

IZA DP No. 8694

The Impact of Age of Entry on Academic Progression

Julio Cáceres-Delpiano
Eugenio P. Giolito

December 2014

The Impact of Age of Entry on Academic Progression

Julio Cáceres-Delpiano
Universidad Carlos III de Madrid

Eugenio P. Giolito
*ILADES, Universidad Alberto Hurtado
and IZA*

Discussion Paper No. 8694
December 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Impact of Age of Entry on Academic Progression^{*}

Using an RD-design and public educational administrative data for Chile, we study the impact of age of entry on children outcomes. Different from previous studies, we are able to track this impact on school achievements over eleven years of the school life of a cohort of students. Our results confirm previous findings that a higher age of entry not only has a positive effect on GPA and the likelihood of passing a grade but also that this impact tends to wear off over time. However, we also find that this impact on school achievement is still present eleven years after a child has started school. Moreover, we show that this decrease in the impact on GPA masks a return associated to a higher age of entry in other dimensions. First, we show that age of entry reduces the probability of being enrolled in a public school. Secondly, during secondary school, children delaying school entry are more likely to follow an academic track and we present evidence that these children are more likely to be enrolled in schools where children coming from other schools had a higher than the mean GPA in the school of origin. Finally, also explaining the decline in the impact of age of entry on school's achievements, we find evidence that age of entry is associated to an increase in the probability that a child is enrolled in a school actively engaged in *cream skimming*.

JEL Classification: A21, I24, I25, I28

Keywords: age of entry

Corresponding author:

Eugenio P. Giolito
ILADES, Universidad Alberto Hurtado
Facultad de Economía y Negocios
Erasmus Escala 1835
Santiago
Chile
E-mail: egiolito@uahurtado.cl

^{*} We want to thank the participants at the Department of Economics at Universidad Alberto Hurtado (July, 2014) for comments and suggestions. Financial support from the Spanish Ministry of Education (Grant ECO2009-11165) and from Fondecyt, Chile (Grant 1130565) is gratefully acknowledged. Errors are ours.

1 Introduction

We study the impact of age of entry on school performance using public administrative data for Chile. Different from previous literature, we are able to track this impact on a selected group of outcomes by following a cohort of students for over eleven years of their school life. This not only allows us to understand the evolution of the impact, but also sheds light on alternative channels that explain the pattern over time.

Since Deming & Dynarski (2008) pointed out a trend for the US in delaying school entry age, an increasing number of studies have explored the short and long term effects of age of entry.^{1 2} Despite the observed positive correlation between age of entry and academic achievements (Stipek, 2002), a series of recent studies reveal mixed results for these and other long-term outcomes once the endogeneity of age of entry is addressed.³ For children in primary school, age of entry has been negatively associated with grade retention and positively linked to scores in standardized tests (McEwan & Shapiro, 2006). The evidence for older students shows that age of entry has a negative effect on IQ, teen pregnancy and mental health (Black *et al.*, 2011). For young adults, age of entry is associated to a negative effect on completed years of education (Angrist & Kruger, 1991), lower earnings during their 20s and early 30s (Angrist & Kruger, 1991; Black *et al.*, 2011) and an insignificant effect on the Armed Forces Qualifying Test (Cascio & Lewis, 2006). Therefore, although a positive effect occurs in early school outcomes, age of entry seems to have an inconclusive long term effect.

Explaining this spectrum of results, we find first that for school age children and outcomes measuring school achievements, age of entry is perfectly correlated with age at test (Black *et al.*, 2011). That is, measuring the effect of age of entry for a child in a particular grade (or years since she/he started primary school) will jointly measure the impact of the age of entry and the age at test.⁴ Black *et al.* (2011), by disentangling these effects, show that while age at test has a positive impact on IQ, the impact of age of entry is negative. The majority of the literature using school outcomes, however, estimates the joint effect of age of entry and age at test. Nevertheless, by assuming that the impact of age at test decreases over time or that the impact of age of entry is an increasing function on the duration of education, the importance of the impact of the age of entry has been inferred (McEwan & Shapiro, 2006). Secondly, a delay in the age of entry reflects skill accumulation that takes place before school entry rather than the development of the necessary skills and maturity. However, these pre-school skill differences would diminish as children gain new skills and

¹As mentioned by the authors, one fourth of this change is explained by legal changes, while the rest is attributed to families, teachers and schools.

²A related part of the literature has focused on the impact of age rather than age of entry. Among these studies, we find Kelly & Dhuey (2006), who show that younger children obtain considerably lower scores than older ones at fourth and at eighth grades for a sample of OECD countries.

³Parents decide to delay a child's school entry based on individual benefits and costs. On the side of benefits, the literature has stressed the concept of "readiness" for learning (Stipek, 2002). On the cost side, families must face the cost of the labor income forgone by the household member in charge of childcare or the market cost of childcare for the unenrolled children. These and other factors defining the decision of holding back school entry are not always observed by the researcher, and potentially correlated with school outcomes. For example, mother's labor force participation has been shown to affect child's health, which might affect school performance (Sandler, 2011).

⁴Two children observed over the same time in the school system, but one of them, who had delayed the entry to school by a year, not only entered school at an older age but took every examination at a later age than the other child who did not delay school entry.

knowledge in the school. Consistent with this hypothesis, Elder & Lubotsky (2008) using differences in the entry cut-off among states in the US, find that the impact of age of entry into kindergarten is larger and more persistent among children from higher socio-economic status families who were more likely to accumulate skills prior to the start of kindergarten. Thirdly, the institutional feature that links mandatory education to age, rather than to years of education, leads to delayed age of entry enabling a child to reach the legal age to drop out school with fewer years of education (Angrist & Kruger, 1991). Finally, conditional on a positive or neutral effect on completed years of education, delaying school entry is associated to a delay in entry into the labor market and therefore to a detriment in the accumulation of labor experience (Black *et al.* , 2011).

In this paper, we focus the analysis on the impact of age of entry among children of school age. Our paper is closely related to that of McEwan & Shapiro (2006) who also study for Chile the benefits of delaying enrollment on educational outcomes. They show that an increase in one year in the age of enrollment is associated with a reduction in grade retention, a modest increase in GPA during the first years, and an increase in higher education participation. Also, like McEwan & Shapiro (2006), we address the endogeneity of age of entry by using the quasi-random assignment coming from the discontinuity in age of entry produced by the minimum age requirements. However differently from them, we first use the complete population of students. This is important given the recognized sorting of better students into schools with better teachers (Tincani, 2014). Secondly, we follow a cohort of students for over ten years of their school life. By following this cohort, we do not need to restrict the analysis to a specific grade, which is a function of the number of grades repeated, but rather to the number of years since a child starts primary school. This is particularly relevant for Chile where over 30% of the students in a particular cohort have repeated a grade at least once over their school life. Thirdly, by following a cohort of students for eleven years, we investigate the evolution of the impact of age of entry and its reported decline over the student's school life. Specifically for Chile, different from other countries, we observe laws on mandatory education linked to years of education rather than to student's age. In fact, for the period under analysis, the completion of secondary education is mandatory. This institutional feature enables us to study the long-term return of delaying school entry over a child's school life even into secondary education without concern that we simultaneously capture the impact of dropping out. Finally, by using other outcome variables going beyond school achievements, we are able to study the channels by which age of entry affects educational performance and understand the evolution of the impact associated to a delay in the age of entry. Specifically, we study the impact on school type, the type of academic track followed by the students and whether or not delaying school entry is associated to being in a school where we observe higher *cream skimming*.

Our findings confirm that delaying school entry not only has a positive effect on GPA, attendance and the likelihood of passing a grade but also this impact tends to wear off over time. Nevertheless, different from previous studies, our findings reveal that this impact is still observed eleven years after a child has started school. Moreover, evidence on the effect of age of entry on school type provides a potential explanation for

the decline in the impact on academic achievement over the school life. Specifically, a higher age of entry decreases the likelihood that a child is enrolled in municipal schools which are characterized by less active selection of students and lower quality teachers. Consistent with these differences in academic selection (competition), children with a later date of entry have a higher probability of being enrolled in schools where children coming from other schools had a higher than the mean GPA in the school of origin. We also find evidence that age of entry has a positive effect on the likelihood that a child follows an academic track in high school. Finally, also explaining the drop in the impact of age of entry on school achievements, we provide evidence that age of entry is associated with an increase in the probability that a child is enrolled in a school which is actively engaged in *cream skimming*.

The paper is organized as follows. In Section 2, we briefly sketch Chile's educational system, present the data set used in the analysis, and define the sample and the selected outcomes in the analysis. Section 3 describes our empirical strategy. In Section 4 we present our results, and Section 5 concludes.

2 Chilean education system, data and variables

Since a major educational reform in the early 1980s⁵, Chile's primary and secondary educational system has been characterized by its decentralization and by a significant participation of the private sector. By 2012, the population of students was approximately three and half millions, distributed throughout three types of schools: public or municipal (41% of total enrollment), non-fee-charging private (51% of total enrollment), and fee charging private schools (7% of total enrollment)⁶. Municipal schools, as the name indicates, are managed by municipalities, while the other two types of schools are controlled by the private sector. Though both municipal and non-fee-charging private schools get state funding through a voucher scheme, the latter are usually called voucher schools⁷.

Primary education consists of eight years of education while secondary education depends on the academic track followed by a student. A "Scientific-Humanist" track consists of four years and it prepares students for a college education. A "Technical-Professional" track has a duration in some cases of five years with a vocational orientation aiming to help the transition into the workforce after secondary education. Until 2003, compulsory education consisted of eight years of primary education; however, a constitutional reform established free and compulsory secondary education for all Chilean inhabitants up to the age of eighteen.

⁵The management of primary and secondary education was transferred to municipalities, payment scales and civil servant protection for teachers were abolished, and a voucher scheme was established as the funding mechanism for municipal and non-fee-charging private schools. Both municipal and non-fee-charging private schools received equal rates tied strictly to attendance, and parents' choices were not restricted to residence. Although with the return to democracy some of the earlier reforms have been abolished or offset by new reforms (policies), the Chilean primary and secondary educational system is still considered one of the few examples in a developing country of a national voucher system which in the year 2009 covered approximately 93% of the primary and secondary enrolment. For more details, see Gauri & Vawda (2003).

⁶There is a fourth type of schools, "corporations", which are vocational schools administered by firms or enterprises with a fixed budget from the state. In 2012, they constituted less than 2% of the total enrollment. Throughout our analysis, we treat them as municipal schools.

⁷Public schools and subsidized private schools may charge tuitions and fees, with parents' agreement. However, these complementary tuitions and fees cannot exceed a limit established by law.

Despite mixed evidence on the impact of a series of reforms introduced as of the early 1980's on the quality of education⁸, Chile's primary and secondary education systems are comparable in terms of coverage to any system we can observe in any developed country.

The primary data source in our analysis comes from public administrative records on educational achievement provided by the Ministry of Education of Chile for the period 2002-2012. These records contain individual information for the whole population of students during the years that a student stays in the system. Moreover, an individual's identification allows each student to be tracked over her/his whole school life.

We define "age of entry" as the age at which a student is observed at the beginning of the school year when she/he was enrolled in first grade. Therefore, we drop from the sample those children who were not observed in first grade the first time that they were enrolled. Secondly, the analysis is focused on the oldest cohort that was eligible to start school the first year for which we have data, that is, children born in 1996. These children, by complying with school's minimum age of entry rule, should have started school either in 2002 (those eligible to start school the year they turned six) or in 2003 (for those students who delayed their entry into school until the year they turned seven). That is, depending on the age of entry in primary school, we observe these children either eleven or twelve times (years) in the records. Given this last constraint, we center the analysis on the impact of age of entry over the first eleven years of a student's school life.

By using students' records we construct two set of outcomes. The first group of variables attempts to characterize the impact of age of entry on school performance. The first variable, attendance, corresponds to the percentage of days that a child has attended school during a given school year. Attendance, however, might well be capturing a school's effort since its funding is a function of students' attendance for those institutions receiving funding through the school voucher. In order to take into account that attendance is able to capture this and other school effects, we define the dummy variable "attendance below the median" as a dummy variable taking a value of one for students with attendance below the median in the class, and zero otherwise. The next variable is the annual average GPA over all subjects. As well as the variable attendance, GPA could reflect a school's characteristics rather than a student's own achievements⁹. As we did with attendance, we define five dummy variables indicating whether the GPA for a student a particular year is over the 90th, 75th, 50th, 25th and 10th percentile of the class. Finally for this group of outcomes, we define the variable "Pass" as a dummy variable taking a value of one when a student passes to the next grade, and zero otherwise.

The second group of variables is composed of variables describing the movement of students between schools and variables related to the school's characteristics. Two dummy variables describe the movement between schools in a given year. The first of these variables takes a value of one in the case of a child changing

⁸The bulk of research has focused on the impact of the voucher funding reform on educational achievements. For example, Hsieh & Urquiola (2006) find no evidence that school choice improved average educational outcomes as measured by test scores, repetition rates, and years of schooling. Moreover, they find evidence that the voucher reform was associated to an increase in sorting. Other papers have studied the extension of school days on children outcomes (Berthelon & Kruger, 2012), teacher incentives (Contreras & Rau, 2012) and the role of information about the school's value added on school choice (Mizala & Urquiola, 2013), among other reforms. For a review of these and other reforms since the early 1980's, see Contreras *et al.* (2005).

⁹Anecdotal evidence exists on grade inflation, which has not been equally observed among all schools

school in the middle of a school year, and zero otherwise. A central idea in a voucher system is to promote competition between schools and enable children to choose the school that fits them better. Recent evidence points as well to an active selection from the school side (Urquiola & Verhoogen, 2009). Independently of which side is actively choosing, the fact that a child changes school in the middle of the school year reflects a significant student-school mismatch. The second dummy variable in this group takes a value of one in the case where a child is observed in two (or more) different schools for two consecutive years.

The rest of the outcomes in this second group characterizes a school in three dimensions. First, we define two dummy variables indicating the type of provider. That is, the variable municipal (voucher) takes a value of one in case where a child is enrolled in a public (privately voucher funded) school, and zero otherwise. A third variable, “Scientific-Humanistic,” takes a value of one in the case where a child who is attending secondary education is enrolled in a school following an academic track to prepare students for college, and zero otherwise. Finally, the rest of the outcomes, since we are able to observe all the classmates over the eleven years that we follow this cohort of students, helps us to measure the degree of *cream skimming* observed in the school, that is, in which degree schools are able to select the better students and get rid of the students in the bottom of the distribution. First, we create a variable with the fraction of students among those coming from other schools who had grades in the previous school above the median. Since we do not have a standardized examination comparable across schools, this variable helps us as well to understand something about the quality of classmates. Specifically, this variable will increase when we lower the rotation of students into the school (smaller denominator), or more of the students that move come from the upper part of distribution in their previous school (larger numerator). Both of these movements could be related to a higher quality of the school. The next dummy variable takes a value of one in the case the GPA in first grade is higher than the median of the students that still remain from the first grade. For a child over the median in first grade, this variable will take a value of one. Nevertheless, in the case where the school is actively *cream skimming*, this variable will be less likely to take a value of one later on. We also define two dummy variables that indicate whether or not a particular student has, first, a GPA higher than the median of students who have ever moved, and the second one, a GPA higher than the students just moving into the school. Finally, the last outcomes correspond to the average GPA of the rest of the classmates not counting in this average the student’s own GPA.

The descriptive statistics are presented in Table 1. Our sample is composed of approximately 250,000 students that we observe for approximately 10.2 years; 13 percent of the students attend schools in rural areas; 91 percent of the children starting primary school do so in a school in their own municipality. The average age of entry is 6.23 years (2282 days), with 40% of the students starting school the year they turn six. In terms of the selected outcomes, the average attendance in a given year is over 91%. The average GPA is 5.6¹⁰ with approximately 93% of the students being promoted to the next class in every period. In a

¹⁰The grade scale in Chile goes from 0 to 7, with 7 being the highest grade. In order to pass a subject or test, students must obtain a grade of 4 or higher.

given year over the period under analysis, approximately 20% of students change school, with approximately 5% changing in the middle of the school year. In relation to the school type, approximately 47% and 45% of the students in a given year are enrolled in a public or voucher school, respectively. Although we have approximately 16% percent of the children following an academic track, this fraction is driven down for the periods in primary school. When we restrict the sample to students in secondary school, this fraction increases to approximately 60%.

3 Empirical specification

The specification of interest in our analysis can be expressed as follows,

$$y_{it} = \alpha_t + \gamma^t Aentry_i + \beta^t * X_i + v_{it}, \quad (1)$$

with y_{it} as one of the educational outcomes observed for a student i , t years after she/he started school. $Aentry_i$ corresponds to the age of entry and X_i , other predetermined variables. The parameter of interest is γ^t , which corresponds to the impact of age of entry on a selected outcome. The time superscript highlights the fact that we allow this impact to change over time. As extensively reported in the literature, we suspect that estimating equation (1) by OLS will produce inconsistent estimates of γ^t . Families who decide to delay school entry are more likely to have relatively higher (lower) gains (costs) associated to this delay. Unobserved variables correlated with these gains and costs such as parents' education, parents' motivation and so on, can also have a direct effect on a student's achievement, that is, the OLS estimates are likely to pick up the impact of these unobserved factors as well as the impact of age of entry.

In order to overcome this omitted variable problem and to estimate the impact of age of entry, γ^t , we make use of the minimum age of entry rules as a source of variation in the enrollment age in first grade of primary school. These rules establish that children, in order to be enrolled in first grade at primary school, must have turned six before a given date in the academic year. Children whose birthday takes place before this cutoff date are entitled to start school the year they turn six. Those whose birthday is after this cutoff must wait until the next academic year to start school. This discontinuity in the age of entry together with the assumption that parents cannot fully manipulate the date of birth, provides us with a potential quasi-experimental variation in the age of entry that constitutes the core of our "fuzzy"¹¹ Regression Discontinuity (RD) strategy.

Chile's official enrollment cutoff is April 1st. Nevertheless, the Ministry of Education provides some degree of flexibility to schools for setting other cutoffs between April 1st and July 1st. In fact McEwan & Shapiro (2006) show for Chile that there are four cutoffs used in practice: April 1st, May 1st, June 1st and July 1st. Given this data generating process, we could express $Aentry_i$ using the following regression model,

¹¹In a fuzzy RD design the probability of treatment (starting primary school at the age of six) changes in magnitude lower than one. On the other hand, in the case where the treatment is a deterministic function of day of birth, the probability of treatment would change from one to zero at the cut-off day. For more details, see Lee & Lemieux (2010).

$$Aentry_i = \rho + \pi_1 C1_i + \pi_2 C2_i + \pi_3 C3_i + \pi_4 C4_i + g(DB_i) + \varepsilon_{it}, \quad (2)$$

with Cs_i as a dummy variable that takes a value of one for students who have their birthday, DB_i , after the cutoff s , and zero otherwise; and DB_i as a variable taking values from 1 to 365, indicating the birthday over the calendar year for an individual i . Finally, the function $g(\cdot)$, is a fully flexible polynomial specification. The parameter π_s corresponds to the discrete change in age of entry associated to the minimum age of enrollment at the s cutoff, which we use to learn the impact of age of entry. The assumption that parents cannot precisely choose (manage) the day of birth makes the “treatment” (age of entry) as if it were “randomly” assigned for those individuals close to the four discontinuities.

Hahn *et al.* (2001), show that the estimation of causal effects in this regression discontinuity framework is numerically equivalent to an instrumental variable (IV) approach within a small interval around the discontinuity.¹² That is, while equation 2 can be seen as the first stage with Cs_i as the instruments in the analysis, the following expression corresponds to the structural equation of interest,

$$y_{it} = \alpha_t + \gamma' Aentry_i + \beta' * X_i + g(DB_i) + \varepsilon_{it}, \quad (3)$$

By including $g(\cdot)$ in the previous equation we recognize that students born at different times over the year might differ in a systematic manner. In fact, Buckles & Hungerman (2013), for the United States, show that season of birth is correlated with some mother’s characteristics. Specifically, they show that children born in winter are more likely to be born to a mother with lower levels of education, who are teen mothers, and Afro-American. The fact that these mother’s characteristics are correlated simultaneously with birthday and child’s educational outcomes, does not invalidate our RD approach. The effect of these observed and unobserved factors not changing discontinuously in the mentioned cutoffs is the basis for our identification.

