Long-run consequences of Parents' Support on schooling decisions[1]

Patrizia Ordine, Giuseppe Rose, Gessica Vella[2]

Abstract. The application of law 30/2003 in Italy, introduces new flexible age requirements, allowing children whose sixth birthday falls on or before April the 30th of the given scholastic year to enroll in the first year of primary school.

In addition to classes composition variation, in terms of more or less "ready for school" children, the new agreements provide parents with further responsibility linked to the choice of early entry option. Our research question develops in this scenario, with the aim to assess whether early entry options have long-term effects on students' scholastic choices given that more or less apprehensive parents have in turn had the opportunity to choose on their children enrollment. Exploiting an RDD setup, we show that given early entry option and parents' support, teenagers students are more likely to be tracked in Lyceum rather than Technical or Vocational schools, and above all, that the jump is higher for female students.

[1] Very Preliminary

[2] Ordine, P.: Department of Business and Legal Sciences, University of Calabria, Calabria, Italy.
Rose, G., Vella, G.: Department of Economics, Statistics and Finance, University of Calabria, Calabria, Italy.
E-mails: patrizia.ordine@unical.it; Giuseppe.rose@unical.it; gessica.vella@unical.it

1. Introduction and literature

The application of law 30/2003 brings to significant changes in the Italian school system, allowing children whose third/sixth birthday falls on or before April the 30th of the given scholastic year to enroll in the first year of nursery or primary school (previously, the cut-off date was 31st December).

These flexible age requirements, lead to two-sided aftermaths. On the one hand there is an increase of learning group inequalities relative to new heterogeneity given by age; on the other, there is a strengthened foresight freedom of parents who in first instance can opt early entry enrollment for their children.

The literature on the relationship between school entry ages and factors as academic achievement is extensive and contradictory. It has largely been documented across countries that older children perform better than youngest of their cohort. Among others, Crawford et al. (2010) show that in England the negative effects of being younger at school entry are substantial even at later ages, Puhani and Weber (2008) highlight similar results for Germany and Ponzo and Scoppa (2014) find that younger children in Italy, score substantially lower than older peers at the fourth, the eighth and the tenth grade and the advantage of older students does not dissipate as they grow older. A number of different factors may contribute to this: children may suffer from the fact that they were too young when they started school; there may be a relative age effect when compared with their peers; some of them may suffer the length of schooling. Cameron and Wilson (1990), however, show that even if youngest pupils do slightly less well than the oldest, differences will in mean disappear by the third grade.

By the other side, Black et al (2008) identify a significant positive effect of early entry enrolment on IQ scores.

While optimum minimum school entry age cut-off is a broader policy concern since this may reduce public expenditure for children care, parents are equivalently and justifiably concerned about consequences on immediate and future outcomes of school entrance decisions for their children.

Ordine et al. (2018) highlight that parental choices lead to better matches between children characteristics and school entry option with consequents beneficial measure. Moreover, Dornubsch et al (1987) reforming Baumrind's definition of Parenting Styles develops and

tests the impact of authoritarian, permissive and authoritative styles on adolescent school performances, showing a negative relation of the latter two with grades of students.

Doepke and Zilibotti (2017) rationalize the choice between alternative parenting styles and demonstrate that parents can affect children' choices via two channels, either by influencing their preferences or by imposing direct restrictions. The way parents take influence on the education of their children represents a crucial point.

Dustmann (2004) analyse the association between parents' education and profession and school track choices of their children, showing a strong relation with the secondary track choice and subsequent educational achievements and positive trend for females to follow higher secondary school tracks regardless of parents styles.

The choices of the type of high school (General vs Technical) assumes a relevant aspect in studying academic performances and indicators. High school type attended, greatly depends upon the family and prior school performances. General high schools (Lyceum) are found to increase the probability of transition to university and to improve results once there; technical or vocational schools improve quality of the school to work transition, both in terms of participation and employment probabilities (Cappellari 2004).

In this view, early entry option and paternalism influences on educational choices could in part explain gender difference in academic decisions and be a source of gender stereotypes. Examples of stereotypical perceptions are that boys have more talent and girls compensate by working hard (Deaux and La France 1998). Girls are then encouraged to pursue "traditionally female studies", (Carr et al 1999).

Lavy and Sand (2015) estimate the effect of primary school teachers' gender biases on academic achievement suggesting that early stage behaviour have long run implication for occupational choices and earnings and parents' characteristics play a crucial role.

To sum up, there is little evidence on how early entry option contributes to long term academic implications, even less on the relation between parenting styles and educational tracking decisions and finally, to the best of our knowledge no contributions uses them in order to explain gender differences in educational choices.

This paper aims to provide an empirical evaluation of a long-term birth-date effect on female teenagers' educational choices given early entry option and parents support.

We made use of the "Programme for International Student Assessment" (PISA), a survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students, which includes background questionnaire to provide some contextual

information, e.g. they are asked about the level of education of their parents and how many possessions there are in their household.

