Too young to enter? Students' age and their sorting into academic and vocational education

Manuel Aepli, Swiss Federal University for Vocational Education and Training, and GLO

Jürg Schweri, Swiss Federal University for Vocational Education and Training, University of Bern, and GLO *

March 28, 2023
(Working Paper Draft, consult before quoting)

Abstract

Studies show that school starting age affects school performance and labor market outcomes. As the channels for long-term effects are not yet clear, we analyze the effect of student age on their sorting into academic and vocational education in adolescence. In regression discontinuity estimations using several full cohorts of Swiss adolescents, we show that students that started school at a younger age are 16 percent less likely to attend baccalaureate school and more likely to attend vocational education. Effects hardly differ by parental background and sex, but by institutions in Swiss cantons. We apply an IV mediation approach in the RD setting to analyze how ability tracking of students in lower secondary school mediates the effect of school starting age on later sorting into academic and vocational programs. We find an indirect effect of school starting age operating via early tracking of about one third of the total age effect. The remaining direct effect is consistent with an important effect of adolescents' age at the time of sorting into academic and vocational education. We thus underline the importance of age in educational sorting during adolescence.

Keywords: educational sorting; mediation analysis; regression discontinuity; school enrollment; tracking; vocational education and training

---

*Contact: Jürg Schweri, Kirchlindachstrasse 79, CH-3052 Zollikofen, juerg.schweri@ehb.swiss.
We thank Uschi Backes-Gellner, Anne Brenøe, Martin Brown, Martin Huber, Andreas Kuhn, Guido Schwerdt, Stefan Wolter, and conference and seminar participants at the German Economic Association, Lisbon Economics and Statistics of Education, Leading House Conference on VET, German Standing Committee on the Economics of Education, Swiss Society for Economics and Statistics, and the Universities of Neuchâtel and Zurich for helpful discussions and comments.

1 Introduction

It is well established that younger students achieve lower test scores in compulsory schooling relative to their older classmates. However, the evidence is mixed on whether school starting age differences continue to affect students' lives after compulsory schooling and which mechanisms could cause such mid- to long-term effects.

In this paper, we focus on the effect of age on an incisive education decision, namely the choice between academic and vocational education at upper secondary education. We study the case of Switzerland, where adolescents and their parents start choosing between academic education and vocational programs at the end of compulsory schooling at around age 14 (contrary to Germany, where the academic track starts at age 10 already). This high-stakes choice makes Switzerland an interesting case to study age effects in adolescence. Also, a majority of Swiss students attend apprenticeship programs, whereas a quarter of students attend highly selective baccalaureate schools at upper secondary level. Therefore, the choice is between two important tracks of good reputation, not between a main track and a clearly inferior and minor one, as is the case in many countries when it comes to vocational education.

There is no consensus in the literature about the labor market effects of academic versus vocational education attendance yet. In Switzerland, graduates of vocational programs show a slight advantage over their academic peers in terms of labor market participation and unemployment, whereas academic graduates earn higher mean wages if one does not account for different years of education (see Aepli, Kuhn, and Schweri, 2021). Also in terms of occupations and job tasks, the careers of graduates from academic and from vocational education differ strongly. Therefore, an effect of age on attending academic and vocational education will likely transmit to individuals' short- and even long-term labor market careers.

There are two main reasons why we expect that age matters for the type of education, academic or vocational, individuals will attend. First, several gate-keepers for education programs use students' current physical and/or mental maturity as criterion in their decisions. As for academic education, entry to a baccalaureate school (broadly similar to grammar schools in the UK) requires either a recommendation from compulsory school teachers, who must assess students' suitability and maturity for baccalaureate school; or passing entrance exams, where age may affect performance. Similarly, when recruiting apprentices, companies assess their suitability and maturity in order to select good apprentices and to prevent premature contract terminations. Therefore, age at the moment of choice is likely to matter for both types of education programs because older students are more likely to convince gate-keepers of their maturity.[1]

Secondly, school starting age differences may affect individuals' learning before they enter secondary education, e.g. their achievement in early grades. In addition, early tracking in Swiss compulsory schooling (starting in grade 6 at age 12) based on performance may also depend on school starting age and thus influence the later sorting into academic and vocational programs. It is a priori unclear whether the later sorting would offset age effects of early tracking, at least partially, or whether it would rather perpetuate and increase early differences.

We use register data on the full population of students including exact birth dates, combined with information on different school starting date cut-off rules in Swiss cantons (i.e., the 26 states of the Swiss confederation) to estimate age effects in a fuzzy regression discontinuity (RD) design. We find that being born right before a cut-off date, thus being up to one year younger than peers born after the cut-off, decreases baccalaureate school attendance by 3.9 percentage points and increases the likelihood of making an apprenticeship by almost the same amount. We probe the robustness of these results and discuss heterogenous effects.

In addition, we use an instrumental variable (IV) mediation analysis to assess the role of early tracking as a mechanism which might predetermine later sorting into vocational and academic education. This analysis is important to clarify whether policies aimed at reducing age effects should be mainly directed at early tracking, or whether they should also focus on the selection process leading to academic or vocational education. We find a substantial indirect effect of school starting age via early tracking, which accounts for about a third of the total age effect. The remaining direct effect is consistent with a substantial effect of age at the moment of choosing between academic and vocational education.

Our paper contributes to the broader literature on age effects in education. We show that the early effects of school starting age may be perpetuated and increased in a system with sorting

[1]Not surprisingly, the role of adolescents' age in education choice has been addressed in Swiss media reports. They discussed our title question whether students are too young to choose between tracks at ages 14 and 15, in particular because of recent cantonal reforms in school entry dates leading to even younger students and apprentices at upper secondary level. See e.g. "Aargauer Zeitung" (2017-11-11), "20 Minuten" (2019-01-31), "SWR info" (2019-04-15).

into academic and vocational education at upper secondary level. In addition, we disentangle the effect of age on earlier tracking in grade 6 from the age effect on the subsequent sorting into academic and vocational education. We use a mediation approach that combines the regression discontinuity design with an instrumental variable for the mediator (early tracking) and find evidence compatible with an effect of age at the time of choice for upper secondary education, i.e., in adolescence. This teen-age effect comes on top of the earlier effects caused by school starting age differences. We conclude that when analyzing effects of age on educational attainment, we should not restrict our focus on years of education, but also consider the type of education attained.

2 Literature on age effects in education and on the outcomes of vocational education

The effects of school starting age (SSA) on students' cognitive and non-cognitive development in school are well documented.[2] Older students typically obtain better test score results than their younger peers in the same class. In much of the literature, this age effect is a combination of the effects of relative age compared to class mates and absolute age, because both depend on the combination of birthdate and school starting cutoff dates in education laws.

It is less clear whether and how these age effects persist into adult life.[3] Black, Devereux, and Salvanes (2011) find modest effects of SSA on years of education for adults in Norway, while Fredriksson and Öckert (2014) find more substantive effects in Sweden. The latter authors also point to the importance of tracking in school as a mechanism for perpetuating early age effects. They show that a school reform that postponed tracking reduced the effect of SSA on educational attainment.

Early tracking plays a prominent role in Germany[4], where students are tracked in school at age 10, including the selection into the academic track (*Gymnasium*) that lasts until the end

[2]See, for example: Bedard and Dhuey (2006), Puhani and Weber (2007), Elder and Lubotsky (2009), Lubotsky and Kaestner (2016), Sprietsma (2010), Navarro, García-Rubio, and Olivares (2015), Attar and Cohen-Zada (2018), Dhuey, Figlio, Karbownik, and Roth (2019), Balestra, Eugster, and Liebert (2020).

[3]Some studies find SSA effects on health outcomes, e.g. Bahrs and Schumann (2020) and Johansen (2021). School enrollment cutoffs may also affect the family, e.g. mothers' labor market attachment (Gangl and Huber, 2021).

[4]Oosterbeek et al. (2021) show an effect of school starting age on tracking in the Netherlands.

of upper secondary schooling (Matthewes, 2021). Mühlenweg and Puhani (2010) show that students born in July, i.e. after the cutoff, in the German state of Hessen are considerably more likely to attend the highest track, i.e. *Gymnasium*. However, the later possibility to enter vocational education at age 16 attenuates the age effect on achievement. Dustmann, Puhani, and Schönberg (2017) use the SSA effects on early tracking in Germany as an instrumental variable to analyze the long-term effects of early tracking. They find that early tracking has no impact on wages, unemployment and occupational choice for marginal students, who are close to the threshold between two tracks. They provide evidence that this is because the education system offers mismatched individuals several possibilities to upgrade or downgrade. In line with this argument, they find no effect of SSA on the probability of adults to hold an apprenticeship diploma. Görlitz, Penny, and Tamm (2022) confirm this result using different data, showing that SSA has no statistically significant impact on completing an apprenticeship or college.

These results suggest that SSA has a strong effect on early tracking, but that the completion of academic versus vocational education is hardly affected in Germany. As we discuss in the next section, the sorting in academic versus vocational education is split into several decisions at different ages in Germany. In Switzerland, this sorting happens at the end of compulsory school when students are in adolescence.

To date, there is only one paper on SSA effects in Switzerland. Balestra, Eugster, and Liebert (2020) study the Swiss canton of St. Gall and find that children with higher SSA develop less special educational needs during compulsory. They do not find an effect of SSA on choosing an academic or vocational track, but a higher probability for a high-quality apprenticeship. We will show that, in line with Balestra, Eugster, and Liebert's (2020) findings, we do not find SSA effect on the choice of academic vs. vocational education in the canton of St. Gall, but substantial effects in many other cantons and in Switzerland as a whole.

Attending vocational or academic education potentially leads to very different individual life courses. The increasing literature on labor market outcomes of vocational education is inconclusive, however. One prominent argument holds that vocational education helps to integrate young adults swiftly into the labor market, but that vocational skills become obsolete at a faster rate than academic skills, leading to better outcomes for academic graduates in the long-run. Several studies found such a trade-off between short-run and long-run results for vocational graduates.[5] On the contrary, studies focussing on particular cases providing natural experiments for single countries often find that vocational graduates have similar outcomes like academic graduates, i.e., controlling for different years of education and selectivity.[6] These different assessments likely stem from the differences in countries included and their institutional settings (e.g., the quality of vocational education), and differences in estimation designs.

In Switzerland, descriptive numbers on labor force participation and unemployment favor vocational graduates throughout careers (Aepli, Kuhn, and Schweri, 2021; Korber and Oesch, 2019). Wages are higher for academic graduates, on average, but rates of return are rather below average when compared to other countries (Saltiel, 2021; Wolter, Cattaneo, Denzler, Diem, Hof, Meier, and Oggenfuss, 2018). In short, it is not clear whether students that are on the margin between academic and vocational education profit more from being sorted into one or the other route, but as the next chapter shows, these routes lead to different learning environments and prepare for different careers.

3 Education system and student age in Switzerland

3.1 School system

Compulsory schooling is quite diverse between Switzerland's 26 cantons (states). In the time period relevant for our study, school cut-off dates varied from from 31 December in the Canton of Ticino – meaning that each child turning 6 in the respective year goes to 1st grade – to 28th February in the Canton of Zug – meaning that only children turning 6 before this date go to school in the respective year.

Students have to attend primary school in the community of residence after kindergarten. Compulsory schooling lasts full nine years, independently of students' age. Again depending on canton, students are tracked after 4th, 5th or 6th grade[7] into typically two tracks, one with standard requirements and one with extended requirements. This tracking is based on school

[5] See Chuan and Ibsen (2022); Golsteyn and Stenberg (2017); Hanushek, Schwerdt, Woessmann, and Zhang (2017).

[6] See Bertrand, Mogstad, and Mountjoy (2021); Birkelund and van de Werfhorst (2022); Hartog, Raposo, and Reis (2022); Malamud and Pop-Eleches (2010); Silliman and Virtanen (2022).

[7] Throughout this paper, we use the grade counts valid before 2015, with 1st grade starting after kindergarten. As of 2015, kindergarten attendance is compulsory and counts as grades 1 and 2, such that the former 1st grade has become 3rd grade.

grades in languages and mathematics, and teacher assessments of students' performance and potential. About x percent attended the lower track in the relevant time period, y percent the higher track. In most cases, students in the higher track form own classes, either in the same school building or in a separate school (*Sekundarschule*).

After 9th grade, i.e. after completing lower secondary school, compulsory schooling ends and students are free to enter the labor market. However, more than 90 percent enter an upper secondary education program immediately or within few years. There are many education programs available at upper secondary level, which can be divided into two main categories: vocational and academic education (also see Figure A1 in the appendix).

The main pillar of vocational education are firm-based apprenticeships.[8] More than 60 percent of a cohort complete an apprenticeship in Switzerland by age 25 (FSO 2022). These last between two and four years and are regulated by a federal training ordinance each, ensuring nation-wide recognition in the education system and on labor markets.[9] Companies provide training places voluntarily and finance the in-company part of the training, whilst also profiting from apprentices' work (Schweri, Aepli, and Kuhn, 2021). Cantons run vocational schools, which apprentices attend during one to two days a week. Apprenticeships aim at preparing learners to enter the labor market as skilled workers directly. However, there are various possibilities to acquire tertiary level credentials afterwards, in particular in higher vocational education and in universities of applied sciences.

The main pillar of academic education are baccalaureate schools (*Gymnasium/lycée/liceo*), attended by roughly one quarter of a birth cohort. Successful completion is awarded with the baccalaureate (*Maturität/maturité/maturità*), which grants direct access to all academic universities and almost all fields of studies. The second type of academic upper secondary school is specialized school, which prepares mainly for careers in social and health care or in teaching. These schools account for few percent of all students of a cohort.