In a context of “intrinsic heterogeneity” (Heckman *et al.* , 2006), the estimated γ' can be interpreted as a weighted “Local Average Treatment Effects” (LATE) across *all* individuals (Lee & Lemieux, 2010). That is, this fuzzy RD design, differently as usually understood, does not estimate the impact of age of entry just for those individuals around the discontinuity but for overall *compliers*. How close this weighted LATE will be to the traditional LATE will depend on how flat these weights are (Lee & Lemieux, 2010). Specifically, in our problem these weights correspond to the ex-ante probability that a child will be born close to any of these four discontinuities, that is, be born in the Chilean autumn. If we consider that all individuals in the population have a similar probability of being born in autumn, our estimated parameter will be close to the traditional LATE. Moreover, since the last cut-off (July 1st), as we will show in the following sections, is associated to a

¹²By focusing on the observations around these four discontinuities, first we concentrate on those observation where the age of entry is as if it were randomly assigned. This randomization of the treatment ensures that all other factors (observed and unobserved) determining a given outcome must be balanced at each side of these discontinuities. Secondly, and for a given parametrization of $g(\cdot)$, the estimated function can be seen as the non-parametric approximation of the true relationship between a given outcome and the variable day of birth, that is, we face a lower concern that the estimated impacts are driven by a incorrect specification of $g(\cdot)$.

practically perfect compliance in the age of entry, the use of only the last discontinuity is closer to sharp RD, where the interpretation of the estimated γ' is a weighted “Average Treatment Effects.” In the last section of the paper, we check the sensitivity of the results to restrict the analysis to the use of only the last discontinuity.

3.1 Bandwidth selection

The basic idea in a RD setting is the quasi random assignment of the treatment in the neighborhood around a discontinuity. Specifically, the group of observations used around the four discontinuities in our setting is defined by the selection of the bandwidth. While a larger bandwidth is associated to a higher precision, it is also linked to a higher bias due to the extrapolation required when the function $g(DB_i)$ ¹³ is estimated.

In order to select the optimal bandwidth, we use two of the most popular methods. The first one, the rule-of-thumb approach (Lee & Lemieux, 2010), is given for the following expression for each of the outcomes in the analysis,

$$h_{rot} = 2.702 * \left(\frac{R * \hat{\sigma}}{\sum_{i=1}^n (m_i'')^2} \right)^{\frac{1}{5}},$$

where R is the range of the running variable (DB_i), n is the sample size, and m'' and $\hat{\sigma}$ are the curvature and standard error in the regression for each of the selected outcomes on a fourth-degree polynomial in DB_i .

The second method is based on the calculation of the cross-validation function (Lee & Lemieux, 2010). For each of the selected outcomes we choose the bandwidth, h , which minimizes the value of the cross-validation function defined for,

$$CV_Y(h) = \frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}(DB_i))^2,$$

with $\hat{Y}(DB_i)$ as the predicted value for a selected outcome coming from a regression of Y on a first-degree polynomial of DB_i . These regressions use observations within the bandwidth h of observation i (only to the left of observation i if i is to the left of one of the cutoffs, and only to the right of observation i if i is to the right of one of the cutoffs), but excluding observation i itself. The optimal bandwidth is then the h that minimizes the value of the cross-validation function.

The results for the optimal bandwidths are reported in Tables 2 and 3 for the rule-of-thumb and cross validation approach, respectively. While the optimal bandwidth suggested when using the rule-of-thumb approach is between 5 and 10 days around the discontinuities,¹⁴ the cross validation bandwidth is approximately 20 days around the discontinuities¹⁵. Given these results we set a bandwidth of 15 days around each of the

¹³Specifically we use a flexible polynomial specification of degree G , which we allow to differ at every side of each of the four discontinuities, $g(BD_i) = \sum_{g=1}^G \alpha_g * BD_i^g + \sum_{s=1}^4 \sum_{g=1}^G \beta_{sg} * C_s(BD_i - X_s)^g$ with X_s the “ s ” cutoff day defining the minimum age requirement to start primary school.

¹⁴These values are obtained by setting the same bandwidth at each side of the four discontinuities. When we allow different bandwidths at each side of the four discontinuities, the optimal bandwidths when using the rule of thumb approach tend to be larger. We opted for the most conservative approach.

¹⁵This pattern of smaller bandwidths when using the rule of thumb over the cross-validation procedure is also reported in Lee &

discontinuities in our baseline specification, although in Section 4 we explore the sensitivity of the results using alternative bandwidths.

4 Results

4.1 Validity of the RD design

Our analysis using an RD design built on the fact that the variation of the treatment can be seen as good as a randomized assignment for those observations near a particular discontinuity. Then, as any random assignment, the pre-determined characteristics should have the same distribution among treated (just above each of the discontinuities) and control group (just below each of the discontinuities). Evidence of a systematic jump for these pre-determined characteristics would compromise the underlying assumption that individuals cannot precisely manipulate the running variable (Lee & Lemieux, 2010). Figure 1 inspects graphically the existence of a potential discontinuity among three baseline characteristics available in the data set: gender (fraction of males), a dummy variable that indicates whether or not the student lives in the same municipality of the school the year starting the first grade, and a dummy variable indicating whether or not they live in a rural area. Additionally, since we observe all the classmates, we can check for these potential differences in the characteristics of classmates in the first grade of primary school. For each baseline characteristic we plot the average for each day of birth across the calendar year.¹⁶ Moreover, we plot the fitted value for a flexible polynomial for each of the five samples defined by the discontinuities.¹⁷ The graphical representation does not show any sizable discontinuity for these selected variables. We also formally test for discontinuities in these baseline characteristics for alternative polynomial specifications and different bandwidths. The p-values are reported in Table 4. Only for eight out of 72 specifications is the null rejected. Specifically, the rejection of the null is constrained to the fraction of students who start school in the same municipality where they live. Moreover, for our baseline specification¹⁸ and selected bandwidth (15 days around the discontinuities) we cannot reject the null for any of the selected outcomes.

The randomization of treatment in the neighborhood of the discontinuities rests on the fact that families cannot precisely select the day of birth of their children. That is, the validity of an RD design could be compromised in cases in which individuals were able to *precisely* manipulate the running variable (day of birth) (Lee & Lemieux, 2010). In fact, we could expect that benefits/costs associated to a delay in the age of entry,

Lemieux (2010). While the rule of thumb approach sets the optimal bandwidth considering just the curvature of the estimated function, the cross validation approach weights as well the precision gain obtained when increasing the bandwidth. For a highly non-linear functions the gain in precision will not compensate the higher bias associated to larger bandwidth, so the rule of thumb and the cross validation function should suggest similar bandwidths. However, for a (more) linear (flat) function the gain in precision will dominate, so the suggested bandwidth using the cross validation procedure will tend to be larger.

¹⁶Every point on these figures corresponds to the average for a given outcome over the calendar year.

¹⁷By estimating a flexible model for the students born between April 1st-April 30th, May 1st-May 31st and June 1st-June 30th we end up with a highly non-linear model due to the higher dispersion in a smaller sample. In order to facilitate the inspection of a discontinuity, we also plot the predicted value for a quadratic model for all the students born before July 1st, the definitive cutoff, and compare it with the predicted value for students born after July 1st.

¹⁸Using Akaike's information criterion (AIC), we get a degree for $g(\cdot)$ that is either one or two depending on the outcome. We use a degree of two as a more conservative alternative.

together with the public knowledge about the minimum-age entry rules, might induce some families to choose the season of birth as a function of these and other individual gains. Moreover, Chile (together with Turkey and Mexico) is one of the countries with the highest rate of c-sections in the world. Also, Buckles & Hungerman (2013) for the United States show that season of birth is correlated with some mother's characteristics. These two observations suggest some power to select the running variable. Nevertheless, the fact that families could sort themselves over the calendar year does not invalidate this quasi-experimental design. The critical identification assumption, however, is that individuals lack the power to *precisely* sort themselves around these discontinuities (Lee & Lemieux, 2010). Under precise manipulation, we would find observations stacking up around the discontinuities, or in other words, we would observe a discontinuous distribution for the day of birth (the running variable). Figure 2 presents the histogram for the day of birth. Despite the high volatility, the figure masks a quite uniform distribution of birth during the calendar year but with a high dispersion among days of the week. We observe an average of 650 births per day with values that move between 800 to 500 births per day. However, dividing the sample by days of the week, Figure 3, the uniform distribution of births across the calendar year is evident and it does not support a discontinuity in the distribution of the running variable¹⁹. Finally, following McCrary (2008) we formally test for a discontinuity in the distribution of the running variable by estimating the density of the variable day of birth and formally testing for a discontinuity for each of these cutoffs. The graphical representation and the estimated discontinuities (with their standard errors) are reported in Figure 4. For any of the four thresholds we are able to reject the null hypothesis which supports our previous graphical analysis, that is, there is no evidence of a precise manipulation of the day of birth.

4.2 Discontinuity in the age of entry

Our RD design as a source of randomization of the treatment should not only ensure that other observed and unobserved factors are uncorrelated with the treatment, but also equally important, it provides a significant variation of the treatment defined as the age of entry in the educational system. Figure 5 presents the source of variation associated to the minimum age of entry using two definitions of the variable of interest: age of entry (in days) and a dummy variable indicating whether or not a child started school in the year she/he turned six. Firstly, we observe that those children born at end of the calendar year are older when starting school and are also less likely to start school the year that they turn six (they are more likely to start school the year they turn seven or more). However, conditional to being eligible to start school the year they turn six (being born before April 1st or being subject to the same eligible rule for those born after April 1st but before July 1st), those students born later but to the left of a particular cutoff are the younger ones in their respective classes.

¹⁹April 1st and July 1st for the cohort of children born in 1996 fell on a Monday. May 1st fell on a Wednesday and June 1st on a Saturday. Comparing the number of births around each of these thresholds with the corresponding day does not show any type of sorting around these thresholds. Probably the only anomaly is observed for Wednesday May 1st where we observe fewer births. However, we can explain this issue by a holiday effect rather than a sorting of births. Moreover, in the case in which families manipulate the timing of the births we would expect that this sorting would be larger around July 1st which is the definitive deadline to enroll a child the year she/he turn six.

Secondly, a distinguishable jump in the age of starting school is observed for children born around April 1st, May 1st, June 1st and July 1st. However for each of these thresholds a discontinuity in the treatment can be observed; the largest jump for those children born around the threshold of July 1st is noteworthy. This large jump around July 1st is explained by the perfect compliance associated to the rule of turning six before July 1st. In fact, the fraction of students starting school the year that they turn six (as opposed to the alternative of turning seven or more) drops to practically zero for those children born after July 1st.

Tables 5 and 6 present the estimated discontinuities (equation 2) for the two measures of the treatment, days at start of primary school and a dummy variable indicating whether or not a child started primary school the year she/he turned six, respectively. For both measures of the treatment, the results confirm the graphical analysis. Being born after the cutoff day pushes some students to delay the age of entry, that is, increases the average age of entry. This impact is significant for each of the cutoffs, however the largest discontinuity is observed for those individuals born after July 1st who experience an average increase of approximately half a year in the age of entry. The average increase for the rest of the cutoffs is between 15 to 45 days approximately. The results are robust to the selection of the bandwidths, degree of the polynomial or the inclusion of other covariates in the specification. Finally, following the equivalence with an IV approach, the value of the F-statistic for the null of the relevance of the excluded instruments is large enough to disregard any concern about weak instruments in all the specifications and selected bandwidths.

4.3 Impact of age of entry on selected outcomes. Graphical analysis

Figures 6 to 21 present the relationship between each of the selected outcomes and birthdays. For each outcome, we fit a flexible polynomial of second degree at every side of the four discontinuities and unique polynomial for the birthdays before July 1st. Overall and consistent with the larger discontinuity observed in the previous section, we observe a clear jump mainly at July 1st for the majority of the outcomes at some point of the school life. First, individuals born just after July 1st compared with the ones born at the end of June, have a higher average grade or are more likely to be in the upper part of the grade distribution in their classes. This discontinuity in school performance is observed for all years, although it is less precise as we get farther away from the first year in primary school. For school attendance we see a jump at the beginning of primary and secondary school. On the one hand, at the start of primary school having a birthday after July 1st is associated with higher attendance. On the other hand, at the 9th year after the start of school, first year of secondary, being born after July 1st is associated with lower attendance. Also for school performance, we observe that children born just after July's cutoff are more likely to go on to the following grade during the first four years of primary school. Nevertheless, the ninth year after starting school, the first year of secondary school for those children who have not repeated a grade, we observe that the students born just after July's cutoff are less likely to pass to the next grade. In terms of the outcomes characterizing the movement of students between schools, we observe among students born just after July's cutoff a higher fraction of them

changing school within the academic year of the first year of secondary school and a higher fraction starting in a different school during the second year of secondary school. In terms of school type, for students born just after July 1st, we observe a discontinuous drop in the fraction of students attending municipal schools and for secondary school, a discontinuous jump in the fraction of students following an academic track. For the outcome trying to capture the quality of the school measured by the fraction of students that come from the upper part of the distribution in the school of origin, no differences are shown for years covering primary school. However, as of the ninth year since the beginning of school life, those children born after July 1st are enrolled in schools with a higher fraction of students belonging to the upper part of the GPA's distribution in their previous school. Finally, Figures 18 to 21 present a picture that is consistent with the hypothesis that children who delay school entry stay or move to schools that are more likely to engage in *cream skimming*. That is, while for the first years of the school the fraction of students who have a higher age of entry (born after July 1st) are more likely to have a GPA higher than their classmates, these differences with respect to the students who are moving into school (or have moved at some point) or in relation to those who have stayed in the school, tend to disappear over time.

4.4 Impact of age of entry on selected outcomes

The impact of age of entry on the two selected groups of outcomes is reported in Tables 7 and 8. The overall picture we get from Table 7 regarding the impact of increasing age of entry on the different school achievement outcomes is not only that it is positive but the impact is still present eleven years after starting primary school. The only exception about this positive impact of increasing age of entry is observed for some of the outcomes (attendance) the ninth year after school entry, which would correspond to the first year of secondary school for those children who had not repeated a grade.

Specifically, first, for the variable “attendance” it is observed that children with a higher age of entry increase their attendance between one and two percentage points during the first two years after starting primary school and the second year of secondary school. This impact in terms of school days, considering an average attendance of 91% in the population, means that increasing age of entry is associated approximately to two to five more days of classes during a specific school year. The exception in the sign of the impact of age of entry is found for the first year of secondary school where a three percentage point reduction in the yearly attendance is observed. However, still for the ninth year after starting school, by looking at the impact on the probability of having an attendance rate higher than the median in the class, we can appreciate a positive effect associated to age of entry. This last finding suggests that the negative effect we capture on the attendance rate during the first year of secondary school reflects some school effects rather than an individual reduction in attendance. That is, while we observe that children with a higher age of entry have a lower attendance rate in relation to other students in other schools, this attendance rate is not lower than the one observed on average in their own class this ninth year after starting primary school.

Second, a higher age of entry is associated to a higher annual GPA for all years, for which we follow this cohort of students. Nevertheless, this impact tends to diminish with time. Specifically, the impact on GPA goes from approximately 0.5 points in the average GPA for a child during her/his first year in the school system to 0.15 points in the GPA eleven years after the entry in the system (the third year of secondary school for those children who have not repeated a grade). These magnitudes are not only statistically significant but also economically important. The magnitude of 0.5 (0.15) points in the increase in GPA is similar to the difference between a student in the median of the GPA distribution and another one in the 75th (60th) percentile of this distribution. Can this observed impact on the GPA be driven by school differences in grade inflation or be due to the differences in the requirements established in the school? The results for the outcomes indicating the place on the distribution of GPA in relation to the students within the same school cohort do not support this hypothesis. A higher age of entry increases by almost 20 percentage points the likelihood that a child is over the 90th percentile in the first year of school and approximately 5 percentage points, eleven years after school entry. Although the same qualitative results are observed for the other variables indicating whether or not a student is over the 75th, 50th, or 25th percentiles, the strongest impacts are observed in the upper part of the distribution of the GPA. In fact, only in the first years of the school life is it observed that a higher age of entry increases the probability of being over the 10th percentile in the GPA distribution of the class. Consistent with this positive impact on GPA, we find that a higher age of entry is associated with an increase in the probability of grade passing that goes from 2 to 5 percentage points.

In this way, and different from other studies, our results show an impact of age of entry that is still observed eleven years after the start of primary school. McEwan & Shapiro (2006) also using data for Chile and exactly the same source of variation in age of entry, although finding that an increase in age of entry reduces the likelihood of a child repeating the first year of primary school, the impact on a national test taken among students in fourth grade in 2002 shows a mild impact on this test. We understand these differences with respect to the analysis of McEwan & Shapiro (2006), first, on the base that this national examination was not taken every year among fourth graders, which implies that children who repeated before reaching fourth grade in 2002, were missing from the sample based on an outcome variable.²⁰ Secondly, due to data restriction, McEwan & Shapiro (2006) focus part of their analysis on publicly funded schools in urban areas²¹. Moreover, recent evidence for Chile points out a positive sorting of the students and schools. Can these sample differences be accountable for a reduction in the impact of age of entry on GPA? We explore this possibility by constructing a sample similar to the one in McEwan & Shapiro (2006). In order to do this,

²⁰The information on age of entry was constructed from seven cross sections for first graders during the period 1997-2004. Children in fourth grade the year 2002 are those who enter school the year 1998 (in the case in which they had not repeated a grade) or earlier (in the case in which they had repeated a grade). A child starting school the year 1998 and repeating a grade will not be observed in fourth grade the year 2002. Although McEwan & Shapiro (2006) do not restrict the sample to individuals who had not repeated, the restriction of the data to first graders for the period 1997-2004 and using this national examination for the year 2002, results in the only repeaters in the sample be those that at most had repeated one time and entered school the year 1997 or had repeated fourth grade the year before the examination.

²¹Data on first graders and the information about age of entry is constructed from data collected by the National School Assistance and Scholarship Board (JUNAEB). This survey excludes all private schools charging tuition and some of the private schools receiving government funding McEwan & Shapiro (2006).

we keep in our sample children who had not repeated a grade, in publicly funded schools and living in urban areas. The comparison of the results are reported in Figure 22. The horizontal axis corresponds to the years since the start of primary school and each line connects the point estimates. The continuous line represents our results while the dashed lines are the point estimates using the constrained sample. For the variable GPA (and almost all the outcomes), we observe that the impact of age of entry disappears after three years following school entry.

Tincani (2014) shows for Chile that private schools (privately and voucher funded) not only specialize in higher ability students but are also able to attract through better salaries higher quality teachers from public schools and from other sectors. In relation to our problem, we have shown that a higher age of entry provides some advantage in terms of school achievement. Thus, in the light of Tincani's findings, these better students (at some point in their school life) will be more likely to sort into better schools. Although the absolute impact of this sorting process on average GPA is ambiguous,²² the relative advantage associated to age of entry should decrease with average quality of classmates. The rest of the analysis seeks to measure the effect of age of entry on school's characteristic and movement between schools which are reported in Table 8.