Exploiting the birth date cut-off generated by the reform of 2003 we are able to construct a Regression Discontinuity Design in order to evaluate if upper secondary schools' decisions is affected by the month of birth. Furthermore, as using RDDs with a discrete running variable are misleading in empirical practice, we made use of the so called "Honest" Confidence Intervals (CIs).

We show that the choice of the upper secondary school track is affected by the early entry option given parents' support. In particular, we find that teenagers students born in January or February are more likely to be tracked in a Lyceum rather than in Technical or Vocational schools. Estimates show that the probability of choosing the Lyceum increases of 0.0283 when the specific student is born within the first two month of the year. Furthermore, the same choice turns out to be significantly higher for female students, showing an increase in the probability of choosing a Lyceum of 0.0593 when the analysis is targeted by gender. These results are potentially interesting in explaining existing gender stereotypes about university decisions and their labor market performances. High schools track decisions, in fact, convey future choices, given that the Lyceum is often associated with academic paths while other institutes are associated with direct work transition. The influence of parents, who can opt or not to early entry their children, has therefore a potential impact on future outcomes, especially for girls.

The paper is organized as follows. Paragraph 2 briefly describes the institutional framework in the Italian education system. Section 3 describes data source and the sample used. The fourth paragraph identifies the strategy, while paragraph 5 reports and discuss results. In the last sections robustness checks and conclusion are spelled out.

2. Institutional framework

School entrance rules for all children at same time in very similar tracks guarantees a non-discriminatory school-time scheme. The deregulation of the temporal control of the scholastic institution and the introduction of early entry options represents a reversal of course.

A significant change in fact, in the Italian school system, took place as a result of Legislative Decree N. 59 of 2004 which, in order to regulate the reform law n.30 of 2003 of the nursery school and the first cycle of education, provides for the possibility of enrolling in the first year of the nursery school and in the first year of primary school, children who turn 3/6 years of

age by April 30 of the reference school year, while, all pupils who turn 6 by August 31 of the reference calendar year are required to enrol in primary school.
The early entry option is, however, gradually been introduced and in the first two academic years after reform (2004/2005 and 2005/2006), schools can accept early enrol only for children born before February 28th (articles 12 - 13 D.L. 59/2004).
In the last instance the school can or less admit the pupil in advance, compatibly with the availability of places, the receptivity of the structures, the functionality of the services and the financial resources of the municipalities, according to the obligations conferred by the law and in compliance with the limits set to municipal finance.
The Italian rules seem to be more rigid than several European countries which implement school early entry policies with various solutions as like as "Kindergarten" allowing school entrance from 5 to 7 years of age (optional in Belgium and Denmark, mandatory in Greece), or in Nederland where the first enrolment at school is 5 years but can even be advanced to 4.

3. Data and sample description

3.1 Data

For the analysis we made use of the Italian 2015 OECD Programme for International Student Assessment (PISA).
PISA is an age-based survey, assessing 15-year-old students in school in grade 7 or higher which take place every three years. The first survey took place in 2000 and the last since now in 2018.For each assessment, reading, mathematics or science is chosen as the major domain and given greater emphasis than the remaining two minor domains. In 2015 the major domain was science. Students in the survey are approaching the end of compulsory schooling in most participating countries, and school enrolment at this level is close to universal in almost all OECD countries.
In addition, PISA uses Student Questionnaires to collect information from students on various aspects of their home, family and school background, and School Questionnaires to collect information from schools about various aspects of organisation and educational provision in schools. In PISA 2015, 18 countries also administered a Parent Questionnaire to the parents of the students participating in PISA and that’s the reason why we use the 2015 year of survey.

The optional Parent Questionnaire was administered on paper and targeted the parents of all students participating in PISA. It enquired about learning contexts, support, and resources at home as well as spending on education and parents.
In Italy, as in the majority of countries, the target population is defined as all students born in 1999 who were attending an educational institution.
A minimum of 150 schools were selected in each country; if a participating country had fewer than 150 schools then all schools participated. Within each participating school, a predetermined number of students were randomly selected with equal probability. The target cluster size selected per school had to be at least 20 students, so as to ensure adequate accuracy in estimating variance components within and between schools – a major analytical objective of PISA.
Approximately in Italy, PISA 2015 data was collected on a sample of over 450 participating schools. The Italian sample was stratified by macro-geographical area and type of education (General Institutes, Technical Institutes, Vocational Institutes, Vocational Training Centres, First-level Secondary Schools). The questionnaires include numerous indicators for reporting over time (trend indicators) or were designed to be used in analyses as single items (for example, gender). However, many questionnaire items were designed to be combined in some way in order to measure latent constructs that cannot be observed directly (e.g., a student's achievement motivation or economic, social and cultural background). To these items, transformations or scaling procedures were applied to construct meaningful indices[3].