Most students enter baccalaureate school after 8th or 9th grade. This is in stark contrast

[8]Vocational education is also offered as full-time school programs, leading to the same federal diploma as firm-based apprenticeships. However, diploma from full-time schools make up only about 10 percent of all vocational diploma awarded, in total.

[9]Vocational ordinances specify the name of the training occupation, the skills to be acquired, the final exam and the federal diploma granted after completion (Schweri, Aepli, and Kuhn, 2021). Ordinances are enacted by the confederation and updated every five years.

to the German model, where Gymnasium usually starts after 5th grade.[10] The principal choice between academic and vocational programs thus happens later in Switzerland than in Germany, which is why we focus on the transition from lower to upper secondary education during adolescence. Swiss federal framework curricula require that students receive "career orientation" lessons in grades 7 and 8, i.e. at a typical age range of 12 to 15 years, to help them decide which path to choose.

3.2 Age and the sorting into academic and vocational programs

Adolescents undergo fast mental and physical development at the age of choosing between academic and vocational programs at upper secondary level.[11] Age differences of up to one year, as introduced by school entry cutoff dates, matter a lot in terms of adolescents' maturity.

Admission mechanisms for the different programs place great emphasis on the maturity, and thus also age, of adolescents. Admission to baccalaureate schools depends on cantonal rules, either requiring a recommendation of teachers in grade 8, and/or passing an entrance exam organized by the canton.These requirements are supposed to select students for baccalaureate school that have the potential to study at a university later.[12] In the case of teacher recommendations, teachers should not only consider performance in certain subjects, but also students' engagement, planning and reflection, use of learning strategies, and autonomy and concentration in solving tasks.[13] Older students likely convey more of this desired behavior than younger students, which may prompt teachers to recommend older students more often for baccalaureate school, even conditional on their actual performance. In the case of entrance exams, older students may simply do better at the exam.

Admission to apprenticeships depends largely on firms, which offer training places. Recruiters look for apprentices that fit into the team, become productive quickly and can be

[10]Some cantons also know "long" baccalaureate schools (*Langgymnasium*), which start with grade 7 instead of 9. We account for these in the subgroup analyses in section 6.3.

[11]Largo, Gasser, Prader, Stuetzle, and Huber (1978) report mean ages for the adolescent growth spurt of 12.2 years for Swiss girls and 13.9 years for boys, with age ranges of 5.7 (girls) and 3.8 years (boys). Peak height velocity reaches 7.1 cm per year for girls and 9 cm for boys.

[12]The canton of Bern states: "Teachers do not primarily assess students' past performance, but their potential with regard to the requirements of baccalaureate school and the *attainment of maturity* for university studies. [own translation and italics]" (https://www.bkd.be.ch/de/start/themen/bildung-im-kanton-bern/mittelschulen/gymnasium/aufnahmeverfahren-gymnasium/aufnahmeverfahren-in-das-erste-jahr-des-gymnasialen-bildungsgangs-gym1-.html)

[13]See e.g. the Mittelschuldirektionsverordnung (MiSDV, 2017) in the canton of Bern.

expected to do well in vocational school. This helps to minimize the risk of premature apprenticeship terminations (Rohrbach-Schmidt and Uhly, 2015). Apprenticeship applicants of young age run the risk of appearing not sufficiently mature in the eyes of recruiters in training firms. This was confirmed by Imdorf (2012), who conducted interviews with Swiss SMEs and found that these prefer candidates that are not too young. In addition, younger students may themselves feel more uncertain which training firm and training occupation to choose, which increases their likelihood to postpone the search for an apprenticeship by entering an intermediate program offered to youngsters with difficulties in finding an apprenticeship position.

Because maturity matters for the admission to both academic and vocational education, it is unclear a priori how age affects sorting into programs. While baccalaureate schools and firms offering apprenticeship positions are interested in more able and more mature students, admission requirements for baccalaureate schools are particularly high. Therefore, we hypothesize that older students are more likely to attend baccalaureate schools, ceteris paribus, which would mean that relatively more younger students enter apprenticeships. This hypothesis encompasses two kinds of age effects: First, as argued above, age at the time of selection into academic or vocational programs is a visible, outward signal for maturity and thus matters for program admission. Secondly, school starting age differences lead to early performance differences, reflected in the early tracking of students in lower secondary school. Therefore, older students are also more likely to attend a higher track, which provides them with higher probability to attend baccalaureate school later.

4 Empirical strategy

4.1 Regression discontinuity design

Our starting point is to compare education choices at upper secondary level between students who were born just before and after the cutoff date for school enrollment. The regression discontinuity (RD) design allows us to quantify this effect at the cutoff while controlling for covariates (Calonico, Cattaneo, Farrell, and Titiunik, 2019):

$$P_{ij} = \alpha_{0j} + \alpha_{1j} AC_i + \alpha_{2j} d_i + \alpha_{3j} AC_i \times d_i + \alpha_{4j} X_i + \psi_{cj} + \epsilon_{ij} \;\; \forall \;\; i \in \{i : |d_i| \leq bw\} \quad (1)$$

Program dummies P_{ij} describe student i's education program choice j with $j = (1, 2, 3, 4)$ representing non-entrance, apprenticeship, baccalaureate school, and specialized school. We regress these dummies separately on an "after cutoff" dummy AC_i being one (zero) if student i was born after (before) the cutoff in student i's canton. The RD setup also requires controlling for the number of days (d_i) lying between student i's birthday and the cutoff, our running variable in the RD design. We include the interaction between the running variable and the AC dummy, $AC_i \times d_i$, thus allowing for different linear trends before and after the cutoff.

Estimations are implemented as local linear regressions (Calonico, Cattaneo, Farrell, and Titiunik, 2019).If being born before or after the cutoff is as good as random, the AC dummy coefficient is a causal estimate of the "reduced form" effect, that is the effect of being born before or after the cutoff on education choices at upper secondary level (see subsection 4.2).

To increase estimation precision, we also include a vector of student characteristics X: personal characteristics encompass gender and Swiss citizen dummies, geographic characteristics control for the type of municipality (urban, intermediate, rural) and the language region (German, French, Italian) students lived in when leaving compulsory school. We further control for canton fixed effects (ψ_c) to account for institutional differences across cantons.

4.2 Identification and RD parameter choices

The reduced form effect measured by α_1 is causal if assignment to the treatment (being born after cutoff) is as good as random. This assumption would be violated if parents manipulated birth dates with respect to cantonal school enrollment cutoffs. While planing a birth precisely to a date just before or after the cutoff would be hard to do because of the variation in conception periods and gestation length, manipulation is feasible when the birth process starts in the final hours before the cutoff day, in which case birth could be accelerated or delayed around midnight. Since Swiss parents can easily delay or bring forward school entry of their children by one year, such behavior seems very unlikely. Nonetheless, we show that our results do not change in a donut specification excluding students born within one day from the cutoff.

Random treatment assignment is also violated if seasonal birth patterns correlate with parent characteristics, such that students from different types of parents are over-represented on one side of the cutoff. Our relatively narrow bandwidth of 30 days around the cutoff mitigates issues due to seasonal birth patterns, which occur between spring and late autumn/winter in Switzerland.[14] In addition, we base our analysis on different cantonal cutoff dates (see table 1) and thus on observations of students born in different seasons. As a more general test, we will show that observable covariates do not jump at the cutoff (section 6.5), which is consistent with random assignment at the cutoff.

Because students born after the cutoff are older when starting school, we expect a positive coefficient for α_1 in the regression for baccalaureate school according to the hypothesis formulated in the previous section. Interpreting the reduced form effect as an age effect assumes that cutoff rules are a valid instrument for school starting age.[15]

Figure 1 confirms that there is a strong jump in age at the cutoff directly at school start, which propagates to the end of compulsory schooling.[16] However, the upper part (a) of the figure also shows that compliance is far from perfect and that there are early and late entries, which curb the age means close to the cutoff.

Figure 1 around here g.rd.ssa.labb.pdf and g.rd.ageatleaving.pdf

This considerable amount of non-compliance with the cutoff rule is not a problem for the reduced form setting, at least at first view. It is implausible that someone would want to enter school early when born after the cutoff, yet would want to enter late when born before the cutoff. Therefore, we can safely exclude the existence of such defiers. However, if we want to interpret the cutoff effect as an age effect, we treat the cutoff discontinuity as an instrument for school starting age. This instrument likely violates the monotonicity assumption (Angrist, Imbens, and Rubin, 1996; Angrist and Krueger, 1992) necessary for identifying the local average treatment effect of age. As e.g. Barua and Lang (2016) argue, students who are born after the cutoff and enter school early are now defiers: they go to school together with the "treated",

[14]See for example Eriksson and Fellman (2000) or the official numbers published by the Federal Statistical Office: https://www.bfs.admin.ch/asset/de/23328842 [retrieved on 16 November 2022]. Similarly, we observe slightly lower birth rates in the last three months of the year among students in our sample.

[15]Note that we refrain from using cutoff rules to implement a fuzzy RD two-stage least squares (2SLS) regression, which would estimate the local average treatment effect of a one year difference in school starting age. The main reason is that we do not observe students' school start in our data, which happens about nine years before the end of compulsory school. We could use younger cohorts in the LAES data for whom we observe school start to estimate two sample two-stage least squares models. However, these younger cohorts have already been affected by recent school reforms coordinated among many cantons ("Harmos"), which have changed cutoff rules and introduced compulsory kindergarten. This may also have affected compliance with the cutoff rules.

[16]The samples for the two different points in time differ, as explained in section 5.1.

who were born in the same year, but before the cutoff. This means that these defiers enter school at a younger instead of an older age than the treated. We need the cutoff instrument to make students born after the cutoff older, yet some also become younger when they enter school. The same violation of monotonicity is present in reverse when students born before the cutoff delay their school entry ("redshirting") and enter school at an older age than the "non-treated", who were born after the cutoff and go to school next year.

This is a problem for the classical estimation approach, which used all students born at any date throughout the year and uses predicted school starting age as instrument for actual starting age. Fiorini and Stevens (2021) show that the RD design offers a solution to the monotonicity problem in the particular setting of birthday cutoffs affecting age. When we restrict the analysis to students born shortly before or after the cutoff, the defiers differ only by few days in age compared to their counterfactual school starting age. In contrast, the compliers differ by almost one year in age. Thus, an RD design with small bandwiths ensures that the age variance is almost completely due to compliers and the monotonicity problem becomes negligible.[17] We will therefore interpret the reduced form RD effect as caused by age.

As is the case in most of the literature, it is not clear whether coefficient α_1 is driven by an absolute age effect, a relative age effect (being older than one's peers in class), or a combination of the two. We will differentiate between absolute and relative age effects by exploiting the different timing of cutoffs among cantons, even though this approach is limited by institutional differences between cantons.

In RD designs, an important choice concerns the optimal bandwidth around the discontinuity (see for example Cattaneo, Idrobo, and Titiunik, 2019; Imbens and Lemieux, 2008). While many authors favor a data driven process to choose the optimal bandwidth (Calonico, Cattaneo, Farrell, and Titiunik, 2017; Imbens and Kalyanaraman, 2012), our setting profits from the use of small bandwidths to dispel the monotonicity concerns discussed above. We choose a bandwidth (bw) of 30 days around the cantonal cutoff dates for our main specifications. However, our results are robust to smaller bandwidths and to larger ones, also those proposed by the data driven approach of Calonico, Cattaneo, Farrell, and Titiunik (2017). In addition, we

[17] The bias due to monotonicity in a fuzzy RD design vanishes if the effect is identified exactly at the cutoff, that is when the trends at the left and right of the cutoff are correctly specified in the estimation model. However, the narrower the bandwith, the less unbiased estimation depends on the specification of the trend.

try quadratic and fourth order polynomials of the running variable d_i in our estimations instead of the linear specification in equation 1.

Being born before or after the cantonal cutoff date affects school starting age (SSA) and thus also the age at which students leave compulsory school. Yet, we cannot instrument school leaving age with cutoff discontinuities because school starting age affects students already in compulsory school. Therefore, SSA may affect education program choice at upper secondary level via different mechanisms. As explained in section 3, a likely mechanism is early tracking at lower secondary school. This performance tracking happens between school start and upper secondary education choice. Hence, we consider early tracking as a mediator of the age effect on later education choice.

The goal of our mediation analysis is to decompose the reduced form effect from equation 1 into an indirect effect that runs via early tracking and a remaining direct effect of age. The indirect effect can be estimated from the following two equations (ignoring X and cantonal fixed effects for simplicity):

$$P_{ij} = \beta_0 + \beta_1 AC_i + \beta_2 d_i + \beta_3 AC_i \times d_i + \beta_4 T_i + \eta_{ij} \ \forall \ i \in \{i : |d_i| \leq bw\} \tag{2}$$

$$T_i = \gamma_0 + \gamma_1 AC_i + \gamma_2 d_i + \gamma_3 AC_i \times d_i + \gamma_4 Z_i + \theta_i \ \forall \ i \in \{i : |d_i| \leq bw\} \tag{3}$$

β_4 and γ_1 together identify the indirect effect of SSA on P_{ij} that runs through the mediator T_i (Baron and Kenny, 1986). Equation 1 yields the total effect of SSA on P_{ij}, the direct effect is given by the difference between total and indirect effect.

A key assumption for causal interpretation of these effects is sequential ignorability (Imai, Keele, and Yamamoto, 2010). It requires that η_{ij} and θ_i are uncorrelated. This assumption is hard to justify in our case because unobserved variables that are correlated with tracking would likely also be correlated with later education choice. Our RD design does not address this problem.