We observe mixed results for the outcomes characterizing the movements between schools. While children delaying school entry are less likely to change schools the fourth year after the start of primary school, a year later they are more likely to move to other schools. In fourth grade children take a National Examination (SIMCE). This examination aims to measure the quality of the education provided by the schools which later on might be used by parents to choose schools. Moreover, there is anecdotal evidence indicating that schools would have some margin to select the students taking this examination. Along these lines, this finding that a higher age of entry is not only associated to better grades but also that children with a higher age of entry are less likely to change school the year when this national examination takes place, suggest that schools might have some power to retain these better students at least in the short-run (the year that this examination takes place). Also, during the first year of secondary school we observe that age of entry is associated with an increase in the chance of changing schools. In fact, between the ninth (first year of secondary school) and tenth year after starting primary school, child's age of entry is associated to an increase of approximately 6 percentage points in the likelihood of changing school between these two years. Moreover, we also observe these children still switching schools during the school year of the second year of secondary school. The start of secondary school in Chile is characterized by approximately 50% of the students in the educational system coming from eighth grade switching school. This amount of friction might induce some students to actively search for a new school. The case in which this search effort is positively correlated with early educational outcomes would explain a higher age of entry increasing the search effort in the periods with higher friction in the system. In fact, this greater search effort is consistent with a reduction in the probability of switching

²²Anecdotal evidence speaks for higher grade inflation among worse schools (<http://www.uchile.cl/portal/presentacion/historia/luis-riveros-cornejo/columnas/5416/inflacion-de-notas>), that is, the effect of age of entry should be negatively correlated with grade inflation, that is, the age of entry will be more likely to be observed among better schools. Now, in the case where better schools set higher standards, on the other hand, it will be harder to observe a stronger effect of age of entry as children move to better schools.

schools between the tenth and eleventh year. Also, consistent with this increase in the fraction of students switching school at the beginning of secondary school is the already reported drop in school attendance during the first year of secondary school.

In relationship to school type, we observe first that a higher age of entry is associated with a decrease in the probability that a child is attending a public school for almost all years under analysis. On the other hand, for only two years we do observe an increase in the probability of attending voucher type schools. While voucher schools are characterized by high heterogeneity in their quality, public schools have been reported to attract children with worse educational outcomes and children coming from disadvantaged socio-economic backgrounds.²³ This last phenomenon would explain our finding that an increase in the age of entry reduces the probability of being enrolled in public schools which are perceived to be of a lower quality, but due to the high heterogeneity among voucher schools the movement might take place within these schools. Secondly, we observe that delaying the age of entry increases by approximately 13 percentage point the likelihood of following an academic track. In terms of the sample, this means that this last impact corresponds to an approximately 25% increase in the fraction of students in a educational track that aims to place students in college.

For the outcomes characterizing the school measured by the classmates choosing the school (who have moved to the school) we observe that increasing the age of entry is associated with a rise in the fraction of students coming from the upper part of the GPA distribution in their previous school, but specifically in secondary school. Therefore, our results not only reveal that higher age of entry is associated with a lower probability of being enrolled in schools linked to lower quality teachers like the one we could find in public schools (Tincani, 2014), but also these students have better quality classmates in secondary school where we observe that they are more likely to follow an academic track. In fact, we can assume that all these factors contribute to the drop in the GPA in secondary school due to a tougher academic track and the less likely grade inflation reported in better schools. It is worth noting that the timing of these impacts is consistent with the drop in the GPA observed when starting secondary school.

The last three outcomes explore the impact of age of entry on the probability of being enrolled in a school that is actively *cream skimming*. We have shown that children with a higher age of entry have a greater likelihood of having a GPA higher than the mean of their classmates. In the case where some of these schools were actively engaged in *cream skimming*, however, the probability of having a GPA higher than the median of the students moving into the school should not only be lower but it should also fall over time. This is what we observe from the impact of age of entry on the outcomes that measure the probability of having a GPA higher than the median of students moving into the school in a given year or who have moved to the school at some

²³On the other extreme in the perception about the quality of schools, we find private schools attracting a higher fraction of students from advantageous socio-economic backgrounds and with better academic records. In an early version we reported the impact on the probability of attending a private school. We found that a higher age of entry is associated to an increase in 3 to 4 percentage points in the likelihood of attending a privately paid school, which in terms of the baseline distribution of students in the first year of primary schools, corresponds approximately to a 50% increase.

point in the past. That is, we observe that age of entry increases the probability of having a GPA higher than the median of students moving into the school, but this increase in the probability is lower than the increase in the probability of having a GPA higher than the median of all students in the class (Table 8) for almost all the years. Secondly, over the school life this probability decreases and when reaching secondary school it is not statistically significant. Also consistent with our hypothesis that age of entry increases the likelihood of being enrolled in a school actively *cream skimming*, when we compare the GPA in first grade over the years, but with the mean of the classmates coming from first year of primary school, this probability decreases over time. Finally, we observe for some years, with the exception of the first year of secondary school, that age of entry increases the probability of being enrolled in a school where the rest of the classmates have a higher GPA.

4.5 Sensitivity analysis

Our previous findings were obtained using a bandwidth of 15 days around the discontinuities and quadratic polynomial specification for the function $g(\cdot)$. In this section, we first explore the sensitivity of the findings to the selection of the bandwidth and the degree of the polynomial.

Figures 23 to 27 present graphically the sensitivity of the results to different values of the bandwidth. In each figure, we observe three lines. The black line corresponds to point estimates for a bandwidth of 15 days for every year since the start of primary education, and the blue and red line, the ones for bandwidths of 10 and 7 days, respectively. Two elements are worth mentioning. First, the point estimates are robust to the selection of the bandwidth with values close to each other. Secondly, as expected, due to the use of fewer observations we observe a wider confidence interval represented by the larger length of the vertical lines. Nevertheless, besides the loss of precision, the main conclusions obtained in the previous section are still sustained with these alternative bandwidths.

Figures 28 to 32 present a similar graphical analysis for the sensitivity of the findings to the polynomial degree. The black line presents the results when using a quadratic specification and the blue and red lines, for a linear and cubic specification of $g(\cdot)$. As well as the previous analysis, the results are robust to the selection of the polynomial's degree.

Finally, Tables 9 and 10 report the estimated impact of age of entry using just the discontinuity at 1st of July as a source of identification. The results are robust to the restriction in the instrument set. Since this last discontinuity is closer to a sharp RD whose estimated impact has can be interpreted as a weighted ATE, this finding rules out usual concerns about the external validity of the estimates.

5 Conclusions

Our findings confirm that delaying school entry not only has a positive effect on GPA, attendance and the likelihood of passing a grade but also this impact tends to wear off across time. Nevertheless, different from previous studies our findings reveal that this impact is still observed eleven years after a child has started school. Moreover, evidence about the effect of age of entry on school type provides a potential explanation for the fading of the impact on academic achievement throughout the school life. Specifically, a higher age of entry decreases the likelihood that a child will be enrolled in municipal schools which are characterized by a less active selection of students and lower quality teachers. Consistent with these differences in academic selection (competition), children with a higher age of entry increase their probability of being enrolled in schools where the children coming from other schools came from the upper GPA distribution in their previous school. We also find evidence that age of entry has a positive effect on the likelihood that a child follows an academic track in high school. Finally, also explaining the drop in the impact of age of entry on school achievements we find evidence that age of entry is associated to an increase in the probability that a child would be enrolled in a school actively engaged in *cream skimming*.

References

- ANGRIST, J. D., & KRUGER, A. B. 1991. Does compulsory schooling attendance affect schooling and earnings? *Quarterly journal of economics*, **106**, 979–1014.
- BERTHELON, M., & KRUGER, D. 2012. Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in chile. *Journal of public economics*, **95**(1-2), 41–53.
- BLACK, S., DEVEREAUX, P., & SALVANES, P. 2011. Too young to leave the nest? the effect of school starting age. *The review of economics and statistics*, **93**(2), 455–467.
- BUCKLES, K., & HUNGERMAN, D. 2013. Season of birth and later outcomes: Old questions, new answers. *Review of economics and statistic*, **95**(3), 711–724.
- CASCIO, E.U., & LEWIS, E.G. 2006. Schooling and the armed forces qualifying test. evidence from school-entry laws. *Journal of human resources*, **41**(2), 294–318.
- CONTRERAS, D., & RAU, T. 2012. Tournament incentives for teachers: Evidence from a scaled-up intervention in chile. *Economic development and cultural change*, **61**(1), 219–246.
- CONTRERAS, D., LARRANAGA, O., FLORES, L., LOBATO, F., & MACIAS, V. 2005. Políticas educacionales en chile: Vouchers, concentracion, incentivos y rendimiento. *In uso e impacto de la informacion educativa en america latina, ed. santiago cueto. santiago: Preal.*, 61–110.

- DEMING, D., & DYNARSKI, S. 2008. The lengthening of childhood. *The journal of economic perspectives*, **22**(3), 71–92.
- ELDER, T.E., & LUBOTSKY, D. H. 2008. Kindergarten entrance age and children’s achievement. *Journal of human resource*, **44**(3), 641–683.
- GAURI, V., & VAUDA, A. 2003. Vouchers for basic education in developing countries. a principal-agent perspective. *World bank policy research working paper 3005*.
- HAHN, J., TODD, P., & VAN DER KLAUW, W. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, **69**(1), 201–209.
- HECKMAN, J., URZUA, S., & VYTLACIL, E. 2006. Understanding instrumental variables in models with essential heterogeneity. *Review of economics and statistics*, **88**(3), 389–432.
- HSIEH, C., & URQUIOLA, M. 2006. The effects of generalized school choice on achievement and stratification: Evidence from chile’s voucher program. *Journal of public economics*, **90**(8-9), 1477–1503.
- KELLY, B., & DHUEY, E. 2006. The persistence of early maturity: International evidence of long-run age effects. *Quarterly journal of economics*, **121**(4), 1437–1472.
- LEE, D. S., & LEMIEUX, T. 2010. Regression discontinuity designs in economics. *Journal of economic literature*, **48**, 281–355.
- MCCRARY, J. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, **142**(2).
- MCEWAN, P., & SHAPIRO, J. 2006. The benefit of delayed primary school enrollment. discontinuity estimates using exact birth dates. *Journal of human resource*, **43**(1), 1–29.
- MIZALA, A., & URQUIOLA, M. 2013. School markets: The impact of information approximating schools’ effectiveness. *Journal of development economics*, **1003**, 313–335.
- SANDLER, M. 2011. The effects of maternal employment on the health of school-age children. *Journal of health economics*, **30**(2), 240–257.
- STIPEK, D. 2002. At what age should children enter kindergarten? a question for policy makers and parents. *Social policy report*, **16**(2), 3–16.
- TINCANI, M. M. 2014. *School vouchers and the joint sorting of students and teachers*.
- URQUIOLA, M., & VERHOOGEN, E. 2009. Class-size caps, sorting, and the regression-discontinuity design. *American economic review*, **99**(1), 179–215.

Table 1: Descriptive statistics.

Attendance	91.00 (15.37)	First grade GPA higher than median classmates coming from first grade	0.73
Attendance over the median	0.57		
GPA	5.57 (1.06)	GPA higher than median grade classmates ever-moved	0.52
GPA over p90	0.15		
GPA over p75	0.30	GPA higher than median classmates just moving IN to school	0.65
GPA over p50	0.56		
GPA over p25	0.80	Average grade other classmates	5.54 (0.62)
GPA over p10	0.94	Become 6 the year of Entry	0.40
Pass	0.93	Age at entry (days)	2282.63 (115.63)
Move within academic year	0.05	Male	0.51
Move school	0.20	Same municipality	0.91
Public school	0.47	Class size first grade	32.17
Voucher	0.45	Rural	0.13
Scientific or humanistic track	0.16	Periods	10.22
Fraction of incoming students over the median in previous school	0.52 (0.26)	Individuals in first grade	253794

Standard deviations in parentheses. Standard deviations for proportions are not reported.

Table 2: Optimal bandwidth. Rule-of-thumb criterion.

	Attendance	Attendance over media	GPA	GPA over p90	GPA over p75	GPA over p50	GPA over p25	GPA over p10	Pass
1	5.7	6.5	7.6	8.6	9.8	7.0	5.5	4.3	6.8
2	6.0	6.8	6.1	10.7	11.5	11.6	10.2	11.1	4.8
3	5.6	7.3	6.9	6.2	5.2	5.3	5.5	8.0	6.8
4	6.0	9.8	7.6	12.3	9.0	7.8	6.8	7.3	7.4
5	5.7	6.9	10.0	5.3	8.6	7.5	7.3	8.0	11.2
6	7.4	6.9	8.0	6.5	9.1	6.9	7.5	10.7	11.6
7	7.0	9.8	7.5	7.7	12.5	9.7	12.4	5.5	6.4
8	4.3	9.1	5.3	10.2	10.9	6.0	7.4	13.7	8.2
9	6.1	5.8	5.8	7.0	7.5	8.9	11.4	7.5	9.2
10	4.8	6.7	8.4	6.6	8.3	10.5	6.4	5.0	6.6
11	6.1	5.6	5.8	6.5	6.2	5.7	6.9	8.0	5.6
Mean	5.9	7.4	7.2	8.0	9.0	7.9	7.9	8.1	7.7
	Move within academic year	Move school	Public school	Voucher	Scientific-humanistic track	Fraction of incoming students over the median in previous school	First grade GPA higher than median classmates coming from first grade	GPA higher than median grade classmates evermoved	GPA higher than median classmates just moving IN to school
1	5.5		6.9	5.3		4.3	7.3	9.4	
2	5.8		5.7	4.9		12.0	8.0	8.2	11.8
3	7.9	7.9	6.9	5.8		6.6	9.4	5.8	9.1
4	13.5	5.3	7.9	6.0		5.7	10.0	6.0	10.5
5	6.4	6.3	6.6	4.6		10.3	7.5	8.7	5.9
6	9.2	4.9	5.7	4.7		9.8	7.4	10.6	6.9
7	5.8	4.7	5.7	4.9		4.9	6.5	9.6	7.0
8	4.4	7.9	6.8	5.9		4.9	6.5	7.2	7.1
9	5.1	7.9	9.3	8.2	5.0	5.5	9.6	7.5	6.8
10	7.8	9.8	9.3	6.5	5.5	11.7	5.4	8.8	5.9
11	6.4	5.8	6.5	6.0	5.1	8.7	6.1	5.7	5.8
Mean	7.1	6.7	7.0	5.7	5.2	7.7	7.6	8.0	7.7

Table 3: Optimal bandwidth. Cross-validation criterion.

	Attendance	Attendance over media	GPA	GPA over p90	GPA over p75	GPA over p50	GPA over p25	GPA over p10	Pass
1	30	29	11	16	11	10	10	30	30
2	20	30	14	23	15	24	19	29	28
3	30	30	17	16	16	17	25	28	30
4	12	30	14	17	17	20	30	30	29
5	19	28	15	29	15	24	24	30	26
6	26	30	17	17	29	29	29	21	30
7	25	30	27	30	29	29	30	27	30
8	24	26	26	27	29	29	30	30	20
9	18	30	27	30	30	24	30	30	23
10	30	28	26	30	30	29	30	30	30
11	22	30	28	29	28	28	26	27	28
Mean	23.3	29.2	20.2	24.0	22.6	23.9	25.7	28.4	27.6
	Move within academic year	Move school	Public school	Voucher	Scientific-humanistic track	Fraction of incoming students over the median in previous school	First grade GPA higher than median classmates coming from first grade	GPA higher than median grade classmates evermoved	GPA higher than median classmates just moving IN to school
1	30		20	27		30	9	10	
2	28	30	20	29		28	19	19	19
3	30	29	27	30		27	20	24	28
4	30	30	18	30		27	20	20	14
5	30	27	17	30		27	27	17	30
6	24	25	28	30		26	19	21	30
7	30	28	20	30		28	19	29	29
8	29	25	20	30		28	30	30	30
9	30	28	30	30	19	27	30	29	29
10	30	28	30	30	28	25	25	30	26
11	28	17	28	30	27	25	29	28	28
Mean	29.0	26.7	23.5	29.6	24.7	27.1	22.5	23.4	26.3

Table 4: Differences in predetermined variables over the discontinuities. p-values.

Degree	Bandwidth	Male	Same Comuna	Rural	Class mates		
					Male	Same Comuna	Rural
P4	15 days	0.709	0.007	0.370	0.848	0.913	0.456
	10 days	0.480	0.001	0.161	0.528	0.313	0.249
	7 days	0.106	0.000	0.269	0.098	0.002	0.367
P3	15 days	0.700	0.695	0.853	0.196	0.461	0.885
	10 days	0.583	0.198	0.226	0.634	0.539	0.240
	7 days	0.026	0.000	0.425	0.059	0.046	0.493
P2	15 days	0.862	0.788	0.929	0.081	0.664	0.925
	10 days	0.884	0.369	0.326	0.537	0.823	0.374
	7 days	0.795	0.000	0.279	0.019	0.445	0.387
P1	15 days	0.969	0.464	0.937	0.922	0.330	0.951
	10 days	0.808	0.731	0.952	0.489	0.376	0.980
	7 days	0.979	0.770	0.663	0.754	0.636	0.671

Table 5: Impact of minimum age entry rules on age of entry (days).

		1			2			3			4			
		15 days	10 days	7 days	15 days	10 days	7 days	15 days	10 days	7 days	15 days	10 days	7 days	
I Unconditional	Born After April 1st	28.3233*** [3.4283]	27.5121*** [4.4860]	31.8843*** [4.1711]	28.7104*** [5.1742]	37.9610*** [3.9449]	40.2898*** [5.0254]	29.4427*** [5.7237]	38.8296*** [4.6506]	41.7091*** [5.9641]	39.4482*** [6.3781]	40.0230*** [5.7847]	40.9451*** [6.3855]	
	Born After May 1st	21.3254*** [3.6022]	22.1842*** [3.8046]	22.6974*** [4.3265]	25.9690*** [5.2597]	26.3100*** [5.4136]	25.9260*** [6.1072]	28.4573*** [6.2929]	28.4133*** [6.3891]	29.3772*** [6.8575]	26.0146*** [7.2710]	25.2956*** [6.9952]	26.9640*** [7.1053]	
	Born After June 1st	47.7396*** [4.3010]	47.4474*** [4.6499]	49.4726*** [5.5082]	47.8233*** [6.9130]	51.3239*** [7.6219]	47.2595*** [8.4934]	50.6623*** [10.1476]	50.8338*** [10.8000]	48.0819*** [12.9671]	54.8340*** [13.3338]	52.2654*** [14.3005]	61.4692*** [14.8077]	
	Born After July 1st	174.1180*** [2.8339]	175.7277*** [3.0380]	174.5104*** [3.1186]	175.2875*** [4.0611]	173.3698*** [3.8915]	176.7629*** [4.6390]	176.7504*** [5.5525]	169.1098*** [4.8457]	172.3772*** [5.8295]	169.5726*** [6.5500]	171.5069*** [5.9557]	163.7984*** [7.7617]	
	Observations	62,937	43,288	30,791	62,937	43,288	30,791	62,937	43,288	30,791	62,937	43,288	30,791	
	R-squared	0.2001	0.2097	0.2107	0.2002	0.2099	0.2107	0.2002	0.2099	0.2108	0.2003	0.2099	0.2109	
	F-statistic	1086.93	939.09	858.82	625.6	765.79	488.96	321.76	390.92	231.66	242.18	284.4	179.18	
	II Conditional	Born After April 1st	29.5483*** [3.3503]	28.7123*** [4.4369]	33.0564*** [3.9279]	29.5923*** [5.2206]	39.5316*** [3.9459]	39.6229*** [5.1033]	29.4691*** [5.8287]	39.5520*** [4.7182]	40.3621*** [6.0876]	40.1387*** [6.1773]	39.6881*** [5.6223]	38.3298*** [6.0367]
		Born After May 1st	22.1573*** [3.6029]	22.8459*** [3.8502]	23.0487*** [4.3750]	26.6489*** [5.2151]	27.3817*** [5.3175]	27.3359*** [5.9482]	27.8893*** [6.0016]	27.9771*** [6.1382]	29.7901*** [6.6256]	27.9677*** [7.2896]	27.9226*** [7.0224]	31.2337*** [7.1966]
		Born After June 1st	47.7814*** [4.1109]	48.0355*** [4.4200]	49.6617*** [5.2006]	48.2743*** [6.5390]	50.8721*** [7.2403]	46.6316*** [8.0494]	51.0012*** [9.7263]	50.0685*** [10.2374]	47.3477*** [12.3223]	52.7347*** [12.5745]	50.0947*** [13.3697]	59.1699*** [15.9308]
Born After July 1st		173.1264*** [2.8338]	174.6430*** [3.0815]	173.2580*** [3.1847]	174.0252*** [4.0713]	172.6457*** [3.7673]	176.4671*** [4.4782]	175.5515*** [5.1979]	169.0771*** [3.9867]	172.1695*** [5.4183]	168.9480*** [5.4995]	170.6162*** [5.4193]	161.7350*** [6.7335]	
Observations		62,922	43,281	30,786	62,922	43,281	30,786	62,922	43,281	30,786	62,922	43,281	30,786	
R-squared		0.2254	0.2343	0.2367	0.2254	0.2345	0.2367	0.2255	0.2345	0.2367	0.2255	0.2345	0.2368	
F-statistic (excluded instruments)		1095.64	938.02	855.54	638.36	769.77	511.2	327.08	424.72	236.28	274.84	306.31	191.38	

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at birth day level.