In PISA 2015 categorical items from the context questionnaires were scaled using IRT (Item Response Theory) modelling. WLEs (Weighted Likelihood Estimates) for the latent dimensions were transformed to scales with a mean of 0 and a standard deviation of 1 across OECD countries (with equally weighted countries), meaning that the average OECD student would have an index value of zero and about two-thirds of the OECD student population would be between the values of -1 and 1. It is possible to interpret these scores by comparing individual scores or group mean scores to the OECD mean.
Negative values on the index for example, means students who responded less positively than the average student across OECD countries. Likewise, students with positive scores are those who responded more positively than the average student in OECD countries.

[3] Pisa 2015 Technical Report, OECD

We therefore came to have relevant information about two types of results (see table 1 for details on variables):

- basic indicators that provide a profile of students' knowledge and skills;
- indicators derived from questionnaires that show how these skills are associated with different demographic, social, economic and educational variables.

3.2 Sample description

The sample consists of 9834 students, all born in 1999 and just over half of them is female. Given the year of birth, the target population is consequently defined as all students aged from 15 years and 3 completed months to 16 years and 2 completed months at the beginning of the assessment period.

At this stage in Italy a student must be regularly enrolled at second year of high school, at grade 10, unless he/she is a repeating student or a redshirter and therefore is in lower grades, or is an early enter student and is in 2015 enrolled at grade 11.

Table 2 gives some description of the sample composition. As can be seen, around 16% of students attend a grade lower than 10 at the assessment period; of these, around 14% because have repeated at least once, while almost 2% for other reasons.

Almost 80% of these students are enrolled regularly. However, among regular students, those who were born between January and February, could have enrolled early in the 2004/2005 scholastic year but their parents opted for regularity, others could be early entry students which repeated a year. For this reason, we exclude repeaters from the estimates. Approximately 4% of students, was enrolled in advance in the first year of primary school, attending therefore grade 11 in 2015.

About 46% attend a Lyceum, almost 33% a Technical school and the remaining part is divided between vocational and private schools.

The largest share of students in the sample attends schools in regions of North-East Italy, followed by North-West with 21%, then South with 19.30% and in finally regions of Center and Islands, 8.30% and 7% respectively.

Passing to table 3, it can be noticed that the level of education of the mothers is in percentage higher than that of the fathers; both among boys and girls, in fact, the percentage of high school degrees is greater for mothers than for fathers. Also in terms of occupation type, the percentage of jobs for which a high educational qualification is required is greater for mothers. This could reasonably depend on the fact that among the jobs classified as low, many blue collars position are included, which are generally associated with men.

The sample does not provide data on the incomes of parents, however, in substitution a variable of particular interest is the one defined as ESCS: index of economic, social and cultural status, which is a composite score built by the indicators of parental education, highest parental occupation and home possessions. As can be seen, the value is on average 0.082 points lower than the international average in the case of female students, while it is 0.0261 higher in the case of students.

4. Identification strategy

This study combines detailed information from a cohort of Italian students and their parents and exploits the natural experiment that stems from the 2003 school entry age reform in order to analyse long-term birth date effect on students making education decisions given parents supports.

Let us formulate in a more precise way. In a first step, we estimate the probability of choosing a Lyceum instead of Technical or Vocational High School on several students and parents characteristics. The predicted values of this probability come from the estimation of the following equation:

$$Prob\ (\ Lyceum)_{ij} = \beta_0 + \beta_1\ \times_i + \beta_2\ Z_i + \ e_{ij} \qquad (1)$$

Where i indicated the generic student and j identifies its parents. The dependent variable Lyceum, is a dummy that equals 1 if student attends a Lyceum, while equals 0 if he/she attends a Technical or Vocational Institute. In the RHS of equation 1 X_i is a vector of students' characteristics which includes a Gender dummy variable (1 if the student is female) and immigration dummy variable (1 if the student is immigrant), three variables indicating students' performance through test scores in Mathematics, Reading and Science; Students' expected occupation status variable, a categorical variable describing the localization of the school in which the student is enrolled expressed in macro-areas and a variable which describes the level of student anxiety at school compared to OECD mean scores. Z_i is a vector of parents' characteristics including parents' education, parents' occupation, a dummy identifying authoritarian parents, a variable describing the level of parents' emotional supports and the ESCS index. β_0, β_1 and β_2 represent parameters, while e_{ij} residuals. We estimate probabilities through a probit analysis in order to use them in the second step as follows:

$$Prob\_Lyceum_i = f(MB_i) + \ u_i \qquad (2)$$

Where f(MB_i) is the student i expectation of the probability of attending a Lyceum conditional on the running variable MB_i which is its Month of Birth. MB_i assumes values from 1 to 12 starting from December (MB_i = 1 if student is born in December - MB_i = 12 if student is born in January). Students are treated if the running variable lies above the cut-off of 11, that happens when month of birth is February or January. The equation is separately estimated for male and female and parameter of interest is given by the jump of the function f(MB_i) at the cut-off as in equation (3):

$$\beta = \lim_{MB \downarrow 11} f(MB) - \lim_{MB \uparrow 11} f(MB) \qquad (3)$$