Instrumenting the mediator provides a way to causal estimation (Huber, 2021). When we assume that Z_i in equation 3 is an instrument, i.e. uncorrelated with θ_i, we can estimate the two equations as a two-stage least squares model.[18] Hence, equation 2 is at the same time a

[18]Non-parametric estimation of this IV mediation model would be hard to do in our setting. Frölich and Huber (2017) discuss non-parametric estimation of natural effects with a discrete mediator and a continuous

reduced form RD equation, as we use AC_i instead of the endogenous variable SSA in a sample of adolescents born around the cutoff, and a 2SLS second stage equation, because we instrument the mediator T_i. The instrument is described in the data section, together with the sample available for the mediation analysis.

5 Data

5.1 Longitudinal analyses in the educational sector

We use educational register data named "Longitudinal analyses in the educational sector (LAES)" provided by the Swiss Federal Statistical Office. The data contains the educational history of every student in formal education in Switzerland since the start of the schooling year 2011/2012. We can thus observe every student's transition from lower to upper secondary level between 2012 and 2018, including the education program they attend at either level, and where the program takes place (municipality). Moreover, LAES contains detailed demographic information on every student, such as their exact birth date, gender, migration status, and – through linking LAES with the Swiss Population and Households Statistics – students' cantons of residence and birth. We transform the LAES spell data into data with students as units of observation. The main variable of interest is students' education program at the upper secondary level, i.e. either non-entrance, starting an apprenticeship, entering a baccalaureate school, or a specialized school.

Main sample: school leaving cohorts

Our analyses focus on the cohorts of students in the data who leave compulsory school. As some students are faster in progressing from start to end of compulsory school than others, these school leaving cohorts are not balanced in birth years. We cannot observe the fastest students of early birth cohorts nor the slowest students of later birth cohorts in the available time window of six years. Therefore, we assume that program choice behavior and the effect of school starting age on progression in compulsory school have not changed between birth

instrument, but the approach has not been implemented to date. Note also that the propensity score weighting approach used for the estimators presented in Frölich and Huber (2017) will not work in an RD setting where the running variable is included in the estimation, because the propensity score becomes deterministic in the running variable, which violates the common support assumption for the mediator.

cohorts. We also show that results based on birth cohorts are similar despite lower numbers of observations available.

We define "first" entry into an education program at upper secondary level as an entry within one year after completing lower secondary school. Students that have not entered a program within one year are considered as non-entries.[19] This requires excluding students who completed compulsory schooling in 2018 (the last year recorded in our data set) from the analysis, because we do not observe a full year after completion for them. Moreover, we exclude very few students entering a non-official educational path at upper secondary level, including some foreign programs.

Table 1 describes the sample. Because we run most estimations with the 74,007 students born within 30 days from the cutoff date (rd sample), table 1 describes both the full and the rd sample with a bandwith of 30 days. Table 1 confirms that starting an apprenticeship after compulsory schooling is by far the most important upper secondary education in Switzerland, chosen by over 60% of all students. Almost one quarter enters a baccalaureate school and roughly 5% a specialized school, while 9% have not entered any track within one year after completing compulsory schooling. A little less than half of our sample are female students and about 86% are Swiss citizens. Students are almost 16 years old when they start their upper secondary education.

Table 1 around here table.desc

The regression discontinuity design (see section 4) relies on a running variable, in our case the number of days between students' birthdate and the cutoff date for school entrance. To construct this variable, we merge the cantonal cutoff dates (see appendix tbc) to every student based on her or his first recorded canton of residence. Unfortunately, the data does neither identify canton of residence at birth nor canton of residence at school entry for all observations[20]. Therefore, we use students' first recorded canton of residence in school to determine the relevant cutoff date. Most students appear in the LAES data before their transition to the upper secondary level and thus their canton of residence is observed some years prior to finishing

[19]Table A.1 in the appendix shows that most program choices happen within the time span of one year.

[20]The LAES contains canton of birth. However, this does not always correspond to the canton of residence at birth, because mothers can give birth in a hospital in a neighboring canton, which happens often in small cantons.

compulsory schooling. In addition, the Swiss Population and Households Statistics records the canton of residence five years ago for roughly two thirds of all students. This allows us to identify their canton of residence at a younger age in many cases. On average, students were a little older than nine years when we first observe their canton of residence. As a robustness check, we will also show estimation results based on recorded canton of birth.

School entrance sample

As the LAES data started in 2011, the time-span covered per person is insufficient to observe both school start *and* transition into upper secondary level. This is one reason why we rely on reduced form estimations, where we look at the effect of being born before or after the cutoff (instead of school starting age) on transition probabilities (see section 4.1).

However, we observe younger individuals in LAES who entered school from 2011 to 2018. This second sample is helpful to study compliance with cutoff rules. We observe 485,125 students entering school between 2011 and 2017. Again, we identify students' first recorded canton of residence to attach the relevant school starting cutoff information. Table A.2 describes this sample among the 85,231 students born within 30 days of the cutoff.

Mediation sample

In most Swiss cantons, the transition from lower to upper secondary education is preceded by tracking in lower secondary schooling after grade six. This first tracking likely affects students' upper secondary education.The time period available in our LAES data allows us to observe lower secondary track for a subset of the main sample described above. With this subsample, we perform a mediation analysis to assess how much of the overall age effect on the education program decision at upper secondary level works through the early tracking happening at lower secondary level. The sample consists of 118,372 students in grade six in the years 2011 or 2012, who finish compulsory schooling between 2013 and 2017. After excluding students from the cantons without early tracking (Obwalden, Ticino and Vallais) and reducing to those born within 30 days from the cutoff, our mediation sample consists of 17,894 students. Roughly 70% of these students attended the advanced track at the lower secondary level, 27% the basic track and very few students attended classes with no level distinction. In the mediation analysis (see section 4.2, the mediated variable is a dummy indicating students who attend the advanced track at lower secondary school.

We instrument the endogenous variable of our mediation analysis, the dummy being one for students attending the advanced track in lower secondary school, with the municipality share of students attending the advanced track. Specifically, the instrument for the mediator of student i is calculated as follows:

$$IV_{ij}^{med} = \frac{students_j^{lowsec=advanced}}{students_j^{lowsectotal}} \tag{4}$$

The instrument for student i attending primary school (grade six) in municipality j equals the share of students entering the advanced track of all students attending primary school (grade six) in municipality j (see table A.9). In order to have sufficient observations also for small municipalities, we rely on all students promoting from grade six to seven between 2011 and 2018. In case we still observe less than 50 promoting students in a municipality, we calculate the instrument on the level of the district (*Bezirk*).[21] We assume that the regional variation of this instrument is mainly driven by exogenous factors affecting the availability of the advanced track including politically defined quota and availability of schooling facilities. To account for possible endogenous factors affecting the share of students entering the advanced track, we control for the municipality share of parents with secondary and tertiary education, respectively. This captures differences in preferences for higher education and different social-economic backgrounds that can affect students' entrance probability into the advanced track.

6 Results

We first discuss how being born on either side of the cutoff affects students transition probabilities into upper secondary level programs. Section 6.2 and 6.3 and look at the results in subsamples by personal characteristics (e.g., sex) and by cantons, which feature different institutional setups. Sections 6.4 analyzes how enrollment in different apprenticeship programs and occupations is affected by age. We then check whether our results are robust to different choices

[21] Note that the instrument is not time-variant; this would be possible but leads to a decrease in the observations per municipality and thus increasing measurement error in the instrument.

in sample definition, RD bandwidth and the specification of trends in section 6.5. Finally, we study the role of early tracking in lower secondary school as a mediator of the age effect on upper secondary program choice.

6.1 Main results

First evidence comes from looking at upper secondary program choices of adolescents born before and after the cutoff. Figure 2 shows their program choices against the running variable, i.e. their birthdate relative to the school enrollment cutoff date in their canton. Each dot in the figure represents the mean share of the respective program among individuals born on the same day. Program choice was observed up to one year after leaving compulsory school.

Figure 2 around here rdplot.tracks

The share of students entering a baccalaureate school jumps at the cutoff (panel c in figure 2). This means that being born after the cutoff increases the likelihood of entering baccalaureate school. In turn, the share of students starting an apprenticeship drops at the cutoff (panel b). For the remaining less frequent options, non-entrances and specialized school, there are no clearly visible jumps in the figure.

Table 2 around here rdest.track.rdrobust.bw30

This interpretation of the graphical evidence is confirmed by table 2, which shows estimates of α_{1j} from equation 1. Being born after the school enrollment cutoff, thus being older at school start on average, lowers the chances to start an apprenticeship by 3.5 percentage points, while it increases the probability to enter baccalaureate school by 4 percentage points (columns 5 and 6 of table 2). In relative terms, the cutoff effect on choosing baccalaureate school is 16%. As expected, controlling for additional covariates (in columns 2, 4, 6 and 8) does not matter for the RD coefficient estimates.

Table 3 around here rdest.track.robust2

The effects on non-entrance or entering a specialized school are close to zero. This changes, however, if we look at program choices that happen right after the end of compulsory schooling, i.e. in the same year. Table 3 shows that being born after the cutoff raises the probability to attend baccalaureate school immediately, but not the probability for apprenticeships. Instead, there are more students that do not yet start an upper secondary education ("no entry"). This result suggests that firms are also sensitive to the age of apprenticeship candidates and shy away from recruiting very young adolescents. Younger students are thus at a higher risk of having to wait a year until they can enter into a formal upper secondary program. In contrast, the results two years after the end of compulsory school, shown in columns (5) to (8) in Table 3, are almost identical to the results after one year.

Figure 3 around here ivdesc.personal

The findings above are driven by the students and parents who comply with the cutoff rule, so the effects hold only for this group. Therefore, we would like to know who are these compliers, compared to the students who always enroll early (always-takers) or late (never-takers). As noted before, we can look at school entry only in younger cohorts than those used in the reduced form estimations. Keeping this caveat in mind, figure 3 presents characteristic means for the three compliance groups based on Marbach and Hangartner (2020).[22] Among the never-takers (i.e. redshirters), men, Swiss nationals and students who speak the local language as first language are overrepresented compared to the full sample, whereas always-takers are more often women and foreigners. Among compliers, students who speak the local language as first language are slightly underrepresented.

6.2 Subgroup results by personal characteristics

Men are more likely to start an apprenticeship than women, who are comparatively more likely to attend a baccalaureate school, as indicated by the sample means in table 4. Nevertheless, there is little difference in the reduced form coefficients for program choice by sex.

Table 4 around here rdest.track.parentbackground

[22]Recall that in the reduced form setting, we assume that there are no defiers, see the discussion in section 4.2.

Differences are slightly more pronounced between students who speak the local language as first language and students who speak a foreign language. While the pattern is the same for both groups, students speaking the local language have higher effects at the cutoff. A possible explanation is that foreign speaking students often belong to groups that are not at the margin between academic and vocational education. Children of expats often attend baccalaureate school (lit.), whereas other immigrant families have more difficulties to excel in the school system and do not gain access to baccalaureate schools, which is reflected in the lower sample mean (lit.).

Finally, table 4 shows results by parents' education, which is available for a subsample only.[23] We distinguish between parents that completed compulsory or upper secondary education as highest education, and those where at least one parent has completed tertiary education. A bit surprisingly, we find no clear picture, as the point estimates for students from parents with tertiary education are *higher*, but the effects are estimated less precisely and not statistically different from those of the other students.

Altogether, we find similar age effects in different subgroups, implying that effect heterogeneity is limited when it comes to (observable) personal characteristics.

6.3 Subgroup results by cantonal institutions

In the literature on school starting age, one particular challenge is to distinguish between relative and absolute age effects (Peña, 2020). The age effects found in table 2 may be due to the fact that students born after the cutoff are older in absolute age than their counterparts born before the cutoff. However, the effects may also be due to the relative age difference between students within classes, because age relative to peers in class may affect students' development. Since cutoff dates differ by cantons, we observe students of different absolute ages in our RD sample, whereas in countries with one national cutoff, absolute and relative age are collinear and the two effects thus indistinguishable.

We compare cantons with an early (May or before) and late (after May) school enrollment cutoff in the first four columns of table 5. Implementing an early cutoff means that students are

[23] The information on parents' education has been merged from the Swiss "structural census", which is a random draw from the population.

younger on average when starting school.[24] The coefficients displayed in table 5 show that point estimates are substantially higher for cantons with late cutoff, i.e. in cantons where students are younger on average. This finding is consistent with an effect of absolute age, even if the 95% confidence intervals for the estimators (columns 1 and 3, columns 2 and 4) overlap. As an additional caveat, cantons enjoy substantial leeway in the implementation of their educational systems (e.g. on the amount and location of different types of schools at upper secondary and tertiary level), which makes it difficult to disentangle differences in students' absolute age from institutional differences between cantons. This difficulty exists also with the two next analyses on cantonal heterogeneity in table 5.

Table 5 around here rdest.track.subsample

A key institutional difference between cantons is the type of admission procedure for baccalaureate schools. Columns (5) to (8) of table 5 split the sample between cantons with baccalaureate school entrance exams and those without, where teachers decide about recommending a student for baccalaureate school. The cutoff effect among these cantons amounts to 5.3 percentage points for baccalaureate school entrance, while the point estimate among cantons with an entrance exam is substantially lower. A possible explanation is that teachers are more sensitive to age differences between students than exam results. However, the point estimate differences are not significant.