Other covariates in the conditional specification are dummies for male, rural status and whether or not the child started school in the same municipality where she/he resided.

Table 6: Impact of minimum age entry rules on the probability to start school in the year the student becomes six years old.

		1			2			3			4			
		15 days	10 days	7 days	15 days	10 days	7 days	15 days	10 days	7 days	15 days	10 days	7 days	
I Unconditional	Born After April 1st	-0.0701*** [0.0062]	-0.0677*** [0.0074]	-0.0752*** [0.0064]	-0.0707*** [0.0081]	-0.0866*** [0.0050]	-0.0902*** [0.0072]	-0.0711*** [0.0081]	-0.0880*** [0.0056]	-0.0928*** [0.0081]	-0.0885*** [0.0083]	-0.0889*** [0.0069]	-0.0906*** [0.0085]	
	Born After May 1st	-0.0527*** [0.0070]	-0.0550*** [0.0071]	-0.0552*** [0.0086]	-0.0596*** [0.0107]	-0.0581*** [0.0103]	-0.0600*** [0.0113]	-0.0681*** [0.0124]	-0.0666*** [0.0115]	-0.0695*** [0.0125]	-0.0627*** [0.0141]	-0.0657*** [0.0134]	-0.0690*** [0.0132]	
	Born After June 1st	-0.1303*** [0.0090]	-0.1282*** [0.0096]	-0.1335*** [0.0112]	-0.1263*** [0.0130]	-0.1352*** [0.0139]	-0.1257*** [0.0151]	-0.1357*** [0.0175]	-0.1379*** [0.0181]	-0.1303*** [0.0227]	-0.1448*** [0.0233]	-0.1361*** [0.0241]	-0.1502*** [0.0282]	
	Born After July 1st	-0.4926*** [0.0076]	-0.4963*** [0.0082]	-0.4925*** [0.0086]	-0.4982*** [0.0109]	-0.4931*** [0.0101]	-0.4982*** [0.0135]	-0.4950*** [0.0145]	-0.4789*** [0.0134]	-0.4873*** [0.0178]	-0.4795*** [0.0179]	-0.4830*** [0.0166]	-0.4706*** [0.0212]	
	Observations	81,859	56,318	40,073	81,859	56,318	40,073	81,859	56,318	40,073	81,859	56,318	40,073	
	R-squared	0.3405	0.3448	0.3456	0.3405	0.3448	0.3456	0.3405	0.3449	0.3457	0.3406	0.3449	0.3457	
	F-statistic	1086.93	939.09	858.82	625.6	765.79	488.96	321.8	390.94	231.67	242.18	284.44	179.19	
	II Conditional	Born After April 1st	-0.0710*** [0.0063]	-0.0692*** [0.0075]	-0.0763*** [0.0061]	-0.0714*** [0.0085]	-0.0890*** [0.0061]	-0.0876*** [0.0080]	-0.0700*** [0.0087]	-0.0883*** [0.0067]	-0.0885*** [0.0089]	-0.0893*** [0.0094]	-0.0870*** [0.0074]	-0.0841*** [0.0086]
		Born After May 1st	-0.0537*** [0.0070]	-0.0556*** [0.0072]	-0.0544*** [0.0087]	-0.0595*** [0.0107]	-0.0589*** [0.0102]	-0.0613*** [0.0109]	-0.0644*** [0.0118]	-0.0626*** [0.0110]	-0.0680*** [0.0121]	-0.0641*** [0.0137]	-0.0685*** [0.0129]	-0.0749*** [0.0132]
		Born After June 1st	-0.1305*** [0.0084]	-0.1299*** [0.0087]	-0.1342*** [0.0103]	-0.1288*** [0.0119]	-0.1360*** [0.0128]	-0.1257*** [0.0140]	-0.1374*** [0.0163]	-0.1370*** [0.0165]	-0.1299*** [0.0208]	-0.1410*** [0.0210]	-0.1317*** [0.0211]	-0.1464*** [0.0254]
Born After July 1st		-0.4906*** [0.0076]	-0.4939*** [0.0082]	-0.4896*** [0.0086]	-0.4948*** [0.0109]	-0.4906*** [0.0101]	-0.4968*** [0.0134]	-0.4921*** [0.0142]	-0.4784*** [0.0124]	-0.4854*** [0.0172]	-0.4779*** [0.0165]	-0.4793*** [0.0156]	-0.4635*** [0.0200]	
Observations		81,842	56,309	40,067	81,842	56,309	40,067	81,842	56,309	40,067	81,842	56,309	40,067	
R-squared		0.3585	0.3623	0.364	0.3585	0.3623	0.364	0.3585	0.3623	0.364	0.3586	0.3623	0.3641	
F-statistic (excluded instruments)		1095.64	938.02	855.54	638.36	769.77	511.2	327.08	424.72	236.28	274.84	306.31	191.38	

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at birth day level.

Other covariates in the conditional specification are dummies for male, rural status and whether or not the child started school in the same municipality where she/he resided.

Table 7: Impact of age entry on selected outcomes. School performance.

	Attendance	Attendance over media	GPA	GPA over p90	GPA over p75	GPA over p50	GPA over p25	GPA over p10	Pass
1	1.9691** [0.8130]	0.1077*** [0.0231]	0.4793*** [0.0524]	0.1904*** [0.0222]	0.2917*** [0.0241]	0.3106*** [0.0243]	0.1946*** [0.0217]	0.0605*** [0.0141]	0.0163 [0.0104]
2	1.1253 [0.7038]	0.0117 [0.0250]	0.3966*** [0.0500]	0.1448*** [0.0142]	0.2033*** [0.0205]	0.2037*** [0.0241]	0.1372*** [0.0169]	0.0551*** [0.0123]	0.0546*** [0.0106]
3	1.1880** [0.5734]	0.0321 [0.0277]	0.3117*** [0.0415]	0.1549*** [0.0188]	0.2190*** [0.0258]	0.1770*** [0.0261]	0.1036*** [0.0209]	0.0752*** [0.0151]	0.0486*** [0.0099]
4	0.1595 [0.5310]	-0.0259 [0.0238]	0.2160*** [0.0403]	0.1109*** [0.0191]	0.1678*** [0.0249]	0.1467*** [0.0226]	0.0955*** [0.0228]	0.0146 [0.0131]	0.0206** [0.0090]
5	0.5353 [0.4952]	0.0374 [0.0275]	0.2060*** [0.0402]	0.0879*** [0.0177]	0.1409*** [0.0348]	0.1355*** [0.0262]	0.0766*** [0.0213]	0.0187 [0.0126]	0.018 [0.0117]
6	0.1043 [0.6490]	0.0443 [0.0282]	0.2199*** [0.0432]	0.0864*** [0.0183]	0.1416*** [0.0182]	0.1503*** [0.0184]	0.0633*** [0.0158]	0.0410*** [0.0126]	0.0259** [0.0106]
7	0.2846 [0.5740]	0.0853*** [0.0273]	0.1696*** [0.0359]	0.0509*** [0.0151]	0.0964*** [0.0194]	0.0880*** [0.0275]	0.0626*** [0.0162]	0.0318** [0.0125]	0.0263** [0.0131]
8	0.4302 [0.6122]	0.0244 [0.0359]	0.0774** [0.0328]	0.0744*** [0.0131]	0.1135*** [0.0221]	0.0945*** [0.0221]	0.0287 [0.0189]	0.0069 [0.0132]	-0.0021 [0.0109]
9	-3.3135*** [0.9667]	0.0662*** [0.0242]	0.0597 [0.0520]	0.0697*** [0.0240]	0.0957*** [0.0239]	0.0661*** [0.0229]	0.0279* [0.0155]	0.0319** [0.0149]	-0.0225 [0.0154]
10	2.1522* [1.2364]	-0.0312 [0.0245]	0.1335** [0.0617]	0.0656*** [0.0228]	0.0601*** [0.0215]	0.0643*** [0.0214]	0.0139 [0.0178]	0.0109 [0.0094]	0.0226 [0.0188]
11	-0.6063 [1.6025]	0.0392 [0.0297]	0.1504* [0.0836]	0.0495** [0.0199]	0.0627*** [0.0192]	0.0724*** [0.0197]	0.0367* [0.0198]	0.01 [0.0089]	0.0259 [0.0174]

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at birth day level.

Other covariates in the specification are dummies for male, rural status and whether or not the child started school in the same municipality where she/he resided.

Table 8: Impact of age entry on selected outcomes. Movement between schools and school type.

	Move within academic year	Move school	Public school	Voucher	Scientific-humanistic track	Fraction of incoming students over the median in previous school	First grade GPA higher than median classmates coming from first grade	GPA higher than median grade classmates evermoved	GPA higher than median classmates just moving IN to school	Average grade other classmates
1	-0.0106 [0.0108]		-0.0725*** [0.0256]	0.0231 [0.0221]			0.3043*** [0.0262]	0.1448*** [0.0142]		0.003 [0.0277]
2	-0.0018 [0.0162]	-0.0271 [0.0172]	-0.0755** [0.0297]	0.0012 [0.0218]		-0.0119 [0.0284]	0.2246*** [0.0296]	0.1641*** [0.0253]	0.1593*** [0.0157]	0.0736*** [0.0186]
3	-0.0066 [0.0107]	0.0071 [0.0227]	-0.0286 [0.0352]	0.0403 [0.0248]		0.006 [0.0143]	0.1370*** [0.0279]	0.1752*** [0.0242]	0.1571*** [0.0354]	0.0151 [0.0199]
4	0.012 [0.0082]	-0.0307* [0.0184]	-0.0814** [0.0349]	-0.0035 [0.0313]		0.0066 [0.0140]	0.1611*** [0.0265]	0.1242*** [0.0298]	0.0823*** [0.0240]	0.0062 [0.0211]
5	-0.0037 [0.0081]	0.0470*** [0.0173]	-0.0830** [0.0392]	0.0655** [0.0305]		0.0021 [0.0148]	0.1262*** [0.0204]	0.1184*** [0.0297]	0.1240*** [0.0348]	-0.0015 [0.0212]
6	-0.0052 [0.0105]	-0.0199 [0.0174]	-0.0722** [0.0337]	0.038 [0.0255]		0.0065 [0.0142]	0.1837*** [0.0201]	0.1169*** [0.0189]	0.0989*** [0.0219]	0.0015 [0.0223]
7	0.0029 [0.0060]	0.0031 [0.0246]	-0.0727** [0.0327]	0.038 [0.0296]		0.0138 [0.0114]	0.1611*** [0.0193]	0.0848*** [0.0209]	0.0650*** [0.0193]	0.0639*** [0.0199]
8	0.0177 [0.0134]	-0.0287 [0.0201]	-0.0823*** [0.0312]	0.0505* [0.0303]		0.0006 [0.0121]	0.1461*** [0.0319]	0.0790*** [0.0241]	0.0643*** [0.0268]	-0.0231 [0.0232]
9	0.023 [0.0176]	-0.0213 [0.0241]	-0.0648*** [0.0215]	0.0312 [0.0273]	0.1457*** [0.0249]	0.0241* [0.0126]	0.1056*** [0.0172]	0.0488** [0.0236]	0.0337 [0.0304]	-0.1122** [0.0449]
10	0.0240* [0.0133]	0.0627*** [0.0199]	-0.0629** [0.0246]	0.0355 [0.0284]	0.1365*** [0.0263]	0.0347** [0.0158]	0.1116*** [0.0158]	0.0699** [0.0273]	0.0769*** [0.0281]	0.0484 [0.0462]
11	0.0087 [0.0172]	-0.1395*** [0.0232]	-0.0548** [0.0242]	0.0131 [0.0256]	0.0947*** [0.0256]	0.0449*** [0.0115]	0.0716*** [0.0169]	0.0751*** [0.0215]	-0.0056 [0.0212]	0.1129** [0.0472]

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at birth day level.

Other covariates in the specification are dummies for male, rural status and whether or not the child started school in the same municipality where she/he resided.

Table 9: Impact of age entry on selected outcomes. RD Design using just 1st of July discontinuity.

	Attendance	Attendance over media	GPA	GPA over p90	GPA over p75	GPA over p50	GPA over p25	GPA over p10	Pass
1	2.2242** [0.8964]	0.0945*** [0.0219]	0.4981*** [0.0557]	0.1976*** [0.0228]	0.3022*** [0.0241]	0.3086*** [0.0246]	0.1913*** [0.0228]	0.0572*** [0.0146]	0.0183* [0.0110]
2	0.7715 [0.7311]	-0.0058 [0.0268]	0.3891*** [0.0526]	0.1470*** [0.0143]	0.2046*** [0.0206]	0.1878*** [0.0262]	0.1335*** [0.0176]	0.0516*** [0.0124]	0.0516*** [0.0110]
3	1.2867** [0.5739]	0.0363 [0.0296]	0.2960*** [0.0452]	0.1620*** [0.0189]	0.2171*** [0.0271]	0.1670*** [0.0251]	0.0907*** [0.0226]	0.0720*** [0.0148]	0.0466*** [0.0099]
4	-0.0465 [0.5396]	-0.0299 [0.0237]	0.2007*** [0.0420]	0.1056*** [0.0202]	0.1658*** [0.0267]	0.1338*** [0.0252]	0.0899*** [0.0245]	0.0053 [0.0135]	0.0159* [0.0082]
5	-0.0428 [0.4156]	0.0285 [0.0279]	0.1629*** [0.0379]	0.0844*** [0.0188]	0.1185*** [0.0382]	0.1166*** [0.0268]	0.0643*** [0.0222]	0.0145 [0.0129]	0.007 [0.0116]
6	-0.3886 [0.6013]	0.0426 [0.0303]	0.2004*** [0.0429]	0.0847*** [0.0194]	0.1417*** [0.0182]	0.1534*** [0.0175]	0.0532*** [0.0155]	0.0420*** [0.0121]	0.0241** [0.0100]
7	-0.0692 [0.5569]	0.0672** [0.0280]	0.1496*** [0.0356]	0.0430*** [0.0152]	0.0888*** [0.0201]	0.0866*** [0.0289]	0.0544*** [0.0149]	0.0285** [0.0124]	0.0316** [0.0138]
8	0.1494 [0.6277]	0.0187 [0.0381]	0.0778** [0.0339]	0.0713*** [0.0124]	0.1085*** [0.0239]	0.0932*** [0.0226]	0.0314* [0.0183]	0.0049 [0.0133]	-0.0015 [0.0112]
9	-3.7819*** [1.0063]	0.0444* [0.0251]	0.0386 [0.0499]	0.0667** [0.0257]	0.0952*** [0.0248]	0.0592** [0.0242]	0.0284* [0.0159]	0.0333** [0.0152]	-0.0324** [0.0146]
10	1.0192 [1.1092]	-0.0326 [0.0253]	0.1127* [0.0596]	0.0673*** [0.0250]	0.0613*** [0.0221]	0.0546*** [0.0189]	0.011 [0.0180]	0.0112 [0.0094]	0.017 [0.0181]
11	-0.7712 [1.6926]	0.0291 [0.0305]	0.1435 [0.0873]	0.0438** [0.0204]	0.0593*** [0.0209]	0.0592*** [0.0170]	0.0313 [0.0202]	0.0073 [0.0090]	0.0224 [0.0175]

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at birth day level.

Other covariates in the specification are dummies for male, rural status and whether or not the child started school in the same municipality where she/he resided.

Table 10: Impact of age entry on selected outcomes. RD Design using just 1st of July discontinuity.

	Move within academic year	Move school	Public school	Voucher	Scientific-humanistic track	Fraction of incoming students over the median in previous school	First grade GPA higher than median classmates coming from first grade	GPA higher than median grade classmates evermoved	GPA higher than median classmates just moving IN to school	Average grade other classmates
1	-0.01 [0.0119]		-0.0688** [0.0276]	0.0174 [0.0232]			0.2946*** [0.0265]	0.1445*** [0.0154]	0.2642*** [0.0207]	0.0117 [0.0305]
2	-0.0063 [0.0178]	-0.0329* [0.0168]	-0.0635** [0.0320]	-0.008 [0.0217]		-0.0089 [0.0306]	0.2140*** [0.0300]	0.1680*** [0.0269]	0.1589*** [0.0149]	0.0774*** [0.0188]
3	-0.001 [0.0104]	0.0031 [0.0243]	-0.0141 [0.0388]	0.0298 [0.0244]		0.0115 [0.0149]	0.1345*** [0.0288]	0.1673*** [0.0240]	0.1381*** [0.0385]	0.0095 [0.0195]
4	0.0115 [0.0083]	-0.0282 [0.0195]	-0.0720* [0.0377]	-0.0107 [0.0335]		0.0105 [0.0154]	0.1559*** [0.0279]	0.1188*** [0.0319]	0.0719*** [0.0257]	0.0071 [0.0219]
5	-0.0054 [0.0081]	0.0295* [0.0151]	-0.0737* [0.0433]	0.0621* [0.0332]		0.0142 [0.0164]	0.1198*** [0.0208]	0.0972*** [0.0320]	0.1021*** [0.0377]	-0.0048 [0.0217]
6	0.0001 [0.0111]	-0.0239 [0.0166]	-0.0723* [0.0381]	0.0376 [0.0278]		0.0191 [0.0155]	0.1733*** [0.0214]	0.1128*** [0.0185]	0.1035*** [0.0220]	0.0054 [0.0242]
7	0.0053 [0.0060]	0.0097 [0.0260]	-0.0625* [0.0346]	0.025 [0.0296]		0.0222* [0.0124]	0.1552*** [0.0186]	0.0701*** [0.0226]	0.0532*** [0.0191]	0.0624*** [0.0204]
8	0.0214 [0.0146]	-0.0221 [0.0201]	-0.0633** [0.0318]	0.0299 [0.0302]		0.0054 [0.0127]	0.1521*** [0.0338]	0.0708*** [0.0242]	0.0396 [0.0259]	-0.0199 [0.0250]
9	0.0263 [0.0191]	-0.0078 [0.0228]	-0.0538*** [0.0203]	0.0188 [0.0283]	0.1447*** [0.0261]	0.0230* [0.0122]	0.1064*** [0.0173]	0.0453* [0.0253]	0.0356 [0.0329]	-0.1174** [0.0452]
10	0.0217* [0.0131]	0.0630*** [0.0174]	-0.0519** [0.0244]	0.0234 [0.0294]	0.1465*** [0.0292]	0.0378** [0.0167]	0.1034*** [0.0138]	0.0598** [0.0276]	0.0806*** [0.0280]	0.0553 [0.0508]
11	0.0081 [0.0160]	-0.1463*** [0.0212]	-0.0550** [0.0245]	0.0115 [0.0263]	0.1050*** [0.0252]	0.0479*** [0.0117]	0.0673*** [0.0164]	0.0689*** [0.0218]	-0.0152 [0.0210]	0.1188** [0.0487]

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at birth day level.