The running variable thus has G=12 support points, of which G+=2 are above the threshold and G-= 10 are below, meaning that the number of support points near the cut-off is too small and there may be not enough observations close to the threshold. In this framework, particular attention should be devoted to the assessment of standard errors and confidence intervals (CIs). As pointed out by Kolesàr and Rothe (2018), discreteness of the running variable causes problems if the number of support points close to the cut-off is small and, in this case, using a small bandwidth is not sufficient to make the bias of the estimator negligible. In the light of this, we made use of the so called “Honest” CIs in the sense of Lee (1989) which achieve asymptotically correct coverage uniformly satisfying the respective assumptions. We construct a two-sided CIs around local linear estimators using bandwidth that is optimized for a given performance criterion[4]. The honesty is realized in the sense that if the regression errors are normally distributed with known variance, the CIs are guaranteed to achieve correct coverage in finite samples. Furthermore, because the CIs explicitly take into account the possible bias of the estimators, the asymptotic approximation doesn’t rely on the bandwidth to shrink to zero at a particular rate.

In this way we estimate the direct long-term effect of the early entry option on teenagers' high school choices and the indirect effect of parental empowerment on the same choices, since the final decision on early entry their children is up to them. Paragraph 5 show results.

[4] BME approach, Kolesàr and Rothe, (2018)

5. Results

Figure 1 and 2 start the evaluation with graphical analysis. The dependent variable, namely the predicted probability of enrolling in a Lyceum is plotted on vertical axis, while Month of Birth lies on the horizontal one.

In Figure 1 all students independently from gender are plotted. Although we cannot establish if the discontinuity at the cut-off separating students born on or before February with the others born from March onwards is due to a direct effect of entry options or instead, it is caused by other factors like attitudes, the figure depicts a clear presence of differences within students. The ones with early entry options, in fact, are more likely to be tracked in a Lyceum than the other group. Furthermore, the jump appears to be higher when only females are examined in Figure 2. This could essentially have two meanings, first of all females are more likely to pursue academic tracks, secondly females are much influenced by parents' support.

Table 4 reports results obtained by implementing a local linear regression without covariates on the bins adjacent to the cut-off point allowing the slope to differ on either side of it.

A common practice to determine CIs in the case of discrete running variable is to use standard errors clustered by the same variable. However, as explained before, discreetness of the running variable causes problems when the number of support points close to the cut-off is small. Using "Honest" CIs, according to the method presented by Kolesàr and Rothe (2018), we construct two-sided CIs around local linear estimators using bandwidth that is optimized for a given performance criterion.

The sense of the honesty is that, if the regression errors are normally distributed with known variance, the CIs are guaranteed to achieve correct coverage in finite samples. The method computes honest CIs when the parameter space corresponds to a second order Holder smoothness class. We then, directly calculate honest CIs (BME method) under the assumption that the specification bias at zero is no worse at the cut-off than away from the cut-off as in section 5.2 in Kolesàr and Rothe (2018).

In table 4 we report estimates, honest CIs, average normalized BME standard errors, implied bandwidth and the effective number of observations. As we can see, results are consistent with Figure 1 and 2.

Students born on or before February are more likely to track a Lyceum. In column 1, estimates point out an increase of the predicted probability of choosing a Lyceum of 0.0283 if all students are considered. Their statistical significance is supported by BME procedure. The result, however, may be related to students' attitudes and different parents' behaviour.

The dependent variable used, however, is given by the predicted values of the probability of choosing a Lyceum estimated on students and parents' characteristics in order to mitigate potential source of bias, such as test scores and level of anxiety for students and parents' emotional support or styles for parents. Since the dependent variable is a dummy for the type of school we do not control for any school characteristics.

In the same table, in column 2, results describe estimation conducted within female students. Interestingly, we find that the month of birth and, indirectly the new arrangement of early entry options, have greater influences on girls. The increase in the predicted probability of choosing a Lyceum instead of other school, conditional on the month of birth is 0.0593. Even in this case, possible source of bias is in part controlled by the variables included in the probit analysis.

The reported estimation therefore, is a clear evidence of a long term month of birth effect, as early entry option positively affects the probability of choosing a Lyceum instead of other schools, given parents support.

Parents' responsibility, therefore, doesn't end by establishing pupils' readiness for school, but goes further, to the concern associated with the idea that current choices may have important future consequences, with greater emphasis in the female cases. Gender stereotype, which see women more prone to intellectual work may be fuelled, shaping again, the way parents take decisions.

6. Robustness Checks

In this section, we investigate whether the results are sensitive to the cut-off specifications providing some falsification exercises. As explained above, we exploit the birth date cut-off generated by the reform of 2003 to construct a regression discontinuity design in order to evaluate if upper secondary schools' decisions are affected by the month of birth. In particular, only students who were born between January and February, could have the early entry option and we used this threshold to estimate variations in the probability of enrolling in a Lyceum instead of other schools. In order to validate our estimate, we set birth date cut-off simulating fake rules generated by the reform. We replicate equation (2) which estimate the probability for a student of attending a Lyceum conditional on the running variable which is its Month of Birth. The equation is set up exactly as before for everything that does not concern the threshold. In a first stage, students are treated if the running variable lies above the cut-off of 10, that happens when month of birth is March (that does not comply with the real rules dictated by the 2003 reform). In a second step, students are

treated if the running variable lies above the cut-off of 3, that happens when month of birth is October.

The idea is that, setting fake cut-off could reduce as much as possible concerns related to students' attitudes and specification issues. If significant jump will also detect in falsification exercises our main results could reflect bias or simply the absence of any impact on the possibility of early entry options on school choices.

In table 5 we report estimates, honest CIs, average normalized BME standard errors, implied bandwidth and the effective number of observations. Results seem to strengthen our analysis. Students born on or before March are less likely to track a Lyceum. In column 1, estimates point out a reduction of the predicted probability of choosing a Lyceum of - 0.0140 if all students are considered. While, setting October as birth-date thresholds results show an insignificant variation of the same probability. In the same table, in column 2 the investigation is targeted to female students. When the fake cut-off is set at March, the coefficients results in a variation of the probability of choosing a Lyceum equal to 0.002 points instead of 0.0593 estimated in the real settings of equation (2), while coefficient equals 0.004 when the fake cut-off is October.

These results seem quite interesting and support our hypothesis according to which early entry option positively affects the probability of choosing a Lyceum instead of other schools, given parents support.

The graphical investigation confirms our hypothesis.

Figure 3 (all students) and 4 (only females) replicate the same analysis as before in which the dependent variable, namely the predicted probability of enrolling in a Lyceum is plotted on vertical axis, while Month of Birth lies on the horizontal one, exploiting however, a fictitious threshold.

The jump depicted in Figures 1 and 2 disappears when the reform cut-off is set to be March instead of February. In this case, no particular differences within students are detected at the threshold.

Implications are something interesting. Firstly, the observed variations in the probability of school track are significantly affected by the possibility of choosing whether or not to enroll pupils in advance. Secondly, parents' behaviours could influence the long-term decisions of their children, conveying any future choices.

7. Conclusions

This paper documents statistically significant long-lasting effects on schooling decisions with backing up school entry cut-off date.

Exploiting the birth date cut-off generated by the reform of 2003, we construct a Regression Discontinuity Design in order to estimate how the month of birth determines the probability of students to enroll in a Lyceum instead of a technical or Vocational school, given parents' support. The predicted values of probabilities related to upper secondary school decisions are estimated using several students and parents' control, including test scores, level of anxiety at tests, mothers and fathers' education and occupation levels, as well as parents' behaviours.

We show that the choice of the upper secondary school track is affected by the early entry option given parents' support. In particular, we find that students born on or before February are more likely to be tracked in a Lyceum rather than in Technical or Vocational schools. Interestingly, the month of birth and indirectly, the new arrangement of early entry options, have greater influences on girls.

Results suggest that parents' behaviour at early stage have long run implications for schooling decisions, and in turn for academic and occupational outcomes. These may have potentially interesting policy implications, as the increased probability of attending a Lyceum carries meaningful consequences because affect the quantity and the quality of academic track in an era in which more and more importance is given to high levels of education in order to easy access to labour market, equally between men and women.

References

1. Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence*, *11*(1), 56-95.
2. Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Too young to leave the nest? The effects of school starting age. *The Review of Economics and Statistics*, *93*(2), 455-467.
3. Cameron, M. B., & Wilson, B. J. (1990). The effects of chronological age, gender, and delay of entry on academic achievement and retention: Implications for academic redshirting. *Psychology in the Schools*, *27*(3), 260-263.
4. Cappellari, L. (2004). High school types, academic performance and early labour market outcomes.
5. Carr, M., Jessup, D. L., & Fuller, D. (1999). Gender differences in first-grade mathematics strategy use: Parent and teacher contributions. *Journal for research in mathematics education*, *30*, 20-46.
6. Crawford, C., Dearden, L., & Meghir, C. (2010). *When you are born matters: the impact of date of birth on educational outcomes in England* (No. 10, 06). IFS working papers.
7. Deaux, K., & LaFrance, M. (1998). Gender. The Handbook of Social Psychology, vol. II.
8. Doepke, M., & Zilibotti, F. (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, *85*(5), 1331-1371.
9. Dornbusch, S. M., Ritter, P. L., Leiderman, P. H., Roberts, D. F., & Fraleigh, M. J. (1987). The relation of parenting style to adolescent school performance. *1987*, *58*, 1244-1257.
10. Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, *56*(2), 209-230.
11. Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, *108*(8), 2277-2304.
12. Lavy, V., & Sand, E. (2015). *On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases* (No. w20909). National Bureau of Economic Research.
13. Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655-674.

14. Ordine, P., Rose, G., & Sposato, D. (2018). PARENTS KNOW THEM BETTER: THE EFFECT OF OPTIONAL EARLY ENTRY ON PUPILS'SCHOOLING ATTAINMENT. *Economic Inquiry*, *56*(3), 1678-1705.
15. Ponzo, M., & Scoppa, V. (2014). The long-lasting effects of school entry age: Evidence from Italian students. *Journal of Policy Modeling*, *36*(3), 578-599.
16. Puhani, P. A., & Weber, A. M. (2008). Does the early bird catch the worm? In *The economics of education and training* (pp. 105-132). Physica-Verlag HD.

Table 1 – Major Variables

GRADE	Grade compared to modal grade in country	11 – Third year of high school; 10 – Second year of high school; <10 – years below the 2th year high school (8 – 9).
GENDER	Students' gender	1 if female
REPEAT	Grade repetition	1 if student has ever repeated a grade
IMMIG	Immigration status	1 if immigrant
MATHEMATICS	Index of performance in Maths	From 1 - very low to 6 - very high
READING	Index of performance in Reading	From 1 - very low to 6 - very high
SCIENCE	Index of performance in Science	From 1 - very low to 6 - very high
M_EDUCATION	Mother's education	ISCED levels
F_EDUCATION	Father's education	ISCED levels
M_OCCUPATION	Level of Mother's occupation	ISCO - ISEI skills level (Ganzeboom and Treiman, 2003)
F_OCCUPATION	Level of Father's occupation	ISCO - ISEI skills level (Ganzeboom and Treiman, 2003)
ESCS	Index of economic, social and cultural status	Composite score built by the indicators parental education, highest parental occupation and home possessions (Cultural possessions at home; Home educational resources; Family wealth; Home possessions.
MACRO-AREA	Students' school localization	1 – North West; 2 – North East; 3 – Center; 4 – South; 5 – Islands.
EXPECTED	Students' expected occupation status	ISCO - ISEI skills level (Ganzeboom and Treiman, 2003)
ANXTEST	Test anxiety	OCSE mean comparison based on four-point Likert scale with the answering categories "strongly agree", "agree", "disagree", and "strongly disagree".
EMOSUPS	Parents emotional support	Students were asked about their perceived emotional support from their parents using questions including items on whether parents are interested in school activities, support the students' educational efforts and achievements, support students when they are facing difficulties at school and encourage them to be confident
AUTHORITARIAN	Parents identified as authoritarian	1 if the answer of students to some questions about their choices was: because my parents want it.

Notes: See Pisa 2015 Technical Report for detailed explanations

Table 2 – Descriptive statistics of all students in the sample

Month of Birth	Freq.	%	Female	Male	<10	10	11	Repeat	Lyceum	Technical	Others	North West	North East	Center	South	Islands
January	855	8,69	50,41	49,59	10,99	59,77	29,24	15,44	46,43	35,32	18,25	20,47	45,85	7,49	19,30	6,90
February	732	7,44	51,09	48,91	14,75	65,85	19,40	15,44	45,49	32,79	21,72	18,17	49,59	7,79	19,54	4,92
March	761	7,74	48,23	51,77	14,98	85,02		14,06	45,20	32,06	22,73	23,00	45,07	6,70	19,05	6,18
April	755	7,68	49,80	50,20	13,91	86,09		12,05	45,96	34,04	20,00	23,18	47,55	8,74	13,91	6,62
May	840	8,54	47,26	52,74	14,88	85,12		13,93	47,26	32,98	19,76	22,62	43,21	8,69	20,00	5,48
June	824	8,38	49,03	50,97	15,78	84,22		15,05	45,87	34,47	19,66	20,87	41,75	10,92	19,17	7,28
July	889	9,04	53,54	46,46	18,00	82,00		15,97	46,57	30,82	22,61	19,91	43,08	9,11	19,01	8,89
August	892	9,07	51,57	48,43	14,46	85,54		12,56	46,52	31,95	21,52	20,52	44,96	9,53	18,83	6,17
September	867	8,82	52,25	47,75	16,96	83,04		13,96	44,87	32,64	22,49	22,72	42,33	6,46	20,42	8,07
October	856	8,70	48,01	51,99	17,17	82,83		13,90	46,14	31,43	22,43	21,03	41,12	7,71	20,56	9,58
November	778	7,91	49,61	50,39	21,21	78,79		15,55	45,12	33,42	21,47	20,31	44,47	6,56	21,85	6,81
December	785	7,98	51,08	48,92	20,38	79,62		13,25	44,33	33,25	22,42	23,44	40,51	9,30	19,62	7,13
Total	9834	100,00	50,19	49,81	16,11	79,91	3,99	14,27	45,84	32,91	21,25	21,34	44,04	8,27	19,30	7,05

Notes: The total number of students is 9834, divided by month of birth. All classification in percentage are related to the total of students by months of birth.

Table 3 – Descriptive statistics of students' parents by gender

	Female	Male
	50,19	49,81
Mother Education		
Low	27.74	24.42
Medium	42.85	38.30
High	29.42	37.28
Father Education		
Low	30.49	28.22
Medium	41.57	39.26
High	27.94	32.52
Mother Occupation		
Very low	8.89	7.64
Low	33.00	31.56
Medium	15.46	15.07
High	42.65	45.73
Father Occupation		
Very low	9.14	8.70
Low	48.60	44.18
Medium	19.39	21.48
High	22.87	25.64
ESCS (mean)	-.0882	.0261

Notes: Father/Mother's educational qualification grouped into three groups according to ISCED levels; Father/Mother's occupation grouped into four groups according ISEI. See ESCS description in Table 1.

Table 4 – RDD with discrete running variables
Non parametric estimates of Discontinuities

Dep. Var.	**Probability of choosing a Lyceum**	
	(1) All	(2) Female
Estimates	0.0283***	0.0593***
BME CI	(-0.01176, 0.0760)	(-0.0388, 0.1220)
BME normalized SE	0.0232	0.0235
Implied Bandwidth	3	4
Effective Observations	894	534

Notes: RDD local nonparametric estimates. The dependent variable is the predicted probability of choosing a Lyceum; BME CI refer to "honest" confidence intervals as discussed in Kolesar and Rothe (2018). BME normalized standard errors are reported. Effective observations indicate the total number of observations just above and below the cut-off. In column (1) all students are considered, in column (2) only female students. No covariates are included. ***Significant at 1%; **significant at 5%; *significant at 10%.

Table 5 – RDD with fake cut-off
Non parametric estimates of Discontinuities

Dep. Var.	**Probability of choosing a Lyceum**	
First falsification	(1) All	(2) Female
Estimates (March)	-0.014***	0.002***
BME CI	(-0.063, 0.0349)	(-0.0517, 0.0562)
BME normalized SE	0.0225	0.0228
Implied Bandwidth	3	4
Effective Observations	1100	628
Second falsification	(1) All	(2) Female
Estimates (October)	0.001***	0.004***
BME CI	(-0.0484, 0.0504)	(-0.0608, 0.0696)
BME normalized SE	0.0227	0.0306
Implied Bandwidth	3	4
Effective Observations	1158	653

Notes: RDD local nonparametric estimates. The first falsification uses March as birth date fake cut-off. The second falsification uses October as birth date fake cut-off. The dependent variable is the predicted probability of choosing a Lyceum; BME CI refer to "honest" confidence intervals as discussed in Kolesar and Rothe (2018). BME normalized standard errors are reported. Effective observations indicate the total number of observations just above and below the cut-off. In column (1) all students are considered, in column (2) only female students. No covariates are included. ***Significant at 1%; **significant at 5%; *significant at 10%.

FIGURE 1 - Probability of enrolling in a Lyceum against month of birth

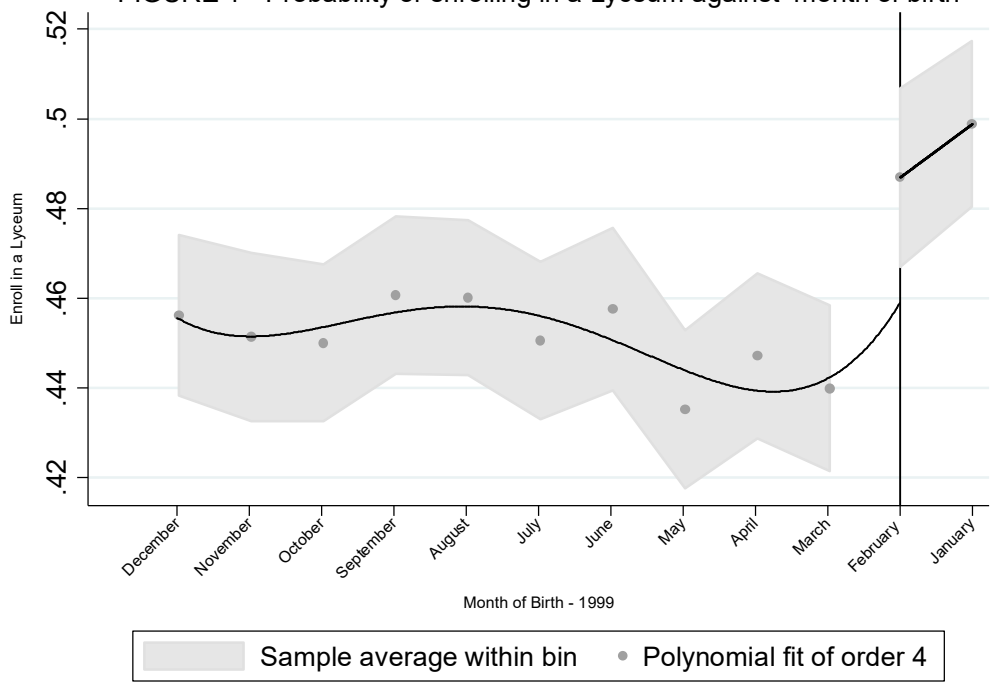

FIGURE 2 - Probability of enrolling in a Lyceum against month of birth - Females

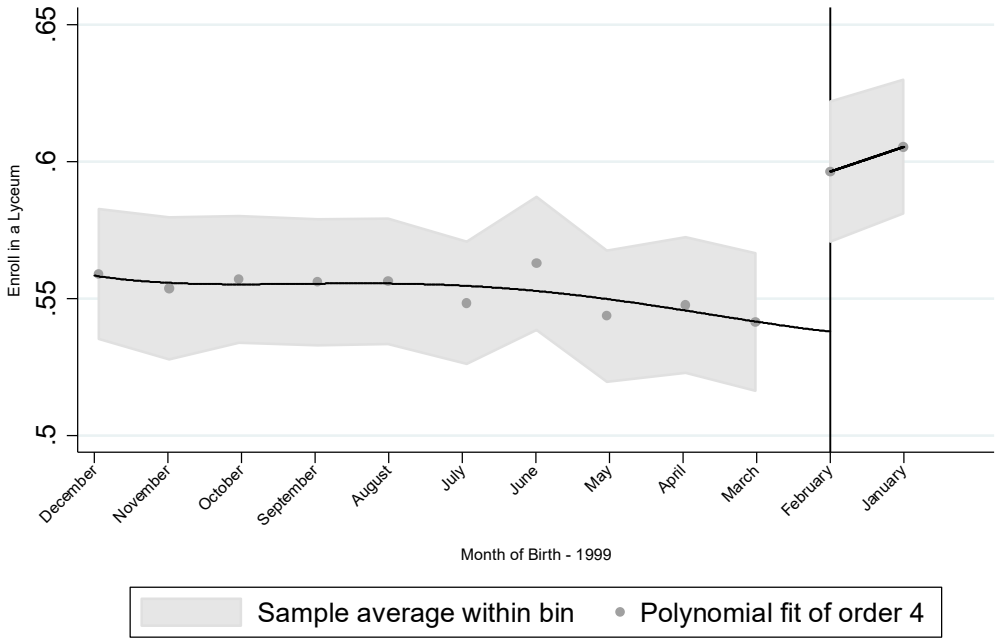

FIGURE 3 - Probability of enrolling in a Lyceum against month of birth (fake cut-off, March)

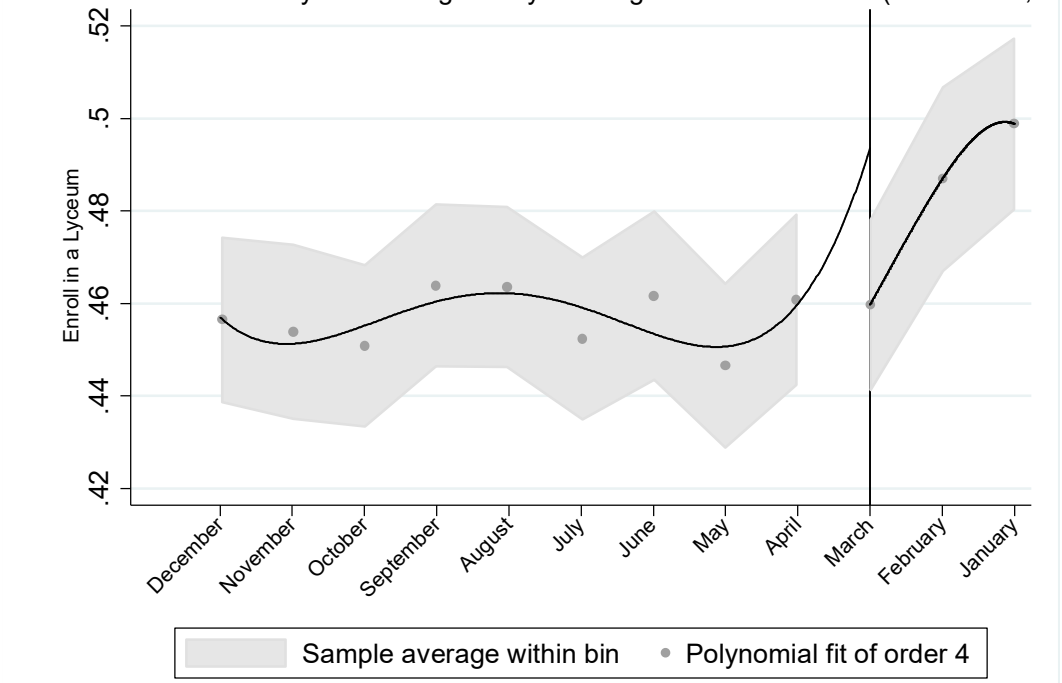

FIGURE 4 - Probability of enrolling in a Lyceum against month of birth - Female (fake cut-off, March)