The last comparison in columns (9) to (12) in table 5 refers to cantons with and without "long" baccalaureate schools. These schools start before grade 9, usually in grade 7. They are more similar to the German model, where baccalaureate schools start in grade 5 already. Cantons with such schools feature hybrid systems, i.e. one part of students enter baccalaureate school early, another (and varying) part enters baccalaureate school later on. Our results show that cutoff effects are stronger in cantons with a system that knows only "short" baccalaureate schools. This is a bit surprising, because the sorting into academic and vocational education happens (partly) earlier in students' life in cantons with long baccalaureate school. On the other hand, these result are in line with the German evidence that found no age effects on completing

[24]The school year starts after the summer holidays in all cantons. There is a small variation in the end date of summer holidays that is not crucial for actual SSA.

academic or vocational education (section 2). A possible explanation is that systems with both short and long baccalaureate school offer students several points in time to enter, which may attenuate age effects.

All interpretations of heterogeneity by cantonal institutions so far hinge on assuming that there are no other differences between the respective two groups of cantons, which could drive the results. In any case, the results in table 5 suggest considerable effect heterogenity across cantons. In figure 5, we therefore look at the reduced form RD coefficients for the 13 largest cantons, which represent 85 percent of all adolescents in our sample, in descending order by sample size. With one exception (Argovia), the blue dots for apprenticeship are to the left of those for baccalaureate school, reflecting our main result that starting an apprenticeship is less likely when born after cutoff. However, the difference between the coefficients indeed varies by canton. In the cantons of Argovia and Saint Gall, the point estimates for apprenticeship and baccalaureate scholl are very close to zero. This is in line with Balestra, Eugster, and Liebert (2020), who found no difference in program choice for Saint Gall. In contrast, differences are substantial in the three largest cantons, where 95% confidence intervals do not overlap (Zurich and Bern) or hardly overlap (Vaud). The smaller the number of observations, the wider the intervals become.

Figure 4 around here coefplot.bb.gym.largecantons

Because cantons differ in many dimensions, it is difficult to explain the differences in cantonal results conclusively. The institutional settings investigated in table 5 provide some explanation. A further explanation is that populations in some cantons comply more with cutoff rules than others. Table (tbc) in the appendix shows compliance rates by cantons in the school entry sample, which vary between more than 80% (Graubunden, Neuchâtel, Geneva) and less than 30% (Ticino, Lucerne, Uri) if we look at persons born in a 30 days window before and after cutoff. These differences are a reflection of school systems that are rooted in cantonal history. For our estimations, we conclude that the the RD cutoff effects are driven by varying parts of the population in different cantons.

6.4 Types of apprenticeship programs

The negative effect of being born after the cutoff on starting an apprenticeship might be more pronounced in particular apprenticeship programs. In table 6, different types of apprenticeships are the dependent (dummy) variables, thus decomposing the overall negative cutoff effect on apprenticeship programs. Columns (1) to (4) show that the effect is strongest in the occupation field of economic and commercial apprenticeships, and second strongest in service sector apprenticeships. This makes sense from a learner's perspective, because apprenticeship programs to become a commercial employee have a good reputation (especially when served in banking and insurance companies) and are widely considered a good alternative to baccalaureate school. The four last columns use occupational demands in the training occupations as dependent variable.[25] The training occupation demands most affected are analytical and social skills, whereas programs demanding robust health or technical skills are hardly affected.

Table 6 around here rdest.track.occfields

In the appendix table A.5, we further show that the negative cutoff effect on entering apprenticeships mainly concerns apprenticeships with three year duration (which are the majority of programs) and apprenticeships without integrated vocational baccalaureate school. When we distinguish between firm-based and purely school-based apprenticeships, the cutoff effect for firm-based apprenticeships is almost the same as for all apprenticeship programs together, while it is small in absolute value and insignificant for the small amount (about 10%) of school-based apprenticeships (not shown).

6.5 Robustness results

First, we check whether our main results change when we include students based on birth cohorts instead of school leaving cohorts (see section 5.1). Table A.3 uses the two birth cohorts for which we have the most observations in the LAES data. Only students born in 1998 and

[25] For this purpose, we scraped the listed "demands" of the official Swiss portal for career guidance, berufsberatung.ch. It provides job postings for apprenticeships in all cantons and occupations. Additionally, it describes every existing training occupation (tasks, learning content, prerequisites, occupational demands, institutional framework). We sorted demands into four categories and one rest category. We then calculated an importance measurement for each demand category and normalized this to mean zero and a standard deviation of one. More precisely, we sum the listed demands for every category $j = (1, 2, 3, 4, 5)$ in occupation o and divide it by the sum of all listed demands in occupation o: $\sum d_{jo} / \sum d_o$.

1999 are included. For both cohorts, we find the same pattern, i.e. students born before cutoff are more likely to attend apprenticeships and less likely to attend baccalaureate school. Point estimates are somewhat higher than in our main analysis, suggesting that we are not overestimating coefficients by mixing several birth cohorts.

Next, we check whether covariates do not jump at the cutoff, which would violate the local continuity assumption. In the appendix, table A.4 shows RD results using covariates as dependent variables. The cutoff coefficients are not significant, except for marginal significance ($p = 0.098$) in the case of Swiss citizenship. We consider this a random result and control for citizenship in models among the personal characteristics. Dropping all foreigners from the sample does not change our results in any substantive way.

Finally, the RD design involves choices on bandwith and the specification of the trends to the left and the right of the cutoff. Figure 6 shows how cutoff effects estimated with and without controls change depending on these parameters. The first takeaway is that in estimations controlling for linear or quadratic polynomials of the running variable, we always find the same pattern as in the main analysis: a significant positive effect of being born after the cutoff on entering baccalaureate school, and an opposing negative effect on apprenticeships. This holds for all bandwiths and whether including covariates in the estimation or not. Confidence intervals of the two effects include zero only in "extreme" models, wich combine a fourth-order polynomial with a bandwith below 10 days. Such models allow for a very flexible fit of the trends around the cutoff in a relatively small sample, which is not recommended in the literature (Gelman and Imbens, 2019). The second takeaway is that point estimates of the cutoff effects increase in absolute value in specifications with small bandwiths of 20 days or less. Our main specification with 30 days might thus underestimate the true cutoff effects, because smaller bandwiths reduce the risk of misspecifying trends around the cutoff. However, smaller bandwiths come at the cost of less precise estimates.[26]

[26] Tables in the appendix complete results on different specifications, e.g. for the optimal bandwith selection procedure proposed by Calonico, Cattaneo, and Farrell (2020); Calonico, Cattaneo, Farrell, and Titiunik (2019). They also show results for a donut specification (in case parents can manipulate birth before or after midnight of the cutoff day), for a sample with students for whom we know the canton of residence before they turned age ten, and for the canton of birth recorded (which can be the canton of the hospital instead of the canton of parents' residence at birth).

7 Mediation analysis of early tracking

7.1 Early tracking

In most Swiss cantons, students are tracked after grade six at the intersection between primary and lower secondary level. While in most cantons (mostly in the western, north-western, and southern part of Switzerland) students are sorted into just two tracks ("real" and "secondary" school) at the lower secondary level, students in some other cantons have the additional option to enter a long-term baccalaureate school.

Figure 5 around here

Figure 5 shows the association between students' birth date (again with respect to the cantonal cutoff applying) and their lower secondary tracks. It becomes apparent that being born after the cutoff decreases (increases) the chances to finish the lower secondary level in the basic (advanced) track.[27]

Table A.10 shows transition patters from the lower to the upper secondary level. Though the links between the lower and the upper secondary tracks are not fixed, there are clear patterns: 90% of all baccalaureate school entrants stem from the higher track at lower secondary school. However, visiting the high track does not guarantee baccalaureate school entrance: 20% start an apprenticeship and a little more than 4% a specialized school. Apprentices, in contrast, come from different tracks, with the basic and the advanced track delivering 41% and 50% of all new apprentices.

Table A.10 around here

We assume that these associations operate through two channels: on the one hand, attending the advanced track and thus being exposed to more able peers throughout lower secondary school might increase students' chances of entering baccalaureate school or a higher-quality apprenticeship place. On the other hand, the lower secondary track is a negative signal for

[27]Note that some students visit schools with no track distinction across classes, or attend schools for students with special needs. Hence, the two shares in (a) and (b) of figure 5 do not sum up to one.

teachers deciding about recommending students for baccalaureate school as well as for recruiters in training firms.[28]

Overall, the strong association between students' lower secondary track and their track entrance probabilities on the upper secondary level requires to take the first tracking into consideration when analyzing the latter transition.

7.2 Mediation analysis

In the mediation analysis, we divide the total reduced form age effect into an indirect effect due to early tracking in lower secondary school, and a direct effect from school starting age on the choice between academic and vocational education at upper secondary level. For this purpose, we reduce the sample to students that were in lower secondary school in 2011 or 2012 (see section 5.1).

Table 7 around here medeff.track.x

The models estimated in Table 7 assume sequential ignorability (Imai, Keele, and Yamamoto, 2010) of the treatment (being born before or after the cutoff) and the mediator (early tracking in lower secondary school). The total or overall effect for baccalaureate school is 4.7 percentage points (in the specification with a bandwith of 30 days), a bit higher than the reduced form estimates in our main specification. The tracking in lower secondary school is statistically significant for the selection into baccalaureate school at upper secondary level. It also accounts for 1.3 percentage points of the total effect on baccalaureate school. Thus, a bit less than one third of the age effect is mediated by early tracking, while two thirds are a direct effect. Direct effects comprise all further channels how school starting age affects track choice at upper secondary level, apart from early tracking. A likely part of such a direct effect is the effect of the age at the time of choosing between academic and vocational education, which is strongly influenced by school starting age.

As described in secion 4.2, it seems unlikely that the assumption of sequential ignorability is met. In order to eliminate a possible bias from unobserved variables correlated with the me-

[28] In this context, it is important to stress that the tracking after the primary level is usually not fixed throughout the lower secondary level and students showing a positive (negative) development can switch to a higher (lower) track, which should damp any effect related to the tracking. However, since we only observe the lower secondary track at the end of compulsory schooling in our data, we take such switches into account.

diator and the outcome (or cutoff and mediator), we instrument the mediator "lower secondary track" with the share of students attending the advanced track in the municipality where a student entered lower secondary school. The distribution of this variable is shown in figure A.1. In addition, we control for the proportion of parents in the municipality that have an upper secondary or tertiary degree, in order to account for residential sorting.

Table 8 around here ivregress.mediation

The first stage regression in Table 8 (column 1) shows, first, that being born after the cutoff increases the likelihood of attending the advanced track. Secondly, it shows that the municipality share of advanced track students affects the likelihood to enter the advanced track strongly. In the models with program choice as dependent variable, the reduced form estimates without mediator (columns 2 and 4) again show a cutoff effect of around 4.7 percentage points on apprenticeships and baccalaureate school. In columns (4) and (5), we add the instrumented dummy for attending the advanced track, which is highly significant. Using these models, we calculate the indirect effect on apprenticeships as -0.017 and the direct effect as -0.029. The indirect effect on baccalaureate school is 0.014, the direct effect 0.033. Thus, the indirect effect again accounts for roughly one third of the overall effect. There remains a substantial direct effect of the reduced form age effect on the sorting into upper secondary programs.

8 Discussion and conclusion

In this paper, we examine the effect of enrollment cutoffs and resulting age differences on the choice between academic and vocational education in adolescence. We use a regression discontinuity design with data on the universe of all Swiss students, which is available for the time period 2012-2019. Reduced form estimations show substantial effects: being born right before the respective cantonal cutoff date decreases the likelihood to attend baccalaureate school by 3.6 percentage points (or 16%) and increases the likelihood for an apprenticeship by 3.3 percentage points. These findings are robust to varying bandwith choices, i.e. to more extensive or restrictive sample choice with respect to the allowed distance between cutoff and students' birthday, as well as to different polynomial specifications.

The results show no variation by gender and are similar by parental education. However, the reduced form estimates are somewhat larger for students whose first language is the local language. They are also larger in cantons with late cutoffs (and thus younger students at enrollment) and in cantons without entry exams for baccalaureate schools, which use teacher recommendations instead. The latter finding suggests that teachers use age as a signal of maturity for baccalaureate school, even though results relying on cantonal differences have to be treated with caution due to the many cantonal differences in educational systems.

Those marginal younger students that do not enter baccalaureate school often seem to attend commercial and, to a lesser extent, other service sector apprenticeships. This is plausible because commercial apprenticeships are considered as a good alternative to baccalaureate school by many Swiss. This finding may also explain the difference to Germany, where several papers have not found SSA effects on the likelihood to attain a vocational degree (Dustmann, Puhani, and Schönberg, 2017; Görlitz, Penny, and Tamm, 2022): in Germany, commercial employees are often students who have completed baccalaureate school and then add a highly recognized apprenticeship at a bank or insurance company, for example. This is very rare in Switzerland, where students typically begin apprenticeships, even prestigious ones, after completing compulsory school without attending baccalaureate schools. Finally, our findings are also consistent with the results in Balestra, Eugster, and Liebert (2020), who find no effect of SSA on upper secondary track in the canton of St. Gall. Indeed, we do not find a significant effect in St. Gall either, but we find sizeable effects in many other cantons, which sum to the reported effects for Switzerland as a whole.