Other covariates in the specification are dummies for male, rural status and whether or not the child started school in the same municipality where she/he resided.

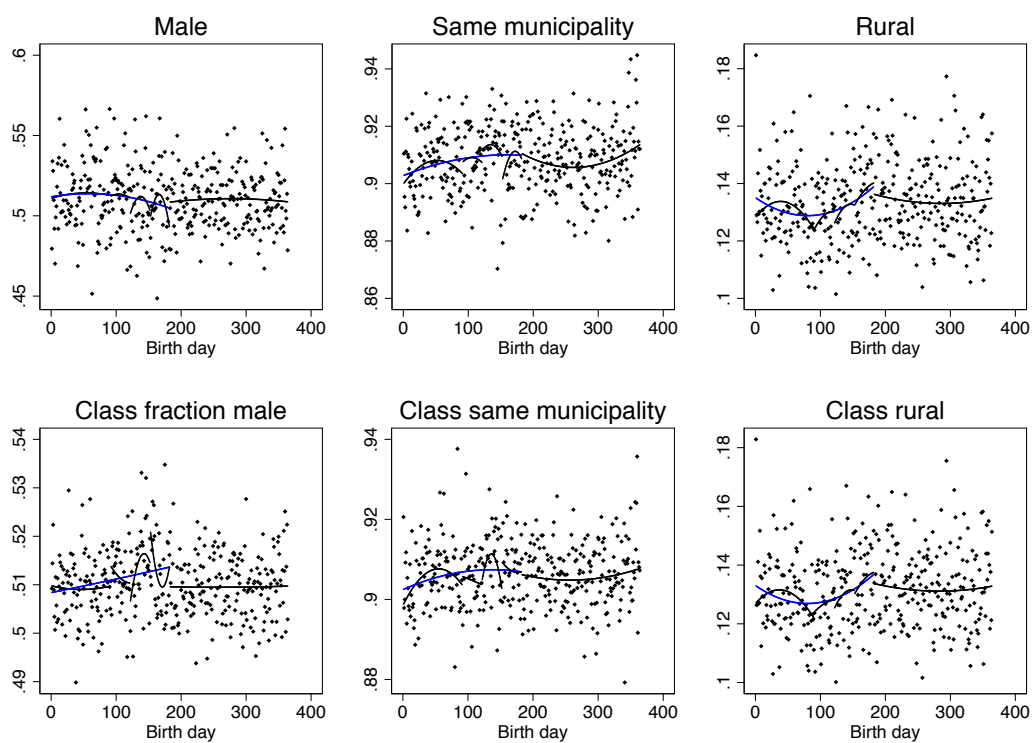


Figure 1: Balancing covariates

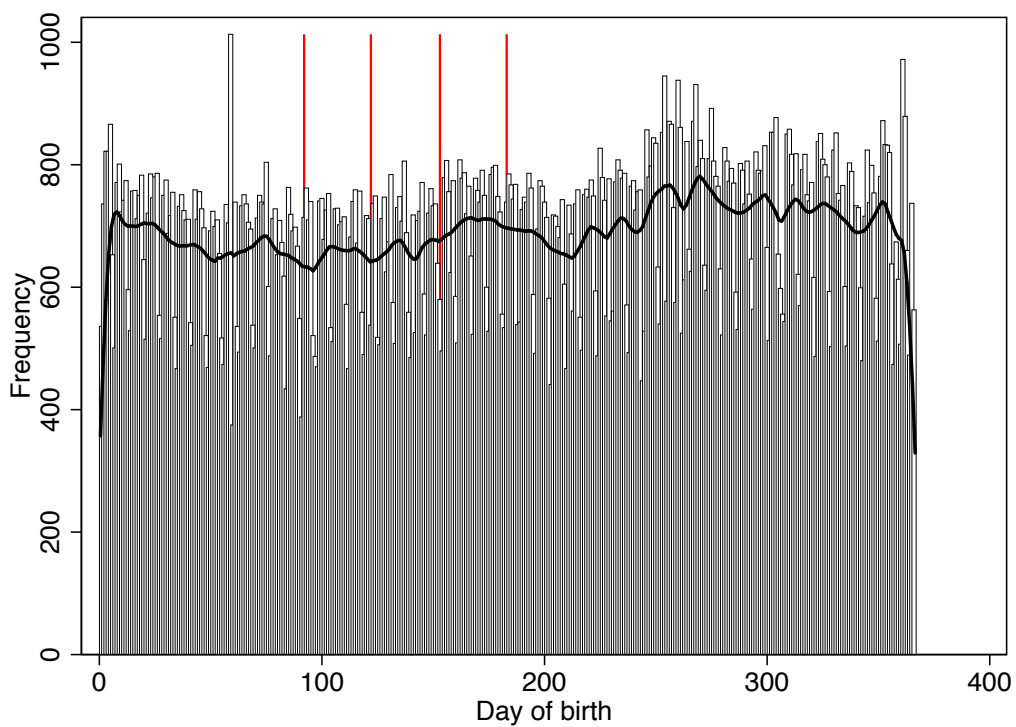


Figure 2: Histogram day of birth.

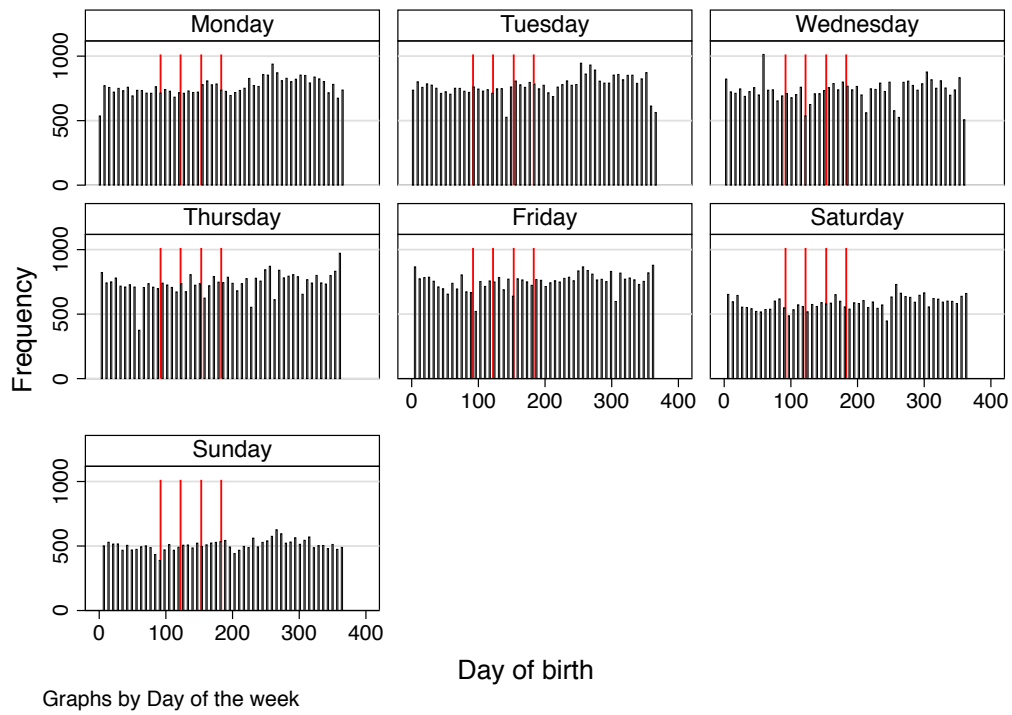


Figure 3: Histogram day of birth by week days

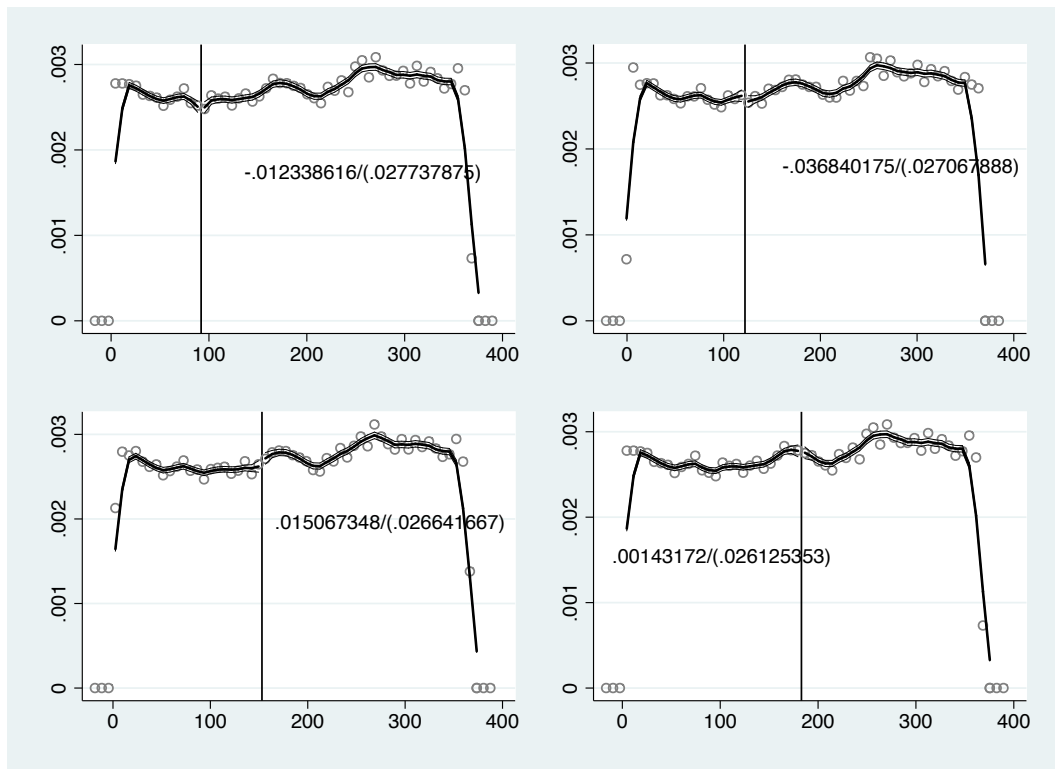


Figure 4: Potential discontinuity in day of birth

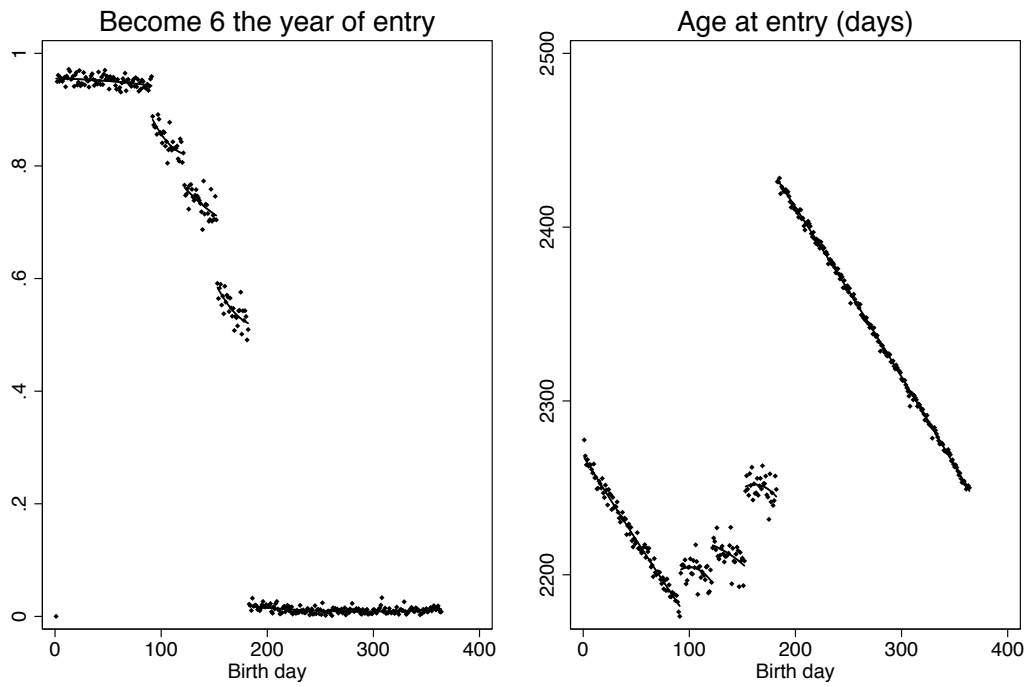


Figure 5: Age of entry and minimum age entry rule.

Attendance

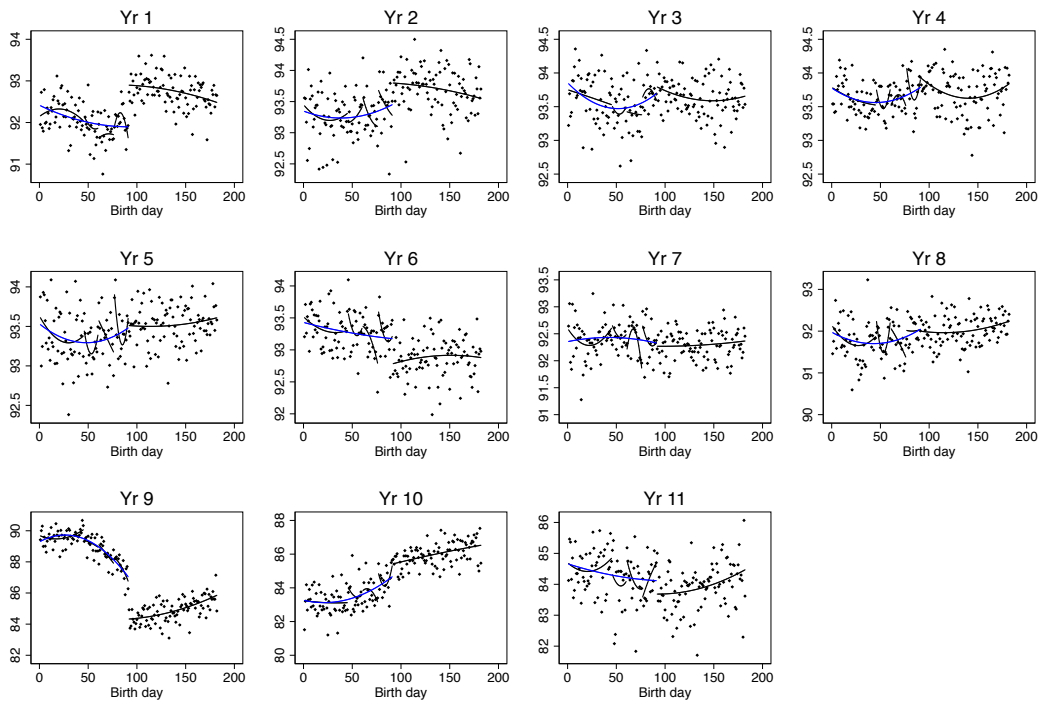


Figure 6: Attendance by day of birth.

Attendance over the median

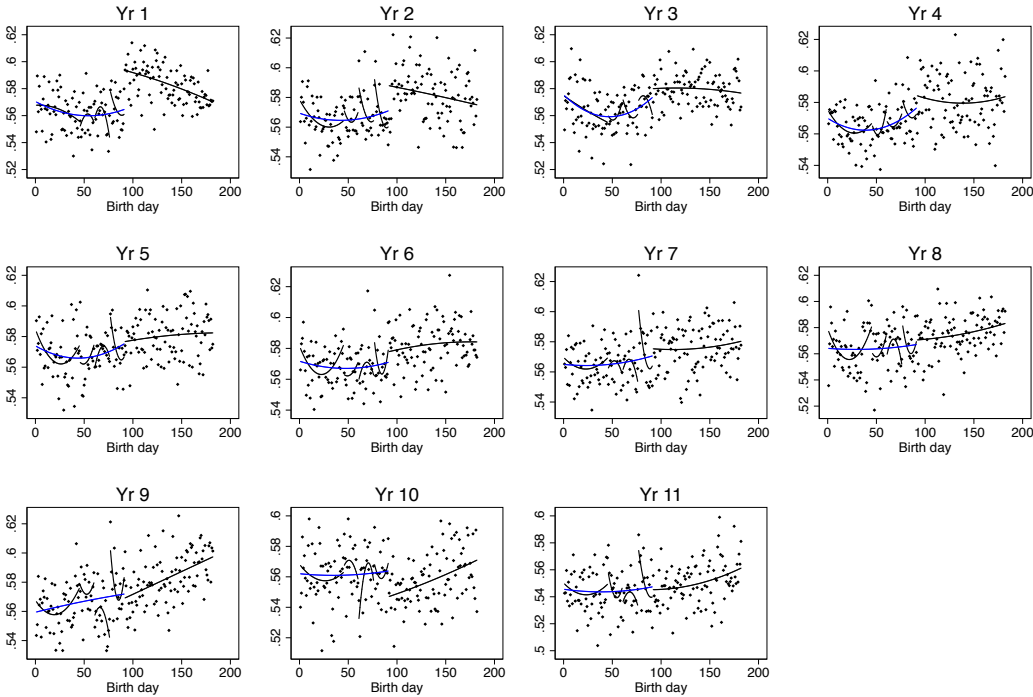


Figure 7: Attendance over the median in the class by day of birth.

GPA

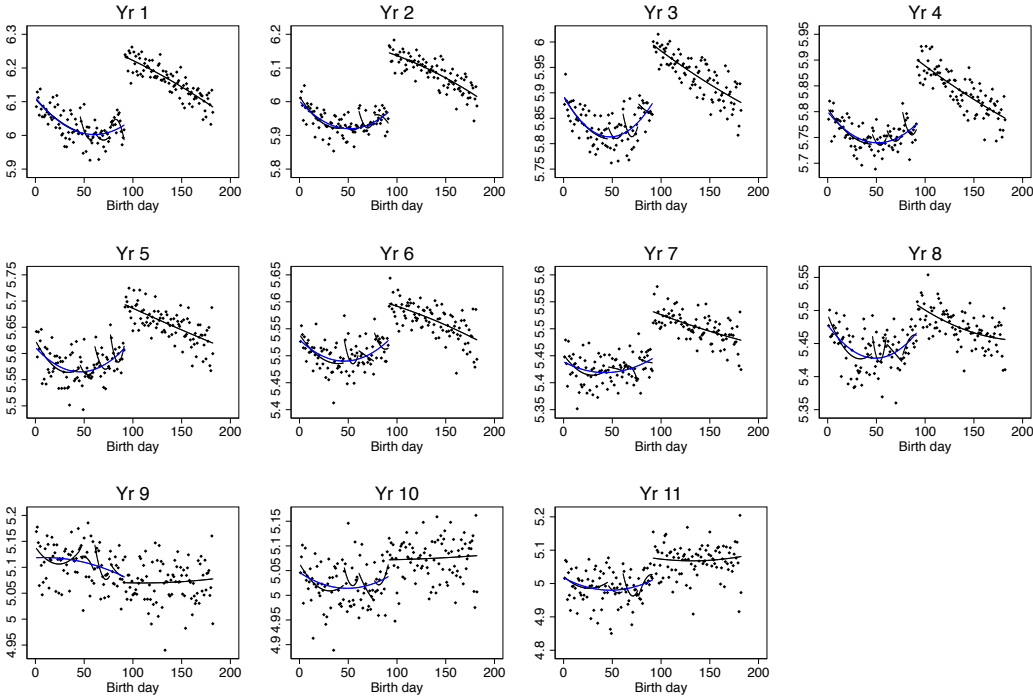


Figure 8: GPA by day of birth.

GPA over p90

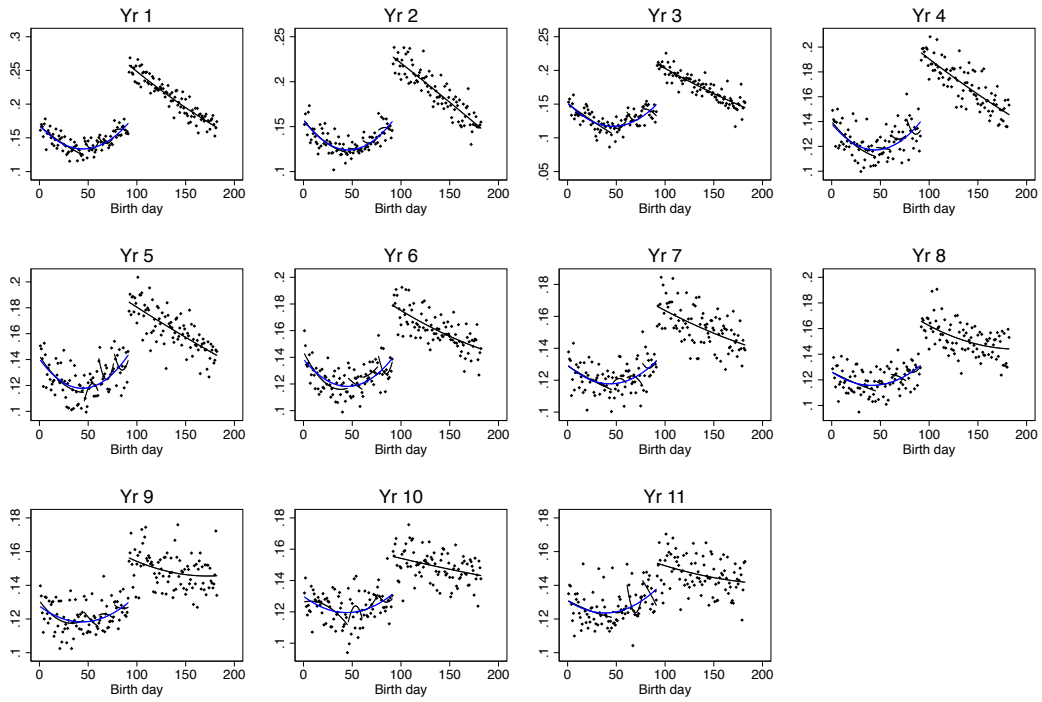


Figure 9: GPA over 90th percentile by day of birth.

GPA over p50

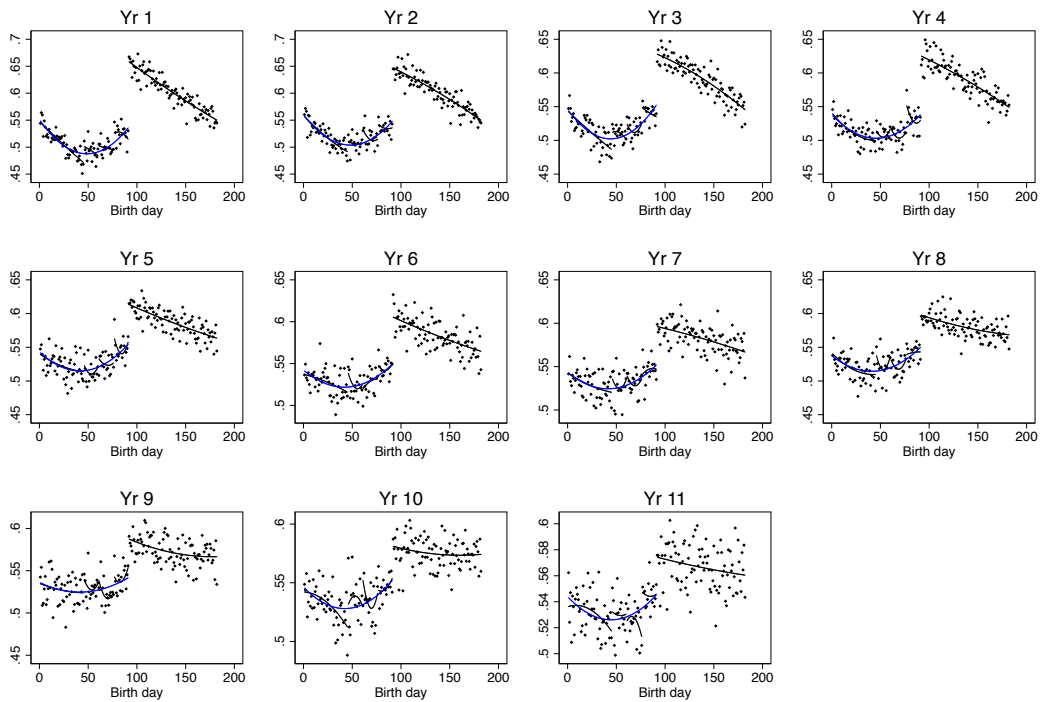


Figure 10: GPA over 50th percentile by day of birth.

Pass

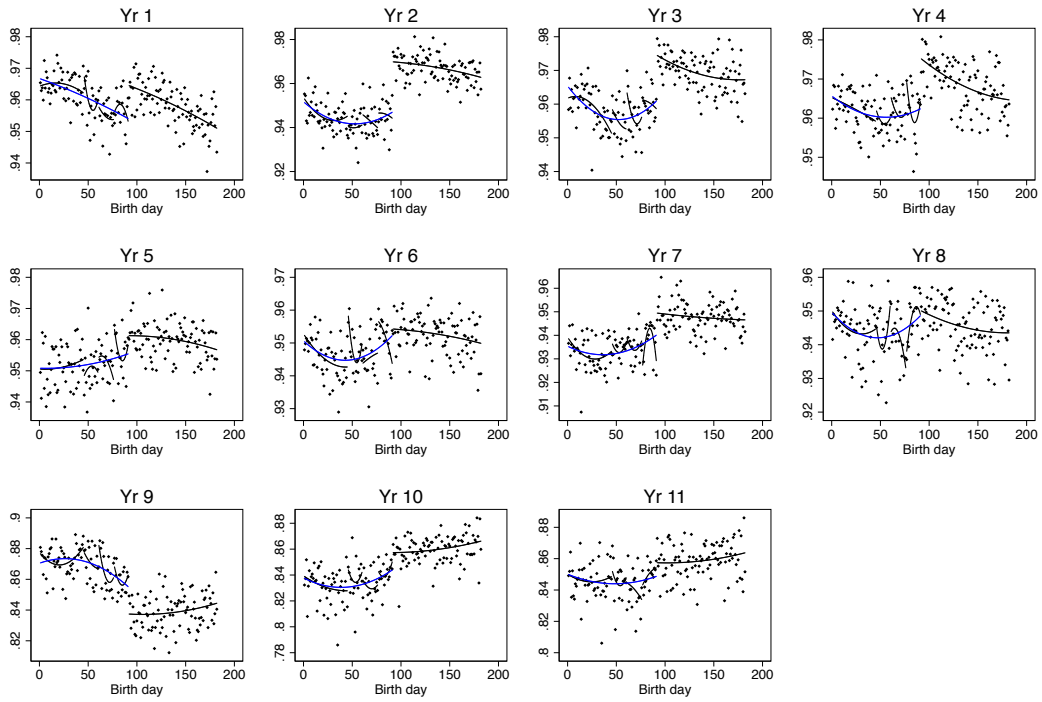


Figure 11: Pass course by day of birth.

Move within academic year

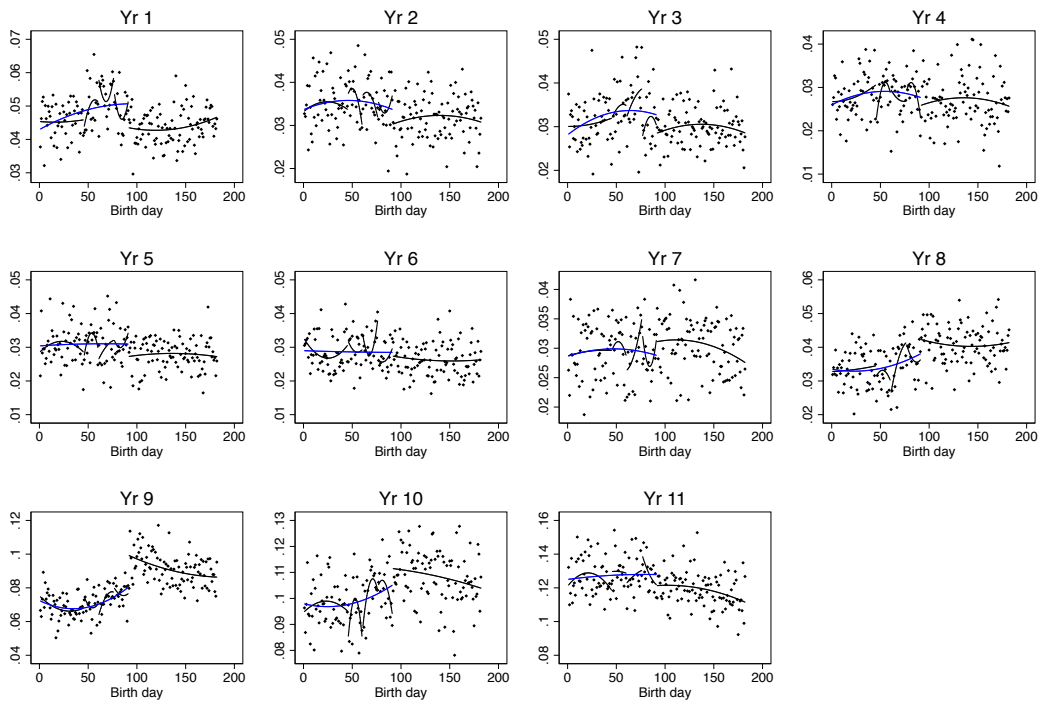


Figure 12: Move school within school year by day of birth.

Move school

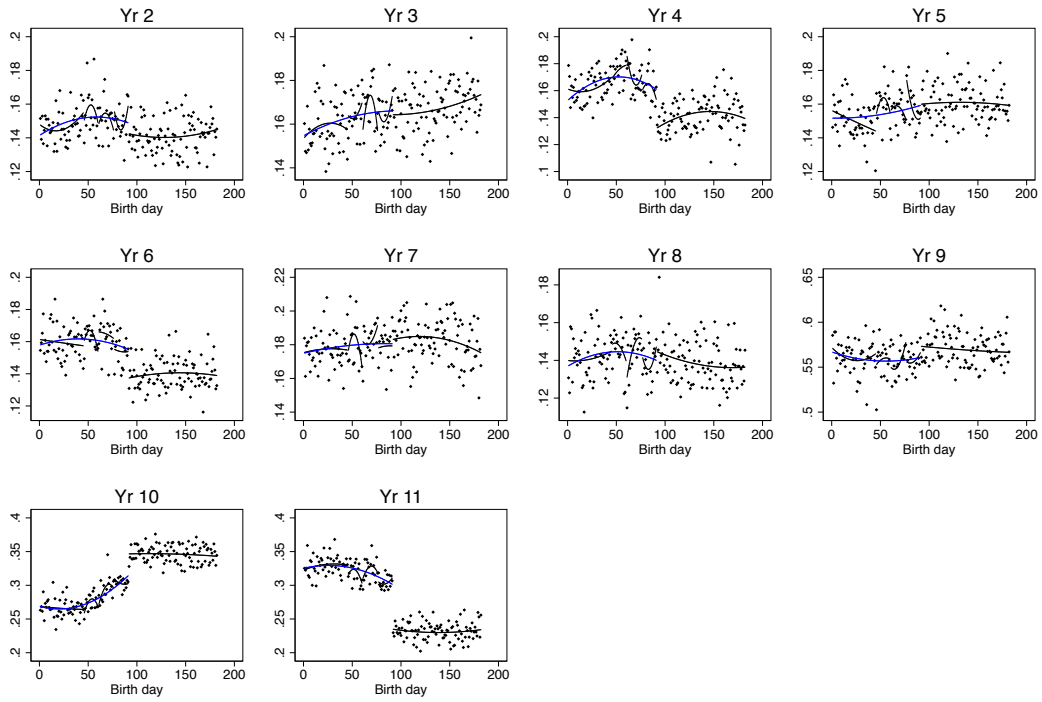


Figure 13: Move school by day of birth.

Public school

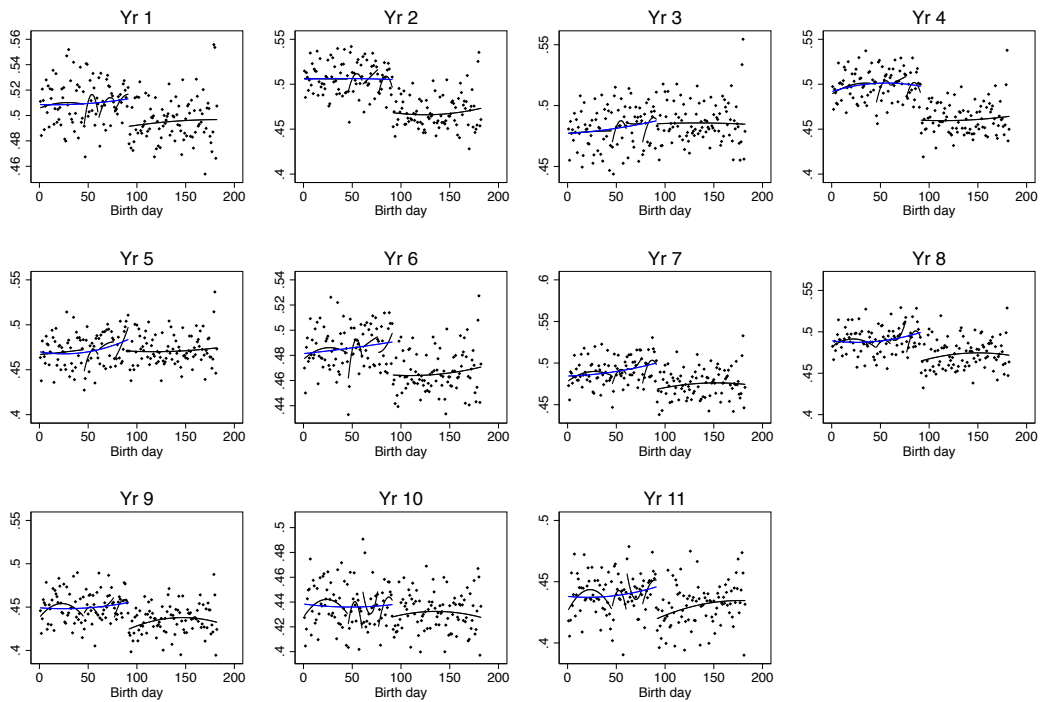


Figure 14: Public school by day of birth.

Voucher

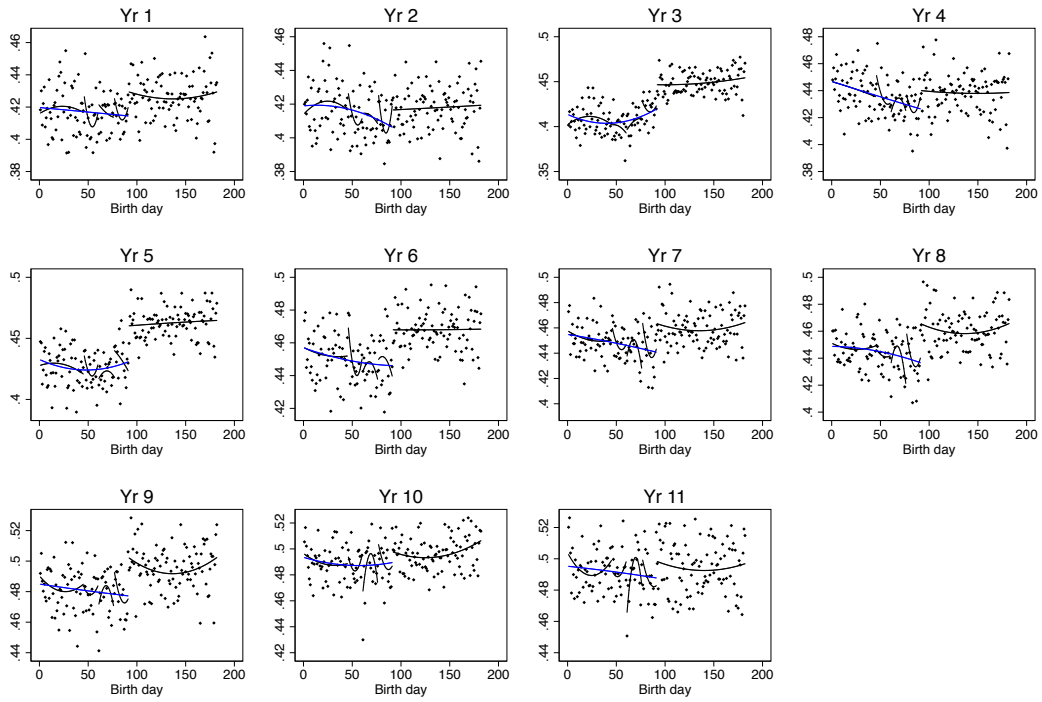


Figure 15: Voucher school by day of birth.

Scientific-humanistic track

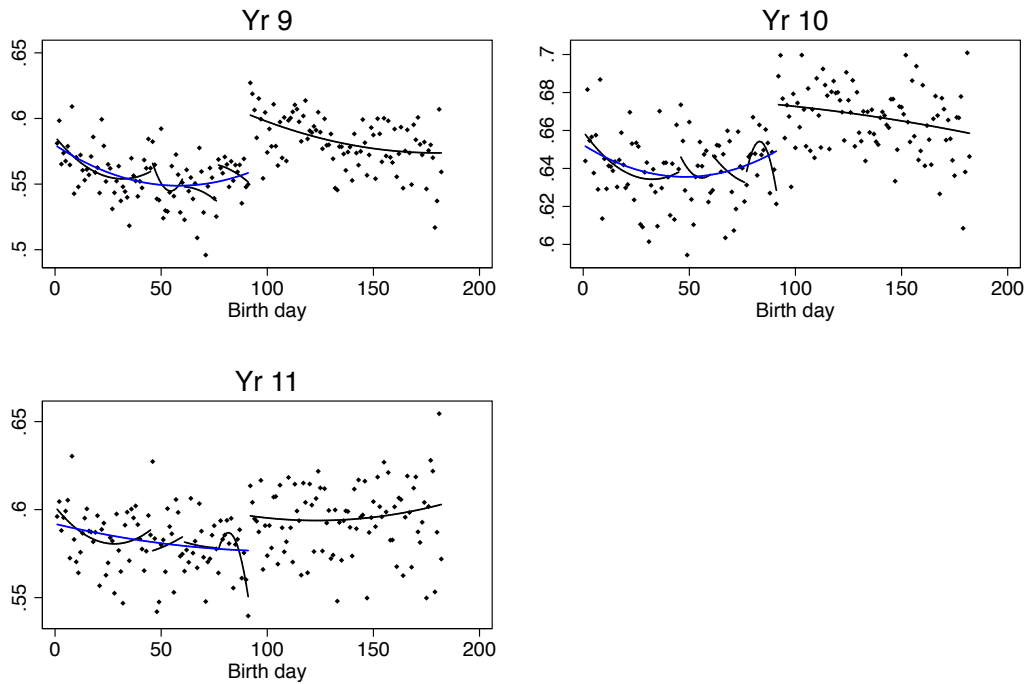


Figure 16: Scientific or humanistic track by day of birth.

Frac. of incoming students over the median in previous school

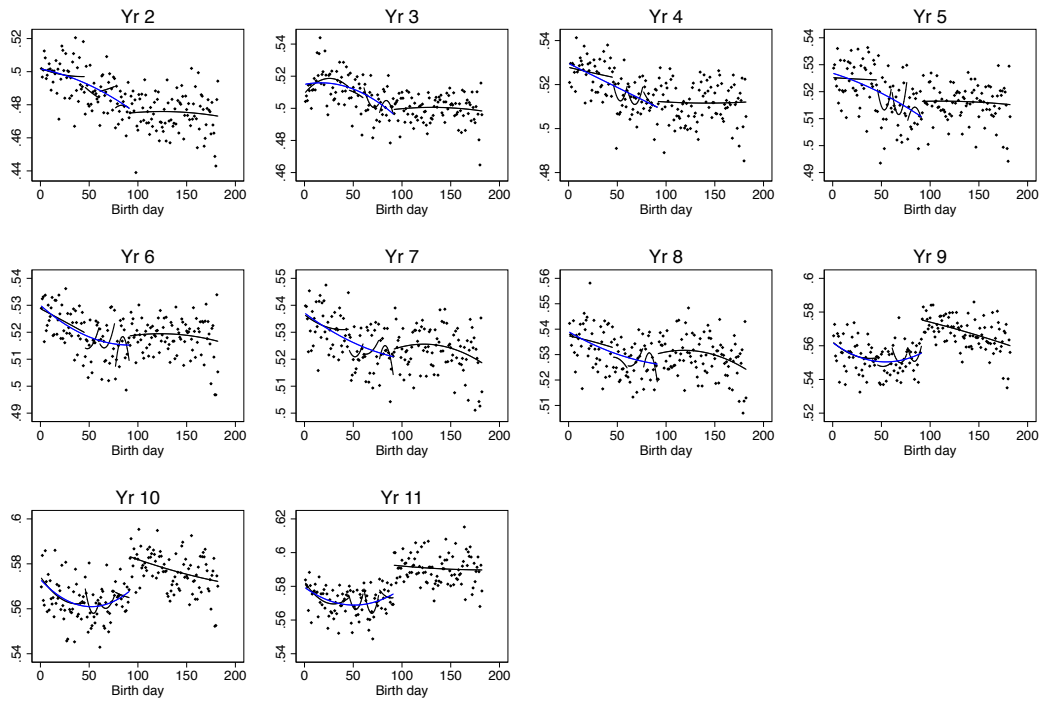


Figure 17: Fraction of incoming students over the median in previous school by day of birth.

GPA (1stG) higher than median classmates coming from 1stG

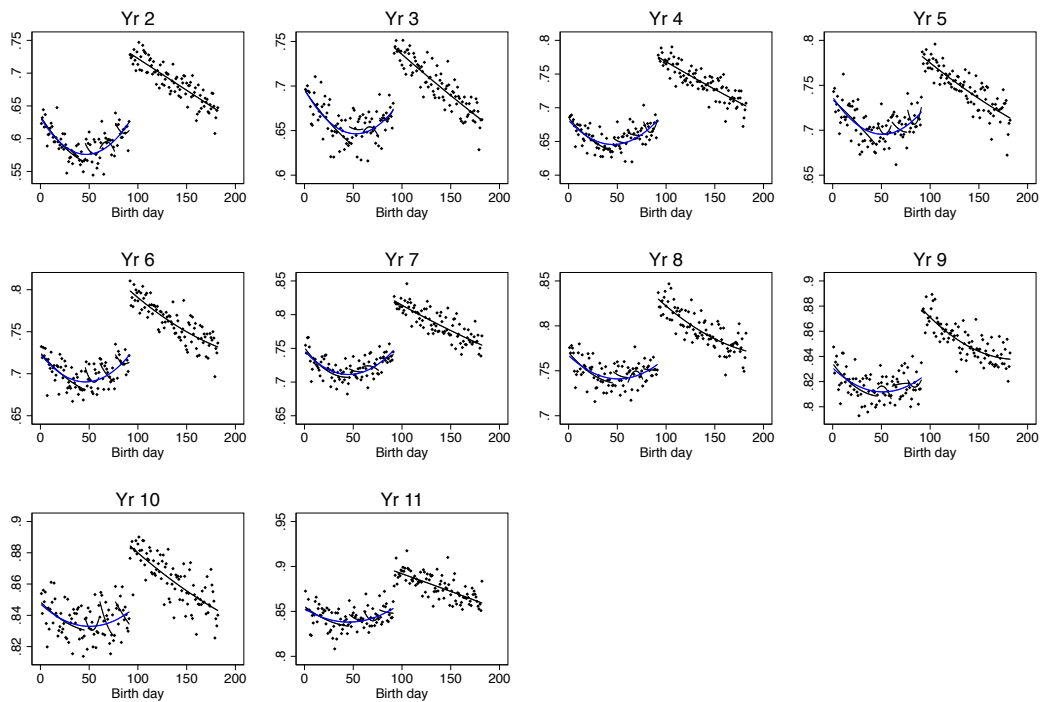


Figure 18: Fraction of students with a GPA in 1st grade higher than median classmates coming from 1st grade by day of birth.

GPA higher than median grade classmates evermoved

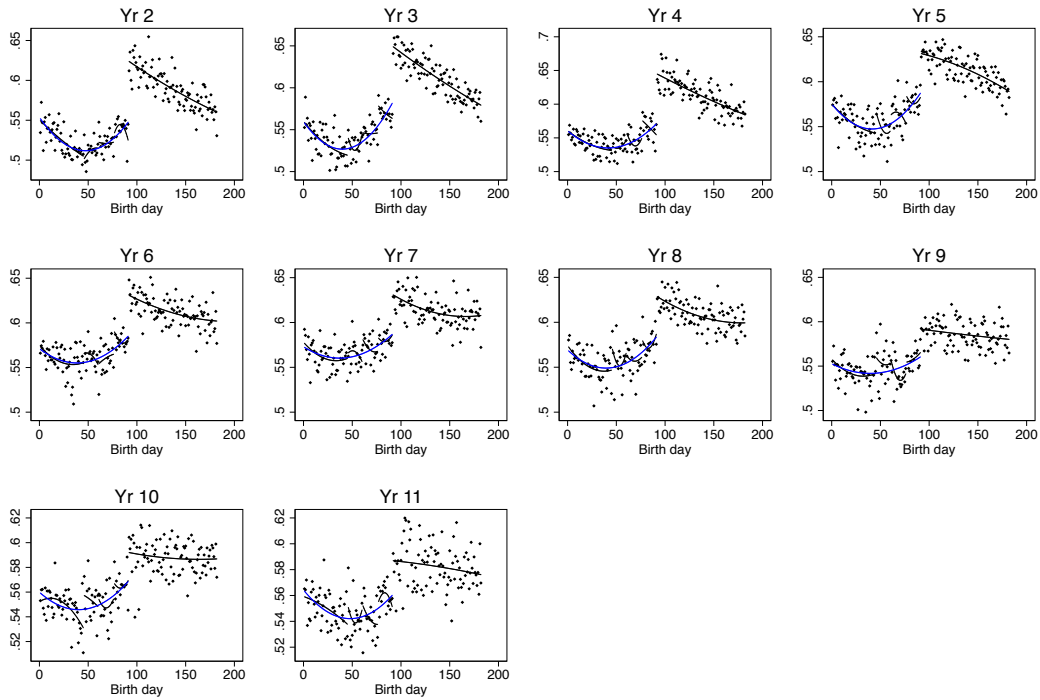


Figure 19: Fraction of students with GPA higher than median grade of their classmates who have ever moved by day of birth.

GPA higher than median classmates just moving IN to school

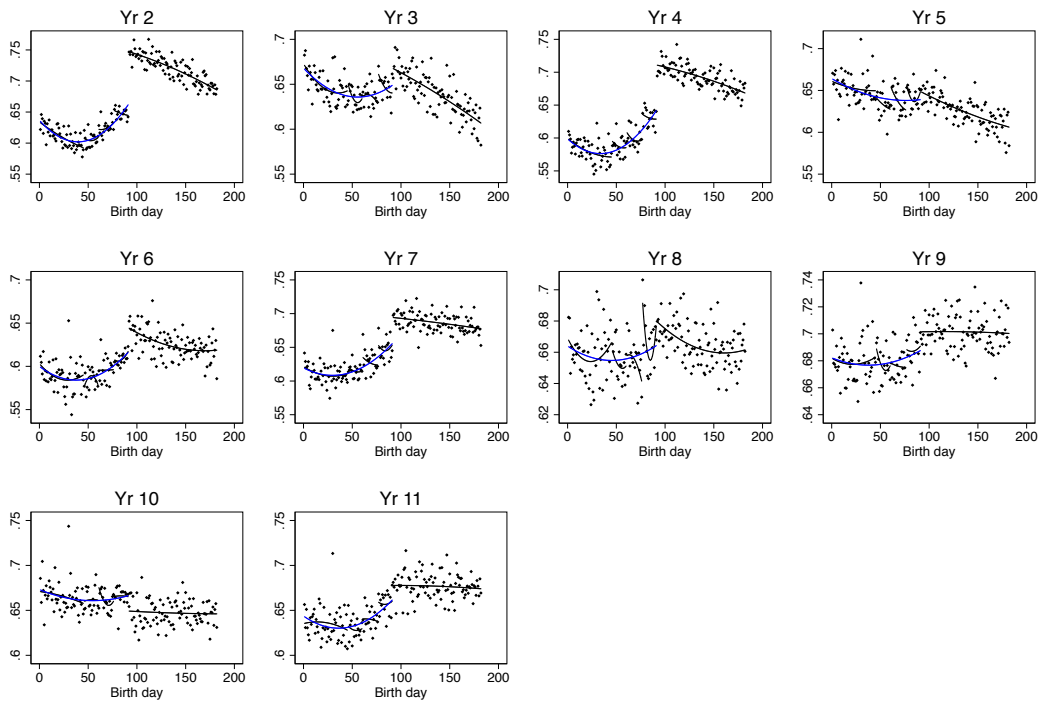


Figure 20: Fraction of students with GPA higher than median grade of their classmates just moving into the school by day of birth.

Average GPA rest of the classmates

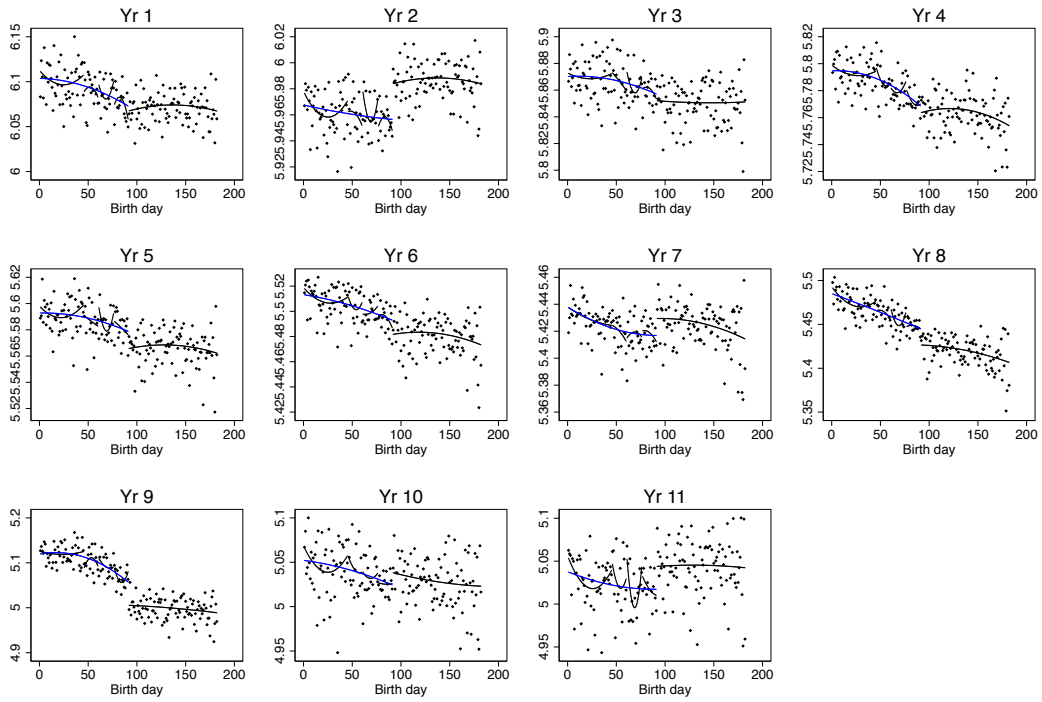


Figure 21: Average GPA of the other classmates by day of birth.

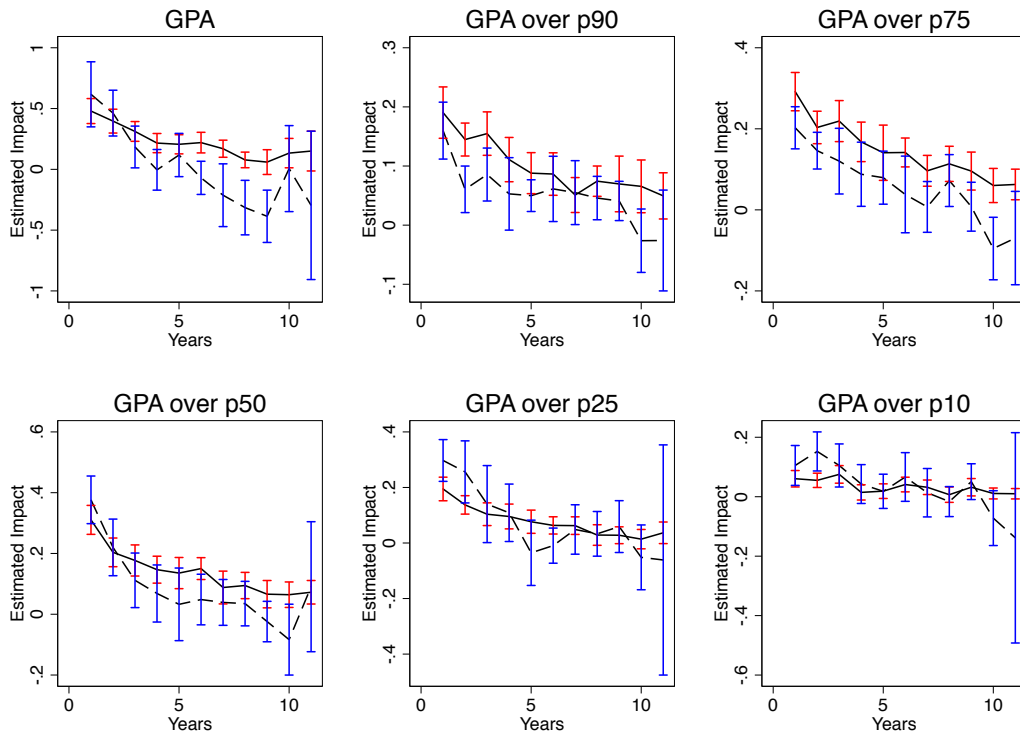
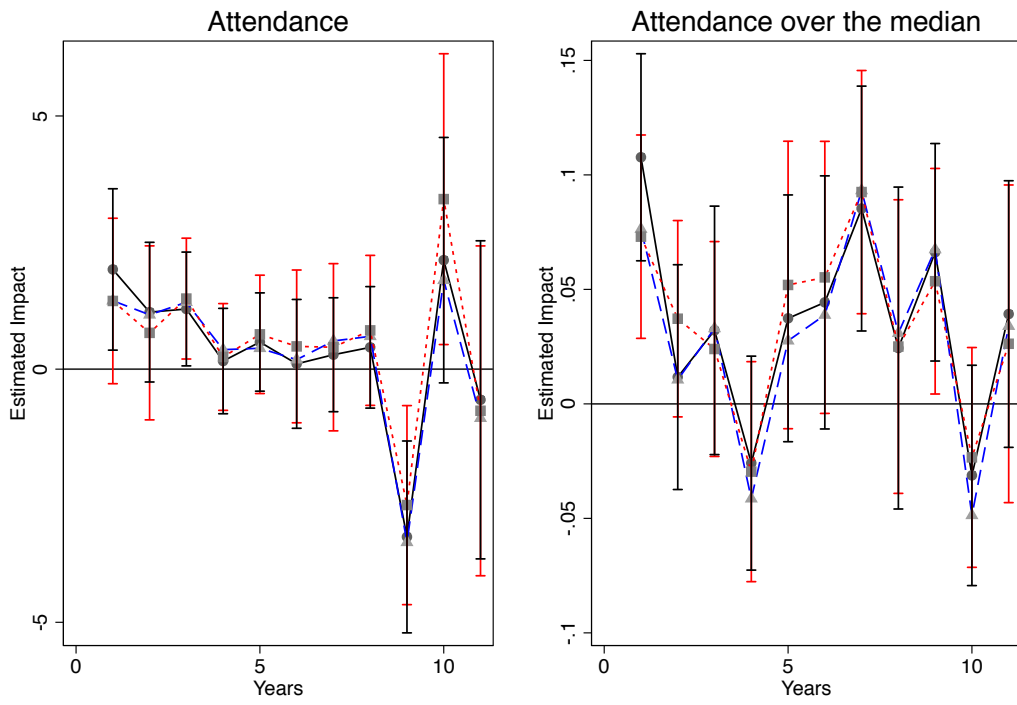
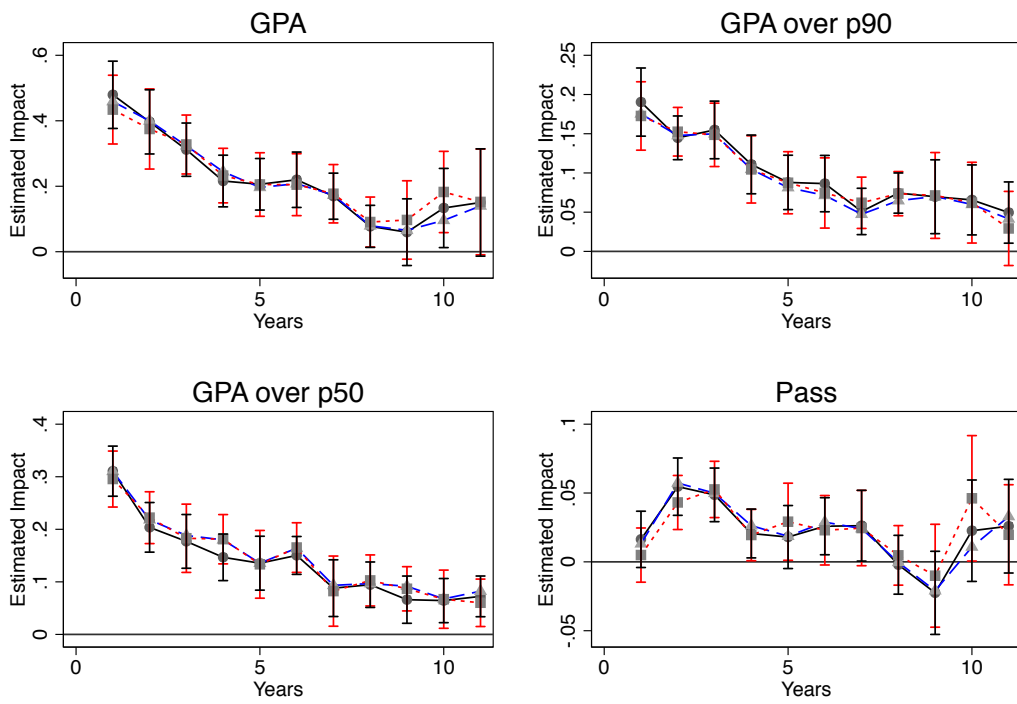


Figure 22: Effect of sample restriction on the impact of age of entry on GPA



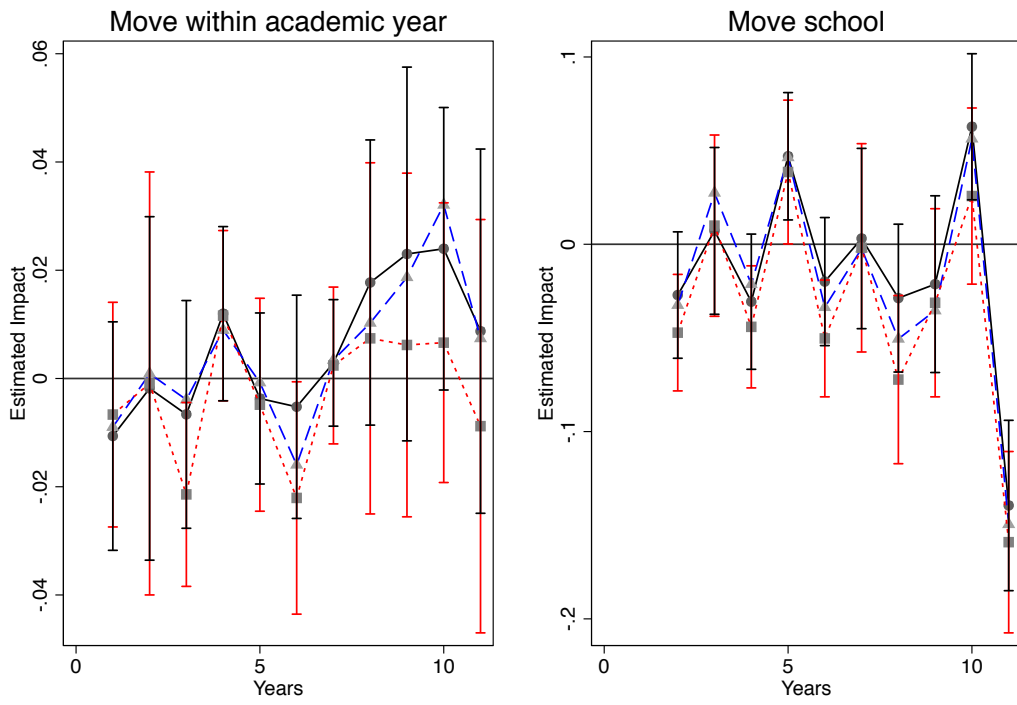
Black, blue, and red lines correspond to the 15, 10, and 7 days bandwidth respectively.

Figure 23: Sensitivity of the results to the Bandwidth. Attendance



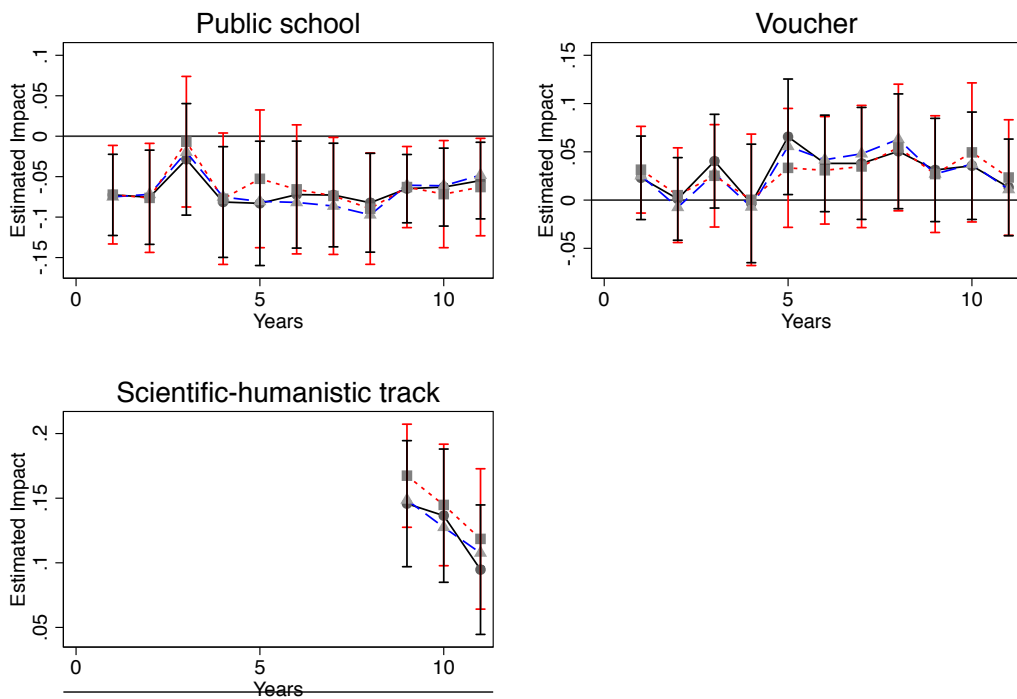
Black, blue, and red lines correspond to the 15, 10, and 7 days bandwidth respectively.

Figure 24: Sensitivity of the results to the Bandwidth. School achievements



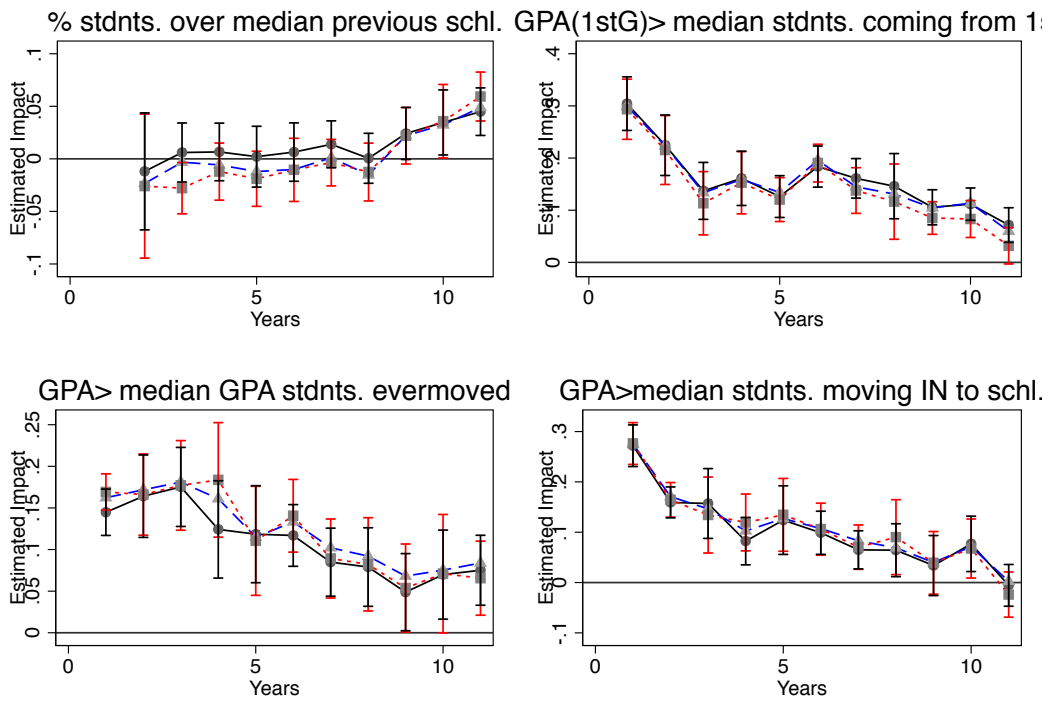
Black, blue, and red lines correspond to the 15, 10, and 7 days bandwidth respectively.

Figure 25: Sensitivity of the results to the Bandwidth. Move school



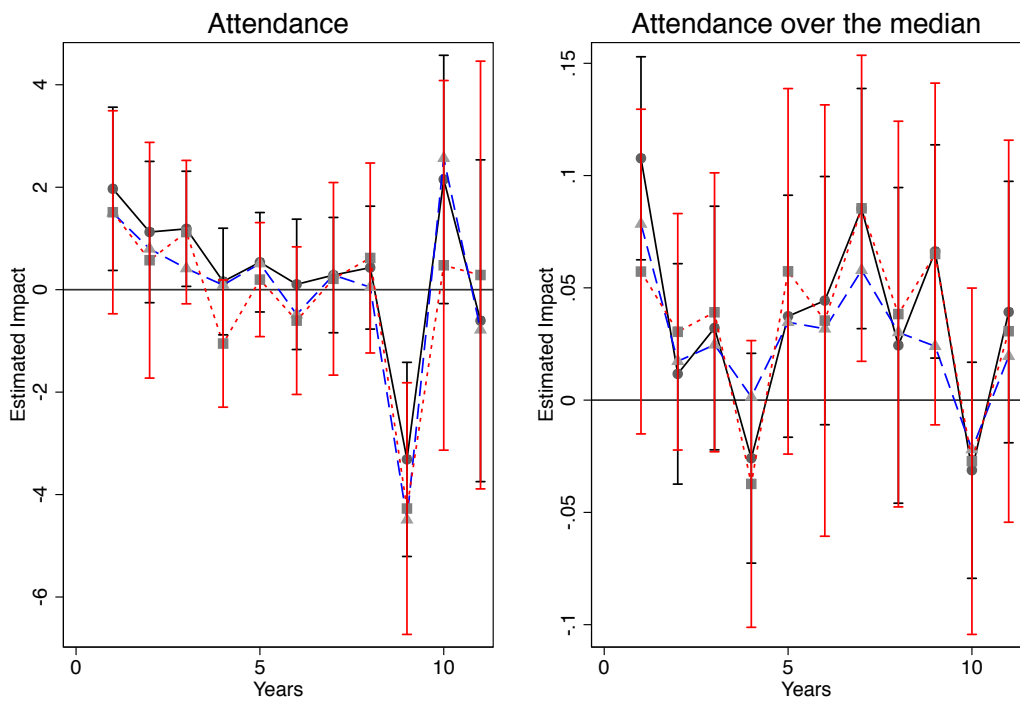
Black, blue, and red lines correspond to the 15, 10, and 7 days bandwidth respectively.

Figure 26: Sensitivity of the results to the Bandwidth. School type.



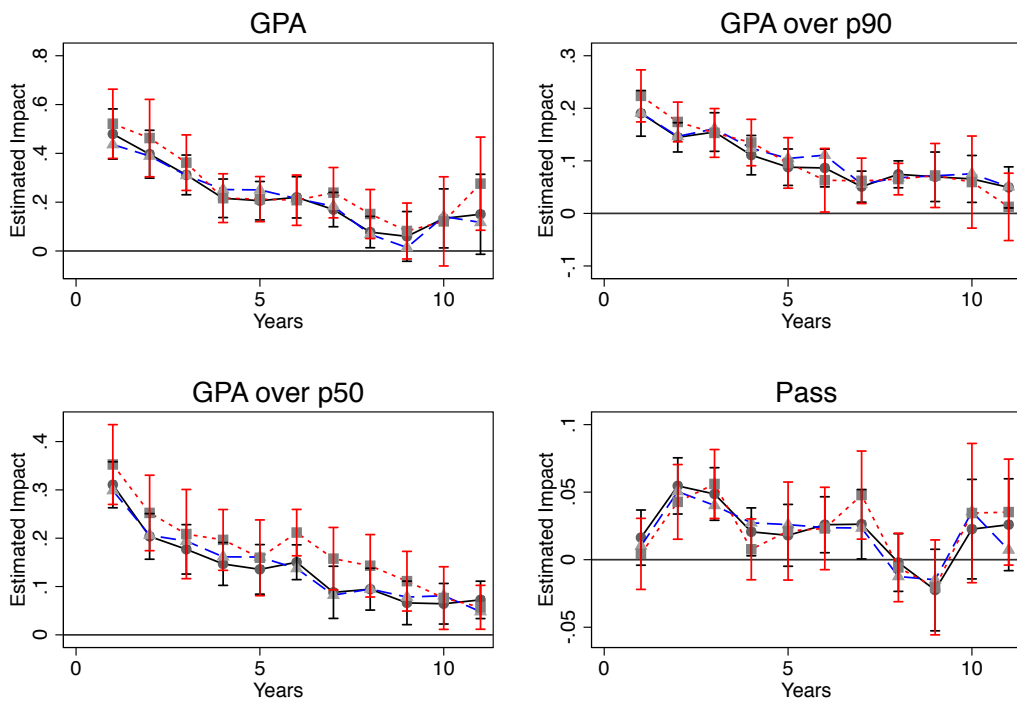
Black, blue, and red lines correspond to the 15, 10, and 7 days bandwidth respectively.

Figure 27: Sensitivity of the results to the Bandwidth. School cream skimming.



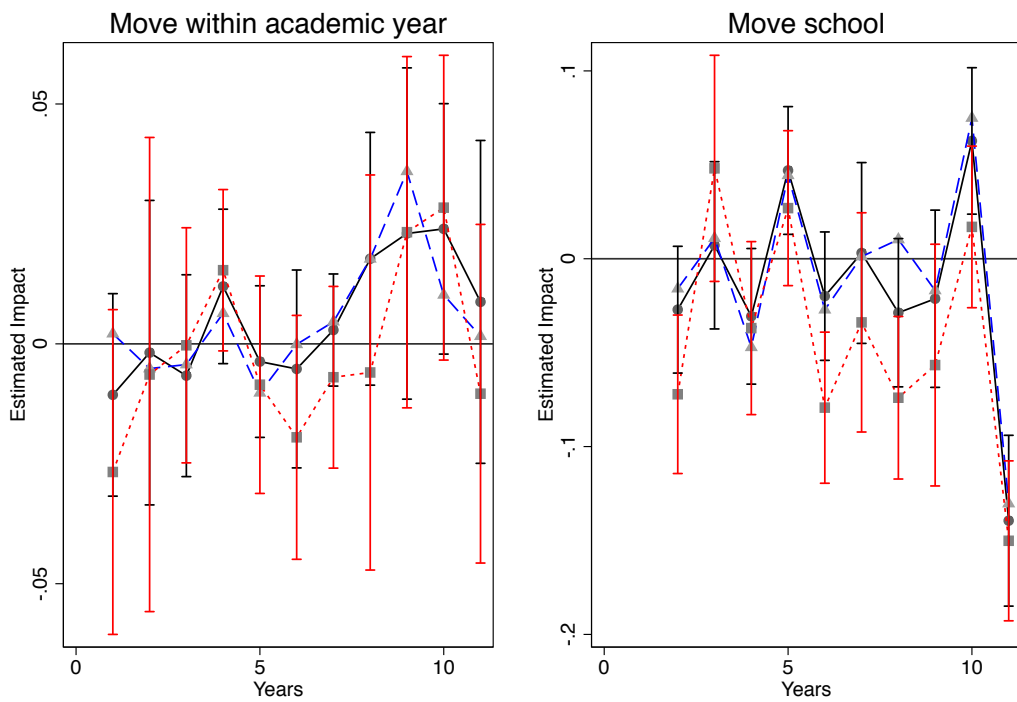
Blue, black, and red lines correspond to the polynomial specification of first, second, and third degree respectively

Figure 28: Sensitivity of the results to the Polynomial's degree. Attendance.



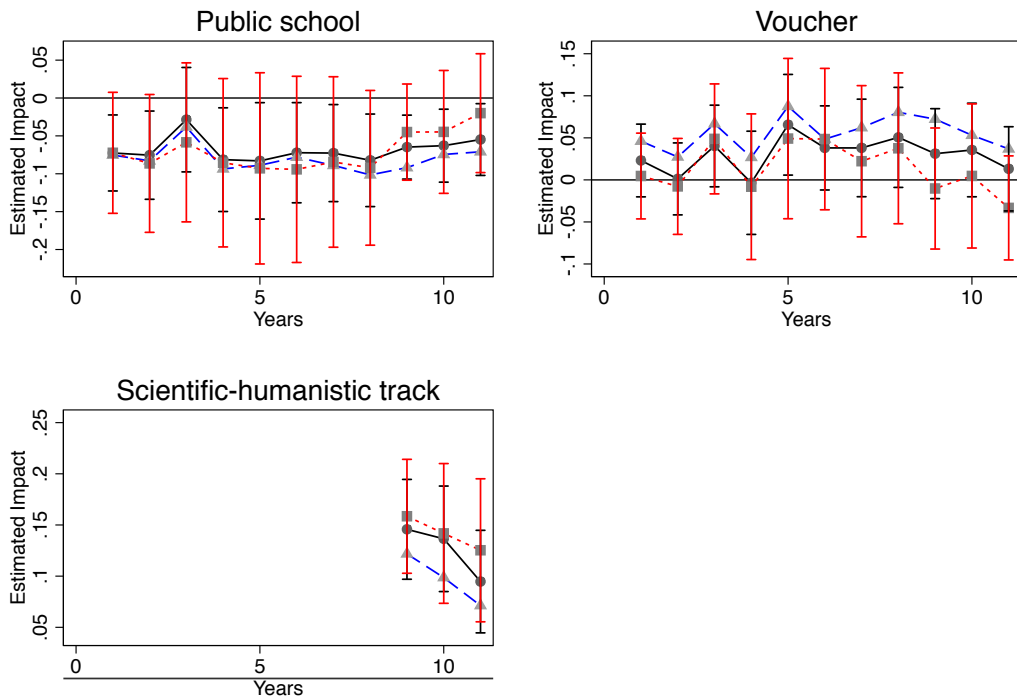
Blue, black, and red lines correspond to the polynomial specification of first, second, and third degree respective

Figure 29: Sensitivity of the results to the Polynomial's degree. School achievements.



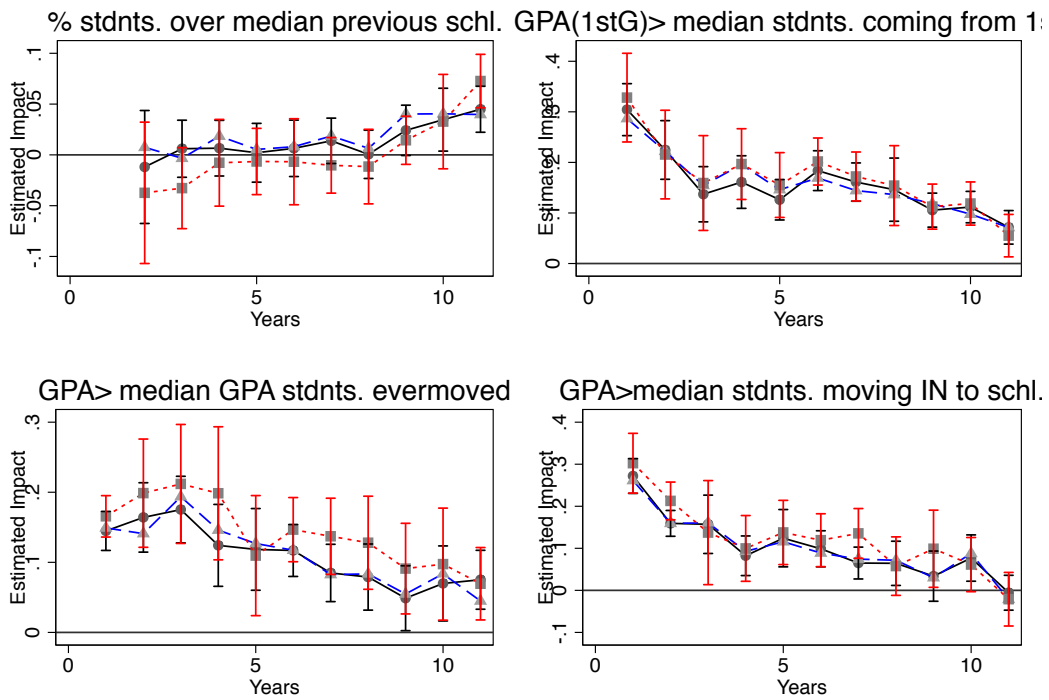
Blue, black, and red lines correspond to the polynomial specification of first, second, and third degree respective

Figure 30: Sensitivity of the results to the Polynomial's degree. Move school.



Blue, black, and red lines correspond to the polynomial specification of first, second, and third degree respective

Figure 31: Sensitivity of the results to the Polynomial's degree. School type.



Blue, black, and red lines correspond to the polynomial specification of first, second, and third degree respective

Figure 32: Sensitivity of the results to the Polynomial's degree. Cream skimming.