Censo Escolar*, authors' calculations.

A.2 Control Variables

Summary statistics for the list of school covariates are given in Table A.5. All variables are taken directly from the *Censo Escolar* and IPEA, except for *Training of teachers* (at levels 1–4 and 5–8). This last variable is an index of mean education levels of teachers teaching at the specified grade levels, coded such that 0 represents less than primary education, 1 is completed primary education, 2 is completed secondary education, and 3 is any form of tertiary training. Summary statistics on the number and education levels of teachers in the school panel are calculated conditional on having at least one teacher teaching at that level.

Table A.5. Time-varying controls

Variable	Mean	Std. Dev.	N
Urban	0.41	0.49	1,372,731
Total teachers	12.12	16.92	1,372,731
Total primary school teachers	9.24	12.14	1,372,731
Number of teachers grades 1-4	4.76	6.11	1,372,731
Ed score of teachers grades 1-4	2.06	0.59	1,281,859
Municipal per capita GDP	4.38	5.09	1,372,731
Municipal population divided by school count	128.21	124.6	1,372,731

Note: Mean values across schools from 1999–2006.
Source: *Censo Escolar*, authors' calculations.

Appendix B The Cycles Policy

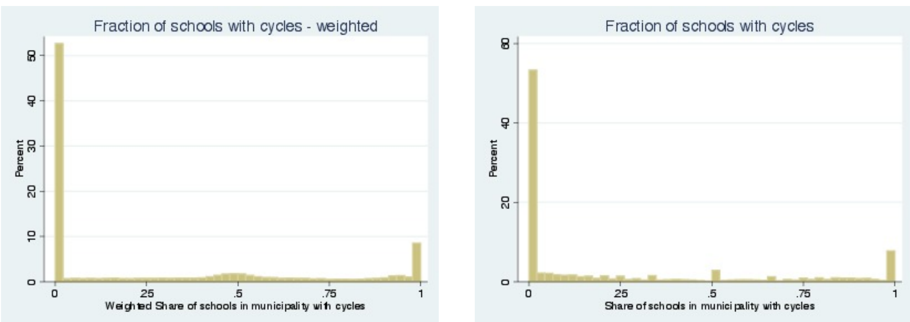
Table B.1 compares average school characteristics between all schools and the schools in our sample. The first row lists average characteristics for all schools which teach primary grades: this is the majority of schools. For the subsequent rows, we limit to public schools with a valid observation for log enrolment of seven year olds. The second row (**Valid**) gives average characteristics for all such schools which also have non-missing values for our control variables (see Table A.5). **Main sample** lists average school characteristics for such schools who also have two or more consecutive years of valid treatment observations (uniform policy at the municipal level). This row is comparable to the first column of Table 6 in the main text.

Table B.1. Comparison of sample schools to all schools.

Sample	Urban	Municipal	State	Private	GDP/c	Pop/school	Obs
All	0.41	0.70	0.19	0.11	4.38	128.21	1,372,731
Valid	0.32	0.84	0.16	0.00	3.88	108.59	1,003,048
Main sample	0.27	0.86	0.14	0.00	3.44	86.41	484,043

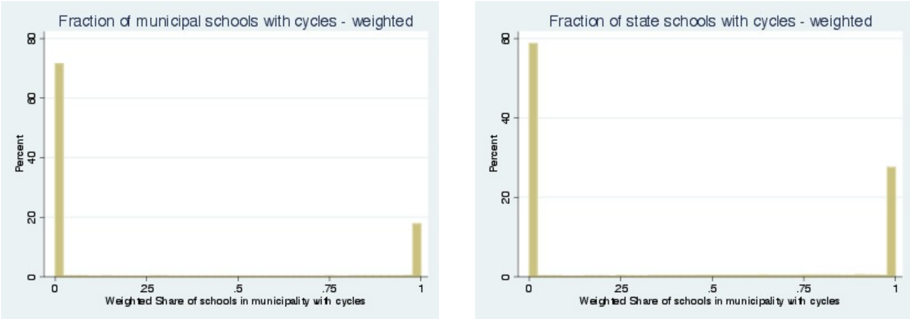
Source: Censo Escolar, authors' calculations. Mean values across schools from 1999–2006.

The share of schools in a municipality which use cycles is highly bimodal. Figure B.1 shows municipal share of schools who report using cycles, aggregated across all years of our panel, both overall (right panel) and weighted by municipal share of student enrolment (left panel). Most municipalities make no use of cycles; those that do, however, commonly adopt entirely. This bimodality is even more pronounced when looking at school jurisdictions individually. Figure B.2 plots the municipal share of schools with cycles (weighted), with municipal-run schools on the left and state-run schools on the right.



(a) (b)

Figure B.1. Distribution of cycle frequency: municipal means.



(a) **(b)**

Figure B.2. Distribution of cycles: municipal vs. state schools (weighted).

B.1 Geographic Variation

The popularity of cycles policies varies considerably across regions. A brief description of the general patterns follows: these overviews are based on a visual inspection of the distribution of municipality-level adoption rates for the years 1999, 2001, 2003 and 2005, for state-run and municipality-run schools separately.

Cycles in the North Municipal schools in the North have low or zero cycle adoption rates over the period. Rates are similarly low in state schools, with a few exceptions: state schools in Roraima report some cycle use (with a few municipalities registering a 100% adoption), while state schools in Tocantins have a range of adoption rates in 1999, diminishing to zero by 2003.

Cycles in the Northeast Municipal schools in the Northeast have low or zero cycle adoption rates, with the exception of Rio Grande do Norte which displays a strong bimodal distribution of municipalities: some adopt at near-census rates, while others avoid the policy entirely (rates peak in 2001–2003). Cycles are more prevalent among state schools. While half of the states have low or zero adoption, Ceará and Bahia display a 'messy' bimodal distribution (with some interior mass) from 1999–2003, and Pernambuco has such an adoption pattern in 2005 only (with no cycles prior to this). State schools in Rio Grande do Norte have a messy bimodal adoption pattern in 1999, which strengthens to a strong level of adoption in 2005 (most municipalities at 100%, and no mass at zero).

Cycles in the South Both municipal schools and state schools in the south adopted cycles at trivial rates, with the stark exception of municipal schools in Paraná. Municipal schools in Paraná display a distinctly bimodal adoption rate: most municipalities either fully adopt, or do not adopt cycles at all.

Cycles in the Southeast Municipal schools in the Southeast display a strongly bimodal distribution of adoption rates (Espírito Santo deviates slightly from this trend in 2005, with more interior points). State schools in general all adopted cycles. Exceptions to this are Minas Gerais in 2001 and 2003, and Espírito Santo in 2005, which are bimodal.

Cycles in the Centre-West The Centre-West region does not seem to follow a common trend. In Goiás, no schools adopted cycles at any point. In Mato Grosso do Sul, from 1999–2003, state schools all had cycles, while municipal schools mostly didn't, with some exceptions (including several with 100% adoption). In 2005, these rates fall to zero in both dependencies. Both municipal and state schools in Mato Grosso display bimodal adoption rates throughout the time period, with non-trivial interior mass among state schools.

Appendix C Extensions

C.1 Repetition rates

Did the policy actually affect retention, as intended? Table C.1 illustrates that the policy did indeed affect repetition rates. The policy has a strong contemporary negative effect on repetition: grade 1 repetition rates are 13.9% lower in cohorts affected by the policy, while the decrease is lower (4.1%, 6.3% and 2.7%, respectively) in grades 2, 3 and 4. Lagged exposure to the policy, in contrast, increases repetition rates. The largest effect of this type is seen in second grade, where exposure to the policy in the previous year is associated with a 4.1% higher repetition rates. The net effect by grade 4 is negative: after 4 years of exposure to the policy a larger share of students are retained in grade 4 than would be otherwise. Similar, though weaker, results hold when the sample is increased to the universe of public primary schools (Table C.2).

C.2 Enrolment by grade

While Table 6 shows that the policy had no effect on enrolment, with the exception of reducing it slightly for young children, Table C.3 shows that the policy did redistribute pupils significantly across grades. As expected, social promotion reduces enrolment in first grade: consequently, lagged social promotion increases enrolment in second grade. The increase in fourth grade enrolment associated with the lagged policy could be due to frequent implementation of a policy of two cycles of two years each (this is suggested by, but cannot be strictly deduced from, Table 2).

Table C.1. Repetition rate by grade.

	(1) Grade 1	(2) Grade 2	(3) Grade 3	(4) Grade 4
Municipal cycle policy	−0.139***	−0.0406***	−0.0629***	−0.0269***
	(0.00391)	(0.00494)	(0.00526)	(0.00610)
L.Municipal cycle policy		0.0373***	−0.00542	0.0233***
		(0.00467)	(0.00521)	(0.00794)
L2.Municipal cycle policy			0.0225***	0.00140
			(0.00459)	(0.00718)
L3.Municipal cycle policy				0.0149***
				(0.00511)
Joint F-test cycle variables	1,253.81	52.13	65.93	8.85
Prob > F	0.0000	0.0000	0.0000	0.0000
Observations	581,934	423,086	305,339	211,625

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, number of junior primary teachers in the school, education index for junior primary teachers in the school, municipal per capita GDP, municipal population per school.

C.3 More generous definition of delay

Our primary definition of grade delay captures students who are nominally two years delayed, half-way through the school year. Table C.4 presents results using a more "generous" definition of delay: students who are nominally delayed by one year, when the survey is carried out in May. Results are broadly similar to those presented in Table 5, although the evidence for persistence of the effect is weaker.

C.4 More restricted sample

Table C.5 replicates Table 5, using a sample of schools restricted further to those in municipalities where ALL schools (both public and private) follow the same repetition policy.

Table C.2. All public schools: repetition rate by grade.

	(1) Grade 1	(2) Grade 2	(3) Grade 3	(4) Grade 4
School cycle policy	-0.0833^{***}	-0.0167^{***}	-0.0250^{***}	-0.00861^{***}
	(0.00114)	(0.00121)	(0.00130)	(0.00154)
L.School cycle policy		0.0149^{***}	0.00337^{***}	0.00279^{*}
		(0.00115)	(0.00124)	(0.00148)
L2.School cycle policy			0.00495^{***}	0.00491^{***}
			(0.00114)	(0.00142)
L3.School cycle policy				0.00490^{***}
				(0.00134)
Joint F-test cycle variables	5,322.53	162.12	135.59	18.05
Prob > F	0.0000	0.0000	0.0000	0.0000
Observations	898,917	723,856	566,885	418,218

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, number of junior primary teachers in the school, education index for junior primary teachers in the school, municipal per capita GDP, municipal population per school.

Table C.3. Log total enrolments by grade.

	(1) Grade 1	(2) Grade 2	(3) Grade 3	(4) Grade 4
Municipal cycle policy	−0.0280**	0.0162	0.0118	0.0236
	(0.0124)	(0.0131)	(0.0142)	(0.0171)
L.Municipal cycle policy	−0.164***	0.0492***	−0.00865	0.0431**
	(0.0119)	(0.0145)	(0.0162)	(0.0201)
L2.Municipal cycle policy		−0.0119	−0.0117	0.0285
		(0.0125)	(0.0181)	(0.0200)
L3.Municipal cycle policy			−0.0187	−0.0174
			(0.0147)	(0.0241)
L4.Municipal cycle policy				0.00688
				(0.0175)
Joint F-test cycle variables	119.55	5.45	0.92	2.50
Prob > F	0.0000	0.0010	0.4491	0.0284
Observations	528,087	393,906	285,191	198,236

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of the target age for this grade per school.

Table C.4. Share of age group "possibly" delayed.

	(1) Age 8	(2) Age 9	(3) Age 10	(4) Age 11	(5) Age 12
Municipal cycle policy	−0.0424***	−0.0282***	−0.0287***	−0.00891*	−0.0165***
	(0.00630)	(0.00732)	(0.00811)	(0.00473)	(0.00482)
L.Municipal cycle policy	−0.0677***	−0.0505***	−0.0216**	−0.00833	−0.0186***
	(0.00643)	(0.00821)	(0.00927)	(0.00603)	(0.00569)
L2.Municipal cycle policy		−0.0157**	−0.0290***	0.00232	0.00102
		(0.00762)	(0.00955)	(0.00421)	(0.00577)
L3.Municipal cycle policy			−0.0133	0.00132	0.000514
			(0.00830)	(0.00411)	(0.00432)
L4.Municipal cycle policy				−0.00179	−0.00643
				(0.00317)	(0.00531)
L5.Municipal cycle policy					−0.00399
					(0.00382)
Joint F-test cycle variables	110.07	27.59	9.67	2.12	3.38
Prob > F	0.0000	0.0000	0.0000	0.0594	0.0024
Observations	516,153	393,046	289,285	201,457	126,462

Notes: Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Joint F-test performed over coefficients on current and lagged policy. Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of this age per school.

Table C.5. Share of age group delayed – alternate sample.

	(1) Age 9	(2) Age 10	(3) Age 11	(4) Age 12
Municipal cycle policy	-0.0416^{***}	-0.0323^{***}	-0.0173	-0.0153^{*}
	(0.00722)	(0.00816)	(0.0111)	(0.00806)
L.Municipal cycle policy	-0.0466^{***}	-0.0192^{**}	-0.0388^{***}	-0.0221^{***}
	(0.00754)	(0.00884)	(0.0120)	(0.00780)
L2.Municipal cycle policy	-0.0328^{***}	-0.00370	-0.00210	-0.0112^{*}
	(0.00601)	(0.00932)	(0.0133)	(0.00587)
L3.Municipal cycle policy		-0.0133	-0.0385^{***}	-0.00988^{**}
		(0.00826)	(0.0146)	(0.00462)
L4.Municipal cycle policy			-0.00997	-0.00588
			(0.0105)	(0.00599)
L5.Municipal cycle policy				-0.00570
				(0.00468)
Joint F-test cycle variables	69.48	7.26	4.37	2.40
Prob > F	0.0000	0.0000	0.0006	0.0253
Observations	377,113	285,567	203,143	128,516

Each column presents results from a fixed effect school-level regression. Standard errors in parentheses, clustered by school (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Joint F-test performed over coefficients on current and lagged policy. Controls (not shown): state-year dummies, urban location dummy, number of teachers in the school, number of primary school teachers in the school, municipal per capita GDP, municipal population per school, estimated municipal population of children of this age per school.