Given earlier results on the effect of age on early tracking in Germany, one may object that the effects of age on the choice between academic and vocational education in adolescence reported so far are predetermined in lower secondary school. There, students are already tracked in a basic and an advanced track based on their school performance. When we plot early track shares against the distance of birthday to cutoff, we indeed find a strong discontinuity at the cutoff. Therefore, we perform a mediation analysis to find out how much of the total effect of age on upper secondary program is due to an indirect effect of age on early tracking. We find that early tracking also predicts education programs at upper secondary level to a considerable extent and accounts for a third of the total age effect. The remaining direct effect thus accounts for almost two thirds of the age effect. Therefore, the moment of choice between academic and vocational education in adolescence emerges as a junction in Swiss students' lives that is influenced by earlier tracking, but also by direct age effects caused by cutoff rules and the one-year rhythm of school classes. This direct effect is consistent with an effect of age at the time of choosing between academic and vocational education, especially with an effect of visible "maturity" as a signal for teachers when making recommendations for baccalaureate schools, and for firms when recruiting for apprenticeship positions.

The main takeaway from our paper is that seemingly small age differences may influence decisions about educational pathways in adolescence, likely reinforced or mitigated by institutional settings. This does not necessarily mean that systems should not allow educational decisions in adolescence. Deciding on one's educational and occupational career is never easy, and postponing such decisions may come with its own problems. However, it suggests that adolescents need support in educational choices, that gate keepers need to account for the process of adolescence, including the fact that maturity is still evolving and correlated with age in adolescence, and that education systems need to be permeable to allow for trial and error and correction of earlier decisions. Future research should address whether the age effect on sorting into academic and vocational programs carries over to different labor market outcomes.

References

Aepli, M., A. Kuhn, and J. Schweri (2021): "Der Wert von Ausbildungen auf dem Schweizer Arbeitsmarkt," Grundlagen für die Wirtschaftspolitik Nr. 31, Staatssekretariat für Wirtschaft SECO.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91(434), 444.

Angrist, J. D., and A. B. Krueger (1992): "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 87(418), 328.

Attar, I., and D. Cohen-Zada (2018): "The effect of school entrance age on educational outcomes: Evidence using multiple cutoff dates and exact date of birth," *Journal of Economic Behavior & Organization*, 153, 38–57.

Bahrs, M., and M. Schumann (2020): "Unlucky to be young? The long-term effects of school starting age on smoking behavior and health," *Journal of Population Economics*, 33(2), 555–600.

Balestra, S., B. Eugster, and H. Liebert (2020): "Summer-born struggle: The effect of school starting age on health, education, and work," *Health Economics*, 29(5), 591–607.

Baron, R. M., and D. A. Kenny (1986): "The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations," *Journal of Personality and Social Psychology*, 51(6), 1173–1182.

Barua, R., and K. Lang (2016): "School Entry, Educational Attainment, and Quarter of Birth: A Cautionary Tale of a Local Average Treatment Effect," *Journal of Human Capital*, 10(3), 347–376.

Bedard, K., and E. Dhuey (2006): "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects*," *The Quarterly Journal of Economics*, 121(4), 1437–1472.

Bertrand, M., M. Mogstad, and J. Mountjoy (2021): "Improving Educational Pathways to Social Mobility: Evidence from Norway's Reform 94," *Journal of Labor Economics*, 39(4), 965–1010.

Birkelund, J. F., and H. G. van de Werfhorst (2022): "Long-term labor market returns to upper secondary school track choice: Leveraging idiosyncratic variation in peers' choices," *Social Science Research*, 102, 102629.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011): "Too Young to Leave the Nest? The Effects of School Starting Age," *Review of Economics and Statistics*, 93(2), 455–467.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020): "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *Econometrics Journal*, 23(2), 192–210.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017): "Rdrobust: Software for Regression-discontinuity Designs," *The Stata Journal: Promoting communications on statistics and Stata*, 17(2), 372–404.

——— (2019): "Regression Discontinuity Designs Using Covariates," *The Review of Economics and Statistics*, 101(3), 442–451.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2019): *A Practical Introduction to Regression Discontinuity Designs*, Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press, Cambridge.

Chuan, A., and C. L. Ibsen (2022): "Skills for the Future? A Life Cycle Perspective on Systems of Vocational Education and Training," *ILR Review*, 75(3), 638–664.

Dhuey, E., D. Figlio, K. Karbownik, and J. Roth (2019): "School Starting Age and Cognitive Development," *Journal of Policy Analysis and Management*, 38(3), 538–578.

Dustmann, C., P. A. Puhani, and U. Schönberg (2017): "The Long–Term Effects of Early Track Choice," *The Economic Journal*, 127(603), 1348–1380.

Elder, T. E., and D. H. Lubotsky (2009): "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *The Journal of Human Resources*, 44(3), 641–683.

Eriksson, A. W., and J. Fellman (2000): "Seasonal variation of livebirths, stillbirths, extramarital births and twin maternities in Switzerland," *Twin Research and Human Genetics*, 3(04), 189–201.

Fiorini, M., and K. Stevens (2021): "Scrutinizing the Monotonicity Assumption in IV and fuzzy RD designs," *Oxford Bulletin of Economics and Statistics*, 83(6), 1475–1526.

Fredriksson, P., and B. Öckert (2014): "Life-cycle Effects of Age at School Start," *The Economic Journal*, 124(579), 977–1004.

Frölich, M., and M. Huber (2017): "Direct and indirect treatment effects–causal chains and mediation analysis with instrumental variables," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 79(5), 1645–1666.

Gangl, S., and M. Huber (2021): "From homemakers to breadwinners? How mandatory kindergarten affects maternal labour market outcomes," arxiv.

Gelman, A., and G. Imbens (2019): "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 37(3), 447–456.

Golsteyn, B. H. H., and A. Stenberg (2017): "Earnings over the Life Course: General versus Vocational Education," *Journal of Human Capital*, 11(2), 167–212.

Görlitz, K., M. Penny, and M. Tamm (2022): "The long-term effect of age at school entry on cognitive competencies in adulthood," *Journal of Economic Behavior & Organization*, 194, 91–104.

Hanushek, E. A., G. Schwerdt, L. Woessmann, and L. Zhang (2017): "General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle," *Journal of Human Resources*, 52(1), 48–87.

Hartog, J., P. Raposo, and H. Reis (2022): "Fluctuations in the wage gap between vocational and general secondary education: lessons from Portugal," *Journal of Population Economics*, 35(2), 643–675.

Huber, M. (2021): "Mediation Analysis," in *Handbook of Labor, Human Resources and Population Economics*, ed. by K. F. Zimmermann, pp. 1–38. Springer, Cham.

Imai, K., L. Keele, and T. Yamamoto (2010): "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects," *Statistical Science*, 25(1).

Imbens, G., and K. Kalyanaraman (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, 79(3), 933–959.

Imbens, G. W., and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615–635.

Imdorf, C. (2012): "Zu jung oder zu alt für eine Lehre? Altersdiskriminierung bei der Ausbildungsplatzvergabe," *Journal for Labour Market Research*, 45(1), 79–98.

Johansen, E. R. (2021): "Relative age for grade and adolescent risky health behavior," *Journal of Health Economics*, 76, 102438.

Korber, M., and D. Oesch (2019): "Vocational versus general education: Employment and earnings over the life course in Switzerland," *Advances in Life Course Research*, 40, 1–13.

Largo, R. H., T. Gasser, A. Prader, W. Stuetzle, and P. J. Huber (1978): "Analysis of the adolescent growth spurt using smoothing spline functions," *Annals of Human Biology*, 5(5), 421–434.

Lubotsky, D., and R. Kaestner (2016): "Do 'Skills Beget Skills'? Evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood," *Economics of Education Review*, 53, 194–206.

Malamud, O., and C. Pop-Eleches (2010): "General Education versus Vocational Training: Evidence from an Economy in Transition," *Review of Economics and Statistics*, 92(1), 43–60.

Matthewes, S. H. (2021): "Better Together? Heterogeneous Effects of Tracking on Student Achievement," *The Economic Journal*, 131(635), 1269–1307.

MiSDV (2017): "Mittelschuldirektionsverordnung: Erziehungsdirektion des Kantons Bern," Bern, switzerland.

Mühlenweg, A. M., and P. A. Puhani (2010): "The Evolution of the School-Entry Age Effect in a School Tracking System," *The Journal of Human Resources*, 45(2), 407–438.

Navarro, J.-J., J. García-Rubio, and P. R. Olivares (2015): "The Relative Age Effect and Its Influence on Academic Performance," *PloS one*, 10(10), e0141895.

Peña, P. A. (2020): "Relative age and investment in human capital," *Economics of Education Review*, 78, 102039.

Puhani, P. A., and A. M. Weber (2007): "Does the early bird catch the worm?," *Empirical Economics*, 32(2-3), 359–386.

Rohrbach-Schmidt, D., and A. Uhly (2015): "Determinanten vorzeitiger Lösungen von Ausbildungsverträgen und berufliche Segmentierung im dualen System. Eine Mehrebenenanalyse auf Basis der Berufsbildungsstatistik," *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 67(1), 105–135.

Saltiel, F. (2021): "Fast-Tracked to Success: Evidence on the Returns to Vocational Education in Switzerland," *Annals of Economics and Statistics*, (142), 5–44.

Schweri, J., M. Aepli, and A. Kuhn (2021): "The costs of standardized apprenticeship curricula for training firms," *Empirical Research in Vocational Education and Training*, 13(1).

Silliman, M., and H. Virtanen (2022): "Labor Market Returns to Vocational Secondary Education," *American Economic Journal: Applied Economics*, 14(1), 197–224.

Sprietsma, M. (2010): "Effect of relative age in the first grade of primary school on long–term scholastic results: international comparative evidence using PISA 2003," *Education Economics*, 18(1), 1–32.

Wolter, S. C., M. A. Cattaneo, S. Denzler, A. Diem, S. Hof, R. Meier, and C. Oggenfuss (2018): *Bildungsbericht Schweiz 2018*. SKBF/CSRE, Aarau.

9 Tables

Table 1: Descriptive statistics

	Full sample			Bandwidth 30 days		
	N	Mean	s.d.	N	Mean	s.d.
Program choice at upper secondary level (1 year after leaving comp. school):						
No entry (yes=1)	447,949	0.088		74,007	0.087	
Enters apprenticeship (yes=1)	447,949	0.616		74,007	0.614	
Enters baccalaureate school (yes=1)	447,949	0.244		74,007	0.247	
Enters specialized school (yes=1)	447,949	0.052		74,007	0.052	
Enters upper sec. level late (yes=1)	447,949	0.229		74,007	0.228	
Age at entering upper secondary education*	423,465	16.1	0.900	69,998	16.1	0.869
Demographic information:						
Female (yes=1)	447,949	0.486		74,007	0.487	
Swiss citizen (yes=1)	447,949	0.859		74,007	0.858	
Not born in Switzerland (yes=1)	447,949	0.085		74,007	0.084	
Local language = first language (yes=1)	447,516	0.786		73,929	0.784	
Parents' education: tertiary level (yes=1)	49,513	0.440		8,305	0.435	
Information on cutoffs:						
Days born before/after cutoff	447,949	0.485	105.2	74,007	-0.138	17.7
Born after cutoff (yes=1)	447,949	0.502		74,007	0.496	
Cutoff January 1 (yes=1)	447,949	0.027		74,007	0.026	
Cutoff May 1 (yes=1)	447,949	0.577		74,007	0.580	
Cutoff June 1 (yes=1)	447,949	0.023		74,007	0.022	
Cutoff July 1 (yes=1)	447,949	0.091		74,007	0.094	
Cutoff August 1 (yes=1)	447,949	0.073		74,007	0.074	
Cutoff September 1 (yes=1)	447,949	0.024		74,007	0.023	
Cutoff October 1 (yes=1)	447,949	0.040		74,007	0.040	
Cutoff November 1 (yes=1)	447,949	0.106		74,007	0.103	
Cutoff December 31 (yes=1)	447,949	0.038		74,007	0.037	
Age when leaving compulsory school:						
Comp. school leaving age in years	447,949	15.889	0.655	74,007	15.933	0.687
Predicted leaving age in years	447,949	15.536	0.364	74,007	15.532	0.508
Leaves school after predicted year (yes=1)	447,949	0.327		74,007	0.381	
Leaves school before predicted year (yes=1)	447,949	0.028		74,007	0.049	
Track at lower secondary school:						
Special curriculum (yes=1)	447,949	0.038		74,007	0.037	
Basic track (yes=1)	447,949	0.271		74,007	0.264	
Advanced track (yes=1)	447,949	0.645		74,007	0.652	
Non-tracked (yes=1)	447,949	0.044		74,007	0.046	
Region of lower secondary school:						
School in urban area (yes=1)	447,949	0.571		74,007	0.571	
School in intermediate area (yes=1)	447,949	0.240		74,007	0.241	
School in rural area (yes=1)	447,949	0.188		74,007	0.187	
School in German speaking region (yes=1)	447,949	0.713		74,007	0.713	
School in French speaking region (yes=1)	447,949	0.245		74,007	0.247	
School in Italian speaking region (yes=1)	447,949	0.040		74,007	0.388	

Notes: table.desc. Due to space restrictions we do not display all cantonal dummies, only the cantonal cutoff dates, which group cantons. Sample restricted to students leaving school in 2012-2017.

* "Age at entering upper secondary education" has lower N due to people not entering upper secondary education until the end of the observation period.

Table 2: RD regression – upper secondary program choice (main specification)

	No entry (yes=1)		Apprenticeship (yes=1)		Bacc. school (yes=1)		Spec. school (yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample mean	0.087	0.087	0.614	0.614	0.247	0.247	0.052	0.052
RD estimate: after cutoff	-0.008	-0.007	-0.035***	-0.035***	0.041***	0.040***	0.002	0.002
	(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.007)	(0.004)	(0.004)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Regional characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Total observations	74007	74007	74007	74007	74007	74007	74007	74007
Observations used (LHS)	37325	37325	37325	37325	37325	37325	37325	37325
Observations used (RHS)	36682	36682	36682	36682	36682	36682	36682	36682

Notes: rdest.track.rdrobust.bw30. Each column displays a separate local polynomial regression discontinuity estimation using the rdrobust command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) of the student's lower secondary school. Sample restricted to students leaving school in 2012-2017 and a bandwidth of 30 days around the cutoffs. Students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: RD regression – program choice 0 and 2 years after compulsory school

	Track choice within 0 years				Track choice within 2 years			
	No entry (yes=1) (1)	Apprenticeship (yes=1) (2)	Bacc. school (yes=1) (3)	Spec. school (yes=1) (4)	No entry (yes=1) (5)	Apprenticeship (yes=1) (6)	Bacc. school (yes=1) (7)	Spec. school (yes=1) (8)
Sample mean	0.228	0.486	0.241	0.046	0.060	0.643	0.246	0.051
RD estimate: after cutoff	-0.040***	-0.005	0.038***	0.006*	-0.005	-0.037***	0.040***	0.002
	(0.007)	(0.008)	(0.007)	(0.003)	(0.004)	(0.008)	(0.008)	(0.004)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	74007	74007	74007	74007	61916	61916	61916	61916
Observations used (LHS)	37325	37325	37325	37325	31316	31316	31316	31316
Observations used (RHS)	36682	36682	36682	36682	30600	30600	30600	30600

Notes: rddummy.sec2.0y.2y Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). The dependent variables are dummies for program choice 0 or 2 years after leaving compulsory school. Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. In columns (5)-(8) we exclude students with leaving year 2017 since we cannot observe them for at least 2 years after finishing compulsory schooling. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: RD regression – program choice by personal characteristic subgroups

	Men		Women		Student language: first = local		Student language: first = foreign		Parent educ.: below tertiary		Parent educ.: tertiary	
	App. (1)	Bacc. school (2)	App. (3)	Bacc. school (4)	App. (5)	Bacc. school (6)	App. (7)	Bacc. school (8)	App. (9)	Bacc. school (10)	App. (11)	Bacc. school (12)
Sample mean	0.688	0.206	0.537	0.290	0.604	0.268	0.651	0.169	0.708	0.154	0.469	0.432
RD estimate: after cutoff	-0.031***	0.043***	-0.039***	0.036***	-0.039***	0.045***	-0.025	0.024*	-0.047*	0.056**	-0.060*	0.055
	(0.010)	(0.009)	(0.011)	(0.010)	(0.009)	(0.008)	(0.016)	(0.012)	(0.028)	(0.023)	(0.035)	(0.036)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	37986	37986	36021	36021	57939	57939	15990	15990	4690	4690	3615	3615
Observations used (LHS)	19206	19206	18119	18119	29237	29237	8056	8056	2255	2255	1721	1721
Observations used (RHS)	18780	18780	17902	17902	28702	28702	7934	7934	2435	2435	1894	1894

Notes: rdest.track.parentbackground. Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017 and being recorded in the Structural Survey. students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: RD regression – program choice by cantonal institutions

	Early cutoff cantons		Late cutoff cantons		Cantons w/o exam admiss.		Cantons with exam admiss.		Cantons w/o long bacc. school		Cantons with long bacc. school	
	App. (1)	Bacc. school (2)	App. (3)	Bacc. school (4)	App. (5)	Bacc. school (6)	App. (7)	Bacc. school (8)	App. (9)	Bacc. school (10)	App. (11)	Bacc. school (12)
Sample mean	0.674	0.206	0.522	0.309	0.545	0.291	0.712	0.185	0.569	0.271	0.702	0.201
RD estimate: after cutoff	-0.031***	0.032***	-0.042***	0.052***	-0.041***	0.053***	-0.026**	0.021**	-0.039***	0.047***	-0.028**	0.025**
	(0.010)	(0.009)	(0.012)	(0.012)	(0.010)	(0.009)	(0.012)	(0.010)	(0.009)	(0.009)	(0.013)	(0.011)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	44887	44887	29120	29120	43308	43308	30699	30699	48742	48742	25265	25265
Observations used (LHS)	22738	22738	14587	14587	21695	21695	15630	15630	24460	24460	12865	12865
Observations used (RHS)	22149	22149	14533	14533	21613	21613	15069	15069	24282	24282	12400	12400

Notes: rdest.track.subsample. Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: RD regression – program choice: apprenticeship occupation

	Apprenticeship (yes= 1)				Normalized share of respective occupational demand			
	Economic/ info./comm. (1)	Engineering/ industry (2)	Health/ social (3)	Service sector (4)	Robust health (5)	Social skills (6)	Work techniques/ analytical (7)	Handcraft/ technical (8)
Sample mean	0.242	0.228	0.087	0.053	-0.003	-0.006	0.005	-0.006
RD estimate: after cutoff	-0.022***	-0.008	0.001	-0.008**	0.012	-0.057***	-0.062***	-0.016
	(0.007)	(0.006)	(0.005)	(0.004)	(0.016)	(0.016)	(0.016)	(0.015)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	74007	74007	74007	74007	74007	74007	74007	74007
Observations used (LHS)	37325	37325	37325	37325	37325	37325	37325	37325
Observations used (RHS)	36682	36682	36682	36682	36682	36682	36682	36682

Notes: rdest.track.occfields. Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: RD regression – program choice, mediation estimates

	No entry (yes=1)		Apprenticeship (yes=1)		Bacc. school (yes=1)		Spec. school (yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample mean	0.063	0.065	0.626	0.627	0.255	0.254	0.055	0.054
Born after cutoff (yes=1)	-0.015	-0.001	-0.006	-0.036***	0.016	0.034***	0.005	0.003
	(0.014)	(0.008)	(0.024)	(0.014)	(0.022)	(0.012)	(0.013)	(0.007)
Lower sec.: advanced track	-0.096***	-0.101***	-0.274***	-0.260***	0.332***	0.327***	0.039***	0.034***
	(0.009)	(0.006)	(0.013)	(0.007)	(0.009)	(0.006)	(0.006)	(0.004)
Overall effect	-0.021	-0.005	-0.020	-0.047	0.033	0.047	0.007	0.004
Lower bound	-0.048	-0.020	-0.068	-0.074	-0.013	0.022	-0.019	-0.010
Higher bound	0.007	0.010	0.028	-0.020	0.075	0.071	0.032	0.018
Direct effect	-0.016	-0.001	-0.006	-0.036	0.016	0.034	0.005	0.003
Lower bound	-0.043	-0.016	-0.054	-0.063	-0.028	0.010	-0.021	-0.011
Higher bound	0.010	0.013	0.039	-0.011	0.057	0.057	0.030	0.016
Indirect effect	-0.005	-0.004	-0.014	-0.010	0.017	0.013	0.002	0.001
Lower bound	-0.010	-0.007	-0.027	-0.017	0.001	0.005	0.000	0.000
Higher bound	-0.000	-0.001	-0.001	-0.004	0.033	0.022	0.004	0.002
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Days from cutoff	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Days × after cutoff (yes=1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	10 days	30 days	10 days	30 days	10 days	30 days	10 days	30 days
Number of observations	5872	17894	5872	17894	5872	17894	5872	17894

Notes: medeff.track.x. Each column displays a separate estimation using the medeff-command (OLS) in Stata and regressing the upper secondary level track dummies on the exogenous treatment (born after cutoff (yes=1)), the mediator (lower sec. school: advanced track (yes=1)), and control variables as indicated. Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural), the language region (German, French, Italian), and the regional share of parents with an upper secondary or tertiary education degree, respectively; all based on the municipality (the share of parents education on the cantonal level in case we observe less than 100 cases in a municipality) where the student's primary school (grade six) is located at. Sample restricted to students attending grade six in 2011 or 2012 and leaving compulsory school 2015-2017. Students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Cantons of Nidwalden, Basel-Landschaft, Aargau, Ticino, and Vallais excluded. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

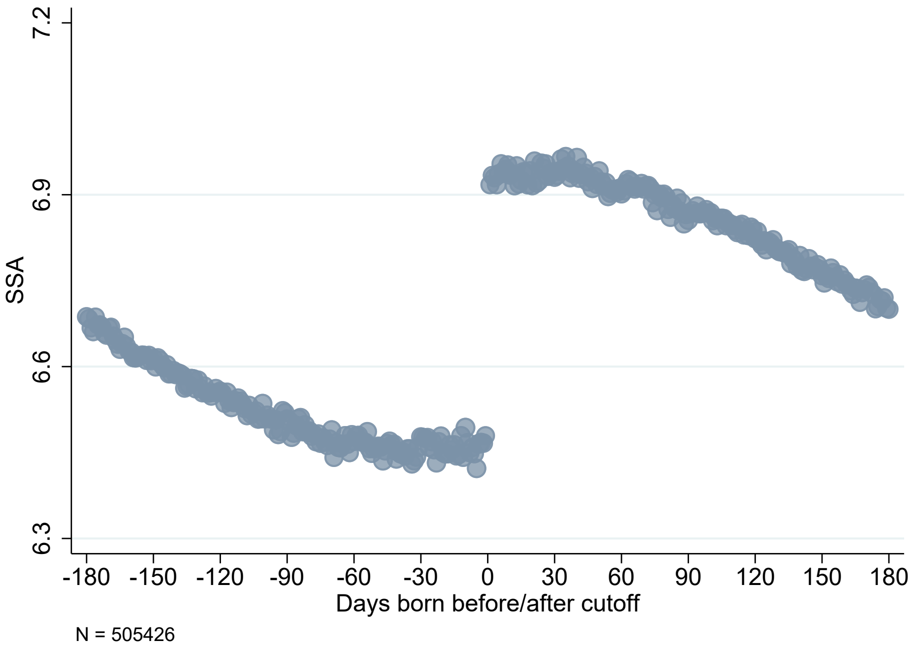
Table 8: Reduced form RD with instrumented mediator

	First stage	Entering apprenticeship		Entering bacc. school	
	(1)	(2)	(3)	(4)	(5)
Sample mean	0.697	0.614	0.614	0.263	0.263
Born after cutoff (yes=1)	0.042***	-0.046***	-0.031**	0.047***	0.035***
	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)
Lower sec. track: advanced			-0.398***		0.344***
			(0.083)		(0.070)
Municip. share entering adv. track	0.707***				
	(0.115)				
Constant	-0.567***	0.545***	-0.358***	0.059	-0.014
	(0.082)	(0.092)	(0.093)	(0.077)	(0.076)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes
Regional characteristics (6th grade)	Yes	Yes	Yes	Yes	Yes
Days from cutoff	Yes	Yes	Yes	Yes	Yes
Days $\times$ after cutoff (yes=1)	Yes	Yes	Yes	Yes	Yes
Bandwidth	30 days	30 days	30 days	30 days	30 days
R-squared	0.075	0.147	0.190	0.087	0.198
Observations	17894	17894	17894	17894	17894
F-statistic	117.1				

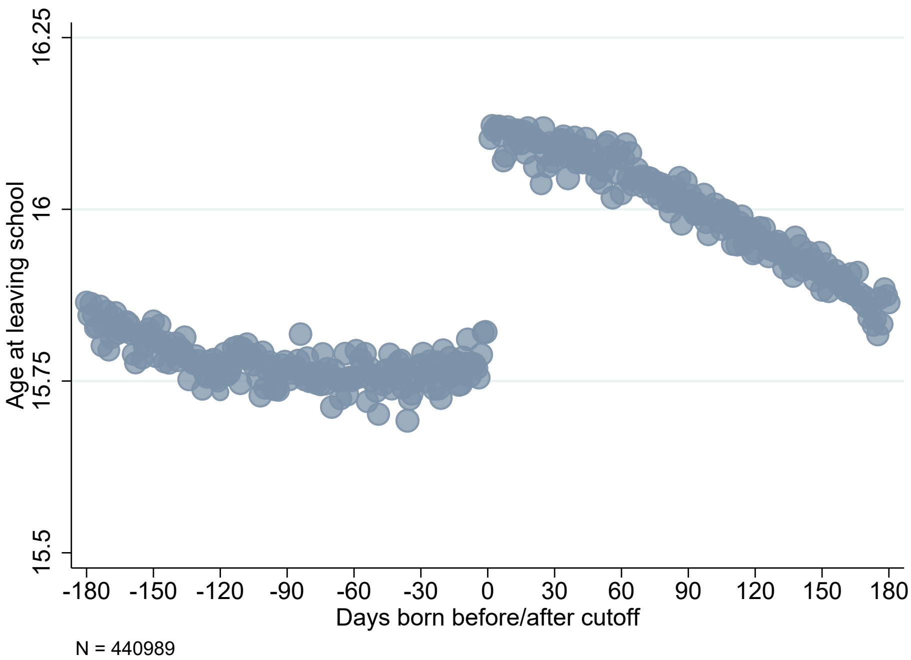
Notes: ivregress.mediation. Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural), the language region (German, French, Italian), and the regional share of parents with an upper secondary or tertiary education degree, respectively; all based on the municipality (the share of parents' education on the cantonal level in case we observe less than 100 cases in a municipality) where the student's primary school (grade six) is located at. Sample restricted to students attending grade six in 2011 or 2012 and leaving compulsory school 2015-2017. Students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Cantons of Nidwalden, Basel-Landschaft, Aargau, Ticino, and Vallais excluded. [a]The dependent variable in the first stage regressions is a dummy being one if a students attends the advanced track on the lower secondary level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10 Figures

Figure 1: Students' age at start and end of compulsory school



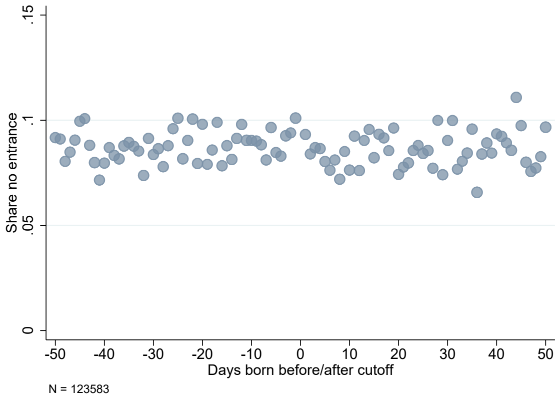
(a) School starting age



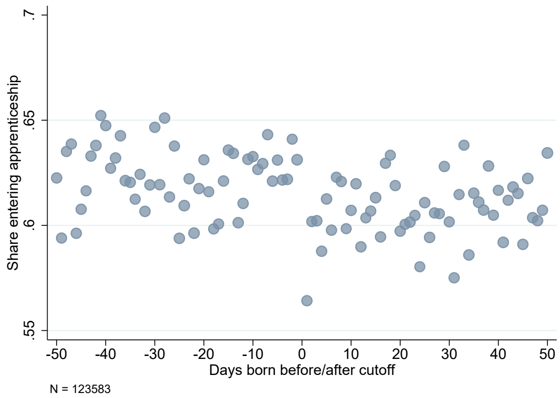
(b) School leaving age

Notes: The figure plots the students' age in days when they (a) start and (b) leave compulsory school against the difference between their birthdate and the cutoff date of their canton. Sample restricted to (a) students starting school in 2011-2014 and (b) leaving school in 2012-2017 and entering the upper secondary level maximum one year after finishing compulsory schooling.
Source: LAES data.

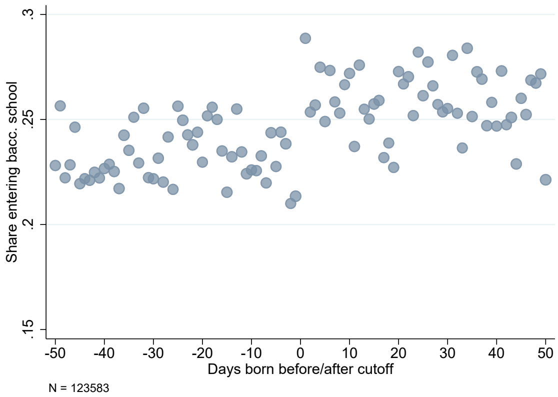
Figure 2: Program choice



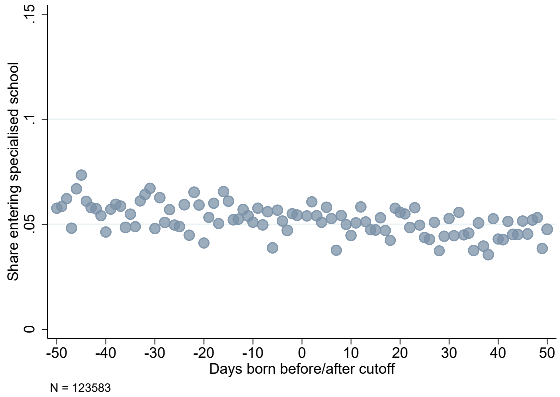
(a) No entrance



(b) Apprenticeship



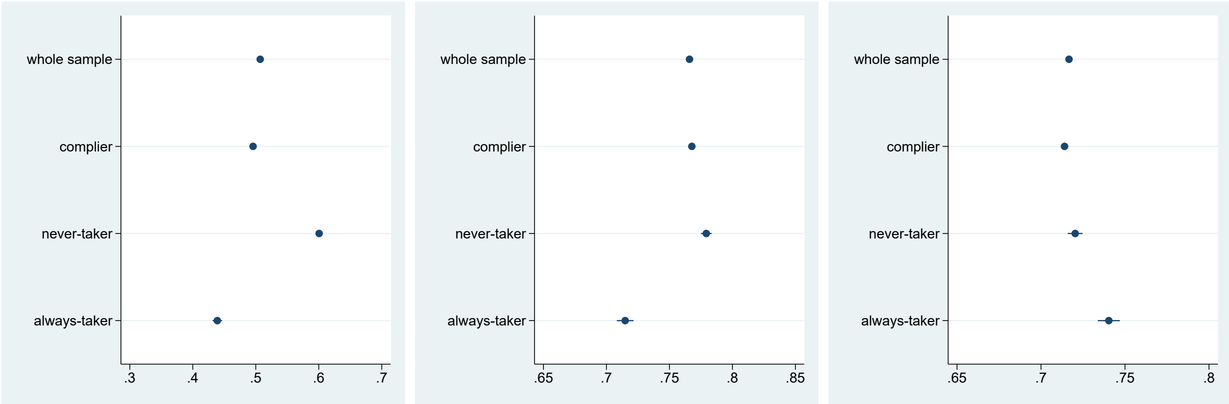
(c) Baccalaureate school



(d) Specialised school

Notes: The figure plots the share of students – aggregated by date of birth in daily bins – entering different tracks after compulsory schooling against the difference between their birthday and the cutoff date of the canton they were born in. Sample restricted to students leaving school in 2012-2017 and entering the upper secondary level maximum one year after finishing compulsory schooling.
Source: LABB data.

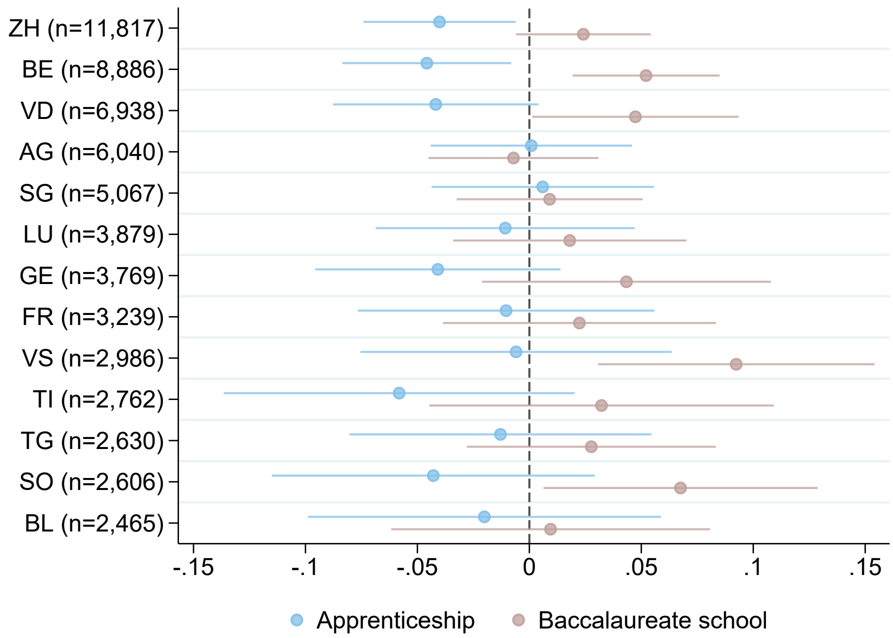
Figure 3: Means of personal characteristics in compliance groups



(a) Men (b) Swiss (c) First = local language

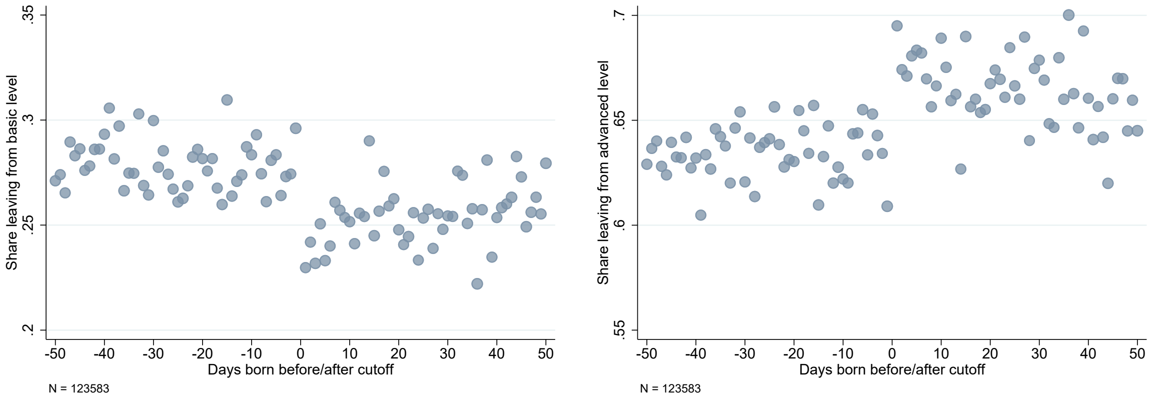
Notes: Figure plots the means of the respective personal characteristic by compliance group. Means calculated with the STATA command ivdesc (Marbach and Hangartner 2020).
Source: LABB data.

Figure 4: RD effects in the 13 largest cantons



Notes: The figure shows point estimates and 95% confidence intervals from separate local polynomial regression discontinuity estimations of program choice on cutoff variables and covariates, estimated separately for the 13 cantons with the largest number of observations. All other specifications are identical to those for the models in table 2 that include covariates. Sample restricted to students leaving school in 2012-2017. Students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries.
Source: LABB data.

Figure 5: Early tracking in lower secondary school around cutoff

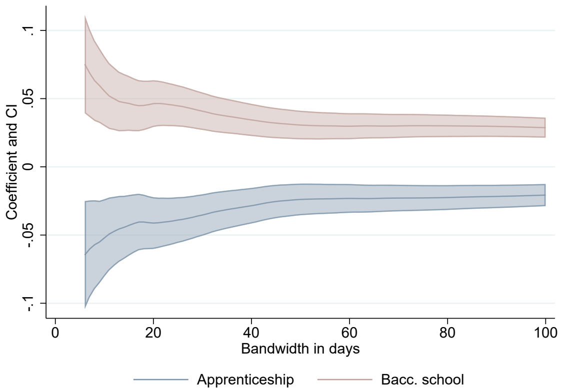


(a) Leaving track: basic

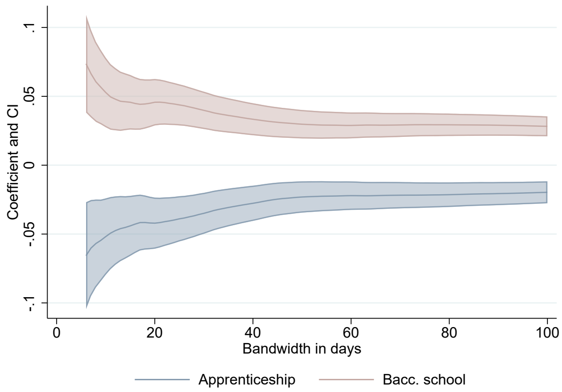
(b) Leaving track: advanced

Notes: The figure plots the share of students – aggregated by date of birth – being in the basic or advanced lower secondary track against the difference between their birthday and the cutoff date of the canton they were born in. Sample restricted to students leaving school in 2012-2017 and entering the upper secondary level maximum one year after finishing compulsory schooling.
Source: LABB data.

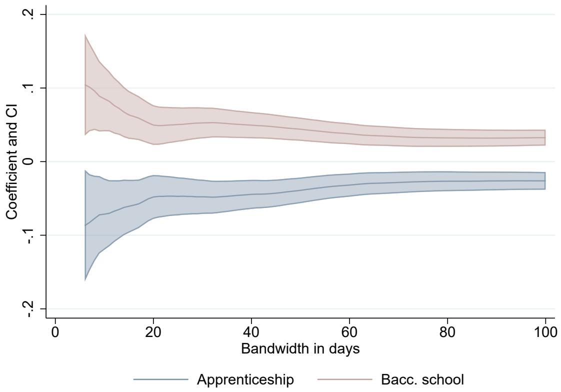
Figure 6: RD coefficients for apprenticeship and bacc. school over bandwidth



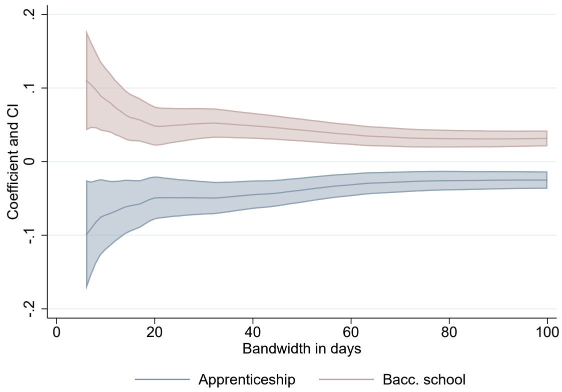
(a) Linear, w/o controls



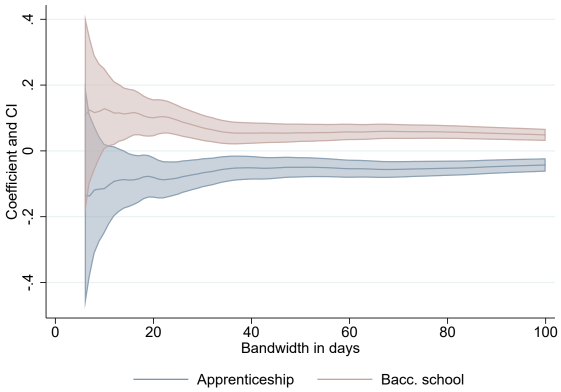
(b) Linear, with controls



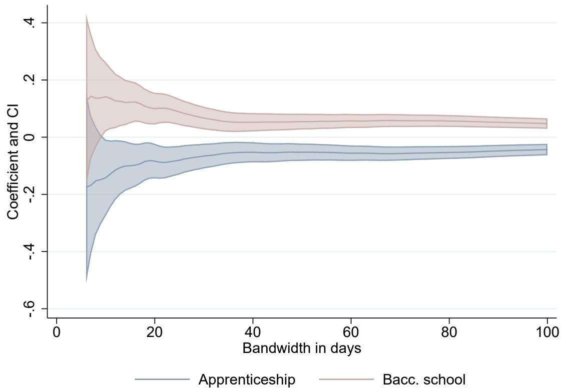
(c) Second polynomial, w/o controls



(d) Second polynomial, with controls



(e) Fourth polynomial, w/o controls



(f) Fourth polynomial, with controls

Notes: The figure shows reduced form estimates derived from equation 1 (with and without covariates) using the rdrobust-command, applying different bandwidths around the cutoff and different polynomials of the running variable (difference between students' birthday and the cutoff in days).
Source: LABB data.

A Appendix: Additional tables and figures

Table A.1: Hazard rate by cohort

	In year						
	2012	2013	2014	2015	2016	2017	2018
Leaving cohort 2012							
No entry (yes=1)	24.51	9.98	6.96	6.08	5.70	5.54	5.41
Entering apprenticeship (yes=1)	48.42	61.73	64.64	65.52	65.89	66.05	66.17
Entering Gymnasium (yes=1)	23.16	23.81	23.87	23.88	23.88	23.89	23.89
Entering specialised school (yes=1)	3.91	4.47	4.52	4.53	4.53	4.53	4.53
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Leaving cohort 2013							
No entry (yes=1)		23.19	8.58	5.64	4.78	4.43	4.27
Entering apprenticeship (yes=1)		49.28	62.66	65.51	66.37	66.71	66.87
Entering Gymnasium (yes=1)		23.30	23.94	23.99	24.00	24.00	24.00
Entering specialised school (yes=1)		4.23	4.83	4.86	4.86	4.86	4.86
Total		100.00	100.00	100.00	100.00	100.00	100.00
Leaving cohort 2014							
No entry (yes=1)			23.08	9.07	6.16	5.26	4.87
Entering apprenticeship (yes=1)			48.75	61.57	64.40	65.29	65.68
Entering Gymnasium (yes=1)			23.76	24.39	24.43	24.43	24.43
Entering specialised school (yes=1)			4.41	4.97	5.01	5.01	5.02
Total			100.00	100.00	100.00	100.00	100.00
Leaving cohort 2015							
No entry (yes=1)				22.70	8.60	5.65	4.77
Entering apprenticeship (yes=1)				48.81	61.69	64.52	65.39
Entering Gymnasium (yes=1)				23.85	24.44	24.51	24.51
Entering specialised school (yes=1)				4.64	5.28	5.33	5.33
Total				100.00	100.00	100.00	100.00
Leaving cohort 2016							
No entry (yes=1)					22.19	8.46	5.57
Entering apprenticeship (yes=1)					48.36	60.83	63.65
Entering Gymnasium (yes=1)					24.50	25.04	25.08
Entering specialised school (yes=1)					4.95	5.67	5.70
Total					100.00	100.00	100.00
Leaving cohort 2017							
No entry (yes=1)						21.63	8.03
Entering apprenticeship (yes=1)						48.76	61.07
Entering Gymnasium (yes=1)						24.58	25.11
Entering specialised school (yes=1)						5.02	5.79
Total						100.00	100.00

Notes: .

Table A.2: Descriptives – school entrants timing

	Share of students starting school		
	Early	On time	Late
School starting age	6.114	6.670	7.181
Male (yes=1)	0.076	0.753	0.171
Female (yes=1)	0.099	0.776	0.125
Foreign citizen (yes=1)	0.104	0.767	0.129
Swiss citizen (yes=1)	0.082	0.764	0.154
German-speaking (yes=1)	0.073	0.752	0.175
French-speaking (yes=1)	0.091	0.817	0.092
Italian-speaking (yes=1)	0.299	0.626	0.075
Romansh-speaking (yes=1)	0.048	0.923	0.029
Urban-area (yes=1)	0.098	0.772	0.130
Intermediate-area (yes=1)Female (yes=1)	0.080	0.759	0.161
Rural-area	0.059	0.745	0.196
School entry 2011 (yes=1)	0.075	0.771	0.154
School entry 2012 (yes=1)	0.084	0.779	0.137
School entry 2013 (yes=1)	0.098	0.766	0.136
School entry 2014 (yes=1)	0.096	0.777	0.127
School entry 2015 (yes=1)	0.093	0.765	0.142
School entry 2016 (yes=1)	0.093	0.747	0.160
School entry 2017 (yes=1)	0.071	0.748	0.181
Zurich	0.121	0.757	0.122
Bern	0.051	0.803	0.146
Luzern	0.065	0.576	0.359
Uri	0.011	0.538	0.451
Schwyz	0.026	0.704	0.270
Obwalden	0.054	0.803	0.144
Nidwalden	0.033	0.806	0.161
Glarus	0.040	0.704	0.256
Zug	0.173	0.756	0.071
Fribourg	0.021	0.750	0.230
Solothurn	0.024	0.876	0.100
Basel-Stadt	0.158	0.773	0.069
Basel-Landschaft	0.088	0.829	0.082
Schaffhausen	0.022	0.858	0.120
Appenzell-Ausserhoden	0.063	0.795	0.141
Appenzell-Innerhoden	0.036	0.728	0.237
St. Gallen	0.014	0.730	0.257
Graubunden	0.040	0.919	0.041
Aargau	0.086	0.777	0.137
Thurgau	0.033	0.653	0.315
Ticino	0.307	0.615	0.078
Vaud	0.179	0.710	0.111
Vallais	0.007	0.862	0.131
Neuchatel	0.027	0.916	0.057
Geneva	0.057	0.914	0.029
Jura	0.002	0.846	0.153
N	7,433	65,150	12,648

Notes: Sample restricted to students entering school in the years 2011-2017 and being born within 30 days of the cutoff date.
Source: LABB data.

Table A.3: RD estimates: birth cohorts 1998 and 1999

	Birth cohort: 1998				Birth cohort: 1999			
	Apprenticeship		Baccalaureate school		Apprenticeship		Baccalaureate school	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample mean	0.618	0.618	0.246	0.246	0.614	0.614	0.246	0.246
RD estimate	-0.061***	-0.059***	0.048***	0.046***	-0.065***	-0.062***	0.058***	0.054***
Standard error	(0.019)	(0.018)	(0.017)	(0.017)	(0.019)	(0.019)	(0.017)	(0.017)
Bias-corr. 95% conf. interval	[-0.099, -0.024]	[-0.095, -0.022]	[0.014, 0.081]	[0.013, 0.079]	[-0.103, -0.028]	[-0.099, -0.025]	[0.024, 0.092]	[0.020, 0.087]
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Regional characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Total observations	12247	12247	12247	12247	11863	11863	11863	11863
Observations used (LHS)	6213	6213	6213	6213	6021	6021	6021	6021
Observations used (RHS)	6034	6034	6034	6034	5842	5842	5842	5842

Notes: Each column displays a separate local linear RD regression (bandwith 30 days from cutoff). Personal characteristics include dummies for being female and being Swiss citizen. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students born in the year 1998 or 1999. Stars (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$) based on standard errors in parantheses. These do not account for estimation bias, the square brackets contain robust bias-corrected confidence intervals.

Table A.4: RD estimates for covariates

	Covariate			
	Female	German speak. region	Parents tertiary education	Swiss citizen
Sample mean	0.487	0.710	0.435	0.858
RD estimate: after cutoff	0.001	-0.002	0.006	0.010*
	(0.008)	(0.003)	(0.025)	(0.006)
Total observations	74007	74007	8305	74007
Observations used (LHS)	37325	37325	3976	37325
Observations used (RHS)	36682	36682	4329	36682

Notes: Each column displays a separate RD regression (bandwith 30 days from cutoff). Covariates from rd main models are here used as dependent variable: (1) being female (vs. male), (2) German language region (vs. French, Italian), (3) having parents with tertiary education (vs. upper-secondary or compulsory), (4) being Swiss citizen. Stars (*p¡.10, **p¡.05, ***p¡.01) based on standard errors in parantheses.

Table A.5: RD regression – type of apprenticeship chosen

	Apprenticeship (yes=1)				
	Length of apprenticeship			Vocational baccalaureate	
	2 years (yes= 1) (1)	3 years (yes= 1) (2)	4 years (yes= 1) (3)	(yes= 0) (4)	(yes= 1) (5)
Sample mean	0.036	0.408	0.171	0.516	0.098
RD estimate: after cutoff	-0.003	-0.024***	-0.009	-0.035***	0.000
	(0.003)	(0.008)	(0.006)	(0.008)	(0.005)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes
Total observations	74007	74007	74007	74007	74007
Observations used (LHS)	37325	37325	37325	37325	37325
Observations used (RHS)	36682	36682	36682	36682	36682

Notes: rdest.track.apptype. Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. students enter the upper secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Technical robustness RD regression – cutoff information

	No entry (yes=1)			Apprenticeship (yes=1)			Bacc. school (yes=1)			Spec. school (yes=1)		
	Donut (1)	Residence canton before 10 (2)	Canton of birth (3)	Donut (4)	Residence canton before 10 (5)	Canton of birth (6)	Donut (7)	Residence canton before 10 (8)	Canton of birth (9)	Donut (10)	Residence canton before 10 (11)	Canton of birth (12)
Sample mean	0.087	0.074	0.086	0.615	0.609	0.622	0.247	0.263	0.242	0.052	0.054	0.050
RD estimate: after cutoff	-0.007	-0.011*	-0.008	-0.027***	-0.035***	-0.024***	0.032***	0.040***	0.030***	0.003	0.006	0.003
	(0.005)	(0.006)	(0.006)	(0.008)	(0.010)	(0.009	(0.007)	(0.009)	(0.008)	(0.004)	(0.005)	(0.004)
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Leaving year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	71477	44327	51543	71477	44327	51543	71477	44327	51543	71477	44327	51543
Observations used (LHS)	36018	23213	25839	36018	23213	25839	36018	23213	25839	36018	23213	25839
Observations used (RHS)	35459	21114	25704	35459	21114	25704	35459	21114	25704	35459	21114	25704

Notes: Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). In the donut specificaitons we exclude students born within one day from the cutoff. In the second type of specifications we exclude students who's canton of residence is not recorded in the data before their tenth birthday. In the third type of specification we merge the cutoff information based on students' canton of birth and exclude students for whom we are missing canton of birth. Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. Students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: RD-robust estimates – program choice, data driven bandwidth selection

	No entry (yes=1)			Apprenticeship (yes=1)			Bacc. school (yes=1)			Spec. school (yes=1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample mean	0.087	0.088	0.088	0.615	0.615	0.615	0.246	0.245	0.245	0.052	0.052	0.052
RD estimate: after cutoff	-0.008*	-0.009*	-0.005	-0.024***	-0.036***	-0.057***	0.029***	0.038***	0.058***	0.002	0.003	0.002
	(0.004)	(0.005)	(0.007)	(0.006)	(0.008)	(0.013)	(0.005)	(0.007)	(0.012)	(0.003)	(0.004)	(0.005)
Polynomials (estimation)	1	2	4	1	2	4	1	2	4	1	2	4
Polynomials (bias)	2	3	5	2	3	5	2	3	5	2	3	5
Bandwidth (estimation)	41.9	63.0	80.7	45.6	53.2	66.7	59.8	57.6	69.5	42.7	66.1	87.9
Bandwidth (bias)	67.6	87.2	103.1	77.0	84.1	92.9	99.0	88.5	95.0	91.4	100.1	114.8
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Leaving year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	447949	447949	447949	447949	447949	447949	447949	447949	447949	447949	447949	447949
Observations used (LHS)	50908	77936	98732	55831	65560	81669	72853	70431	85409	52198	81669	107538
Observations used (RHS)	50308	77769	99197	55308	65442	81545	72814	70448	85324	51496	81545	107894

Notes: Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a triangular kernel to construct the local-polynomial estimator and one common MSE-optimal bandwidth selector for the RD treatment effect estimator). This yields bias-corrected estimates with a selected bandwidth such that the point estimator for the bias-corrected estimate is mean-square-error optimal (**??**). Personal characteristics include dummies for being female, being Swiss citizen, and being immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: RD regression – program choice, smaller bandwidths

	No entry (yes=1)		Apprenticeship (yes=1)		Bacc. school (yes=1)		Spec. school (yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample mean	0.089	0.086	0.612	0.616	0.246	0.246	0.054	0.052
RD estimate: after cutoff	-0.010	-0.004	-0.072***	-0.052***	0.081***	0.053***	0.001	0.003
	(0.014)	(0.009)	(0.023)	(0.014)	(0.021)	(0.013)	(0.011)	(0.007)
Bandwidth (days))	5	10	5	10	5	10	5	10
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Leaving year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	12440	24715	12440	24715	12440	24715	12440	24715
Observations used (LHS)	6243	12457	6243	12457	6243	12457	6243	12457
Observations used (RHS)	6197	12258	6197	12258	6197	12258	6197	12258

Notes: Each column displays a separate local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures using the rdrobust-command in Stata (applying a linear specification of the running variable to construct the point estimator, second order polynomials to construct the bias correction, and a triangular kernel to construct the local-polynomial estimator). Personal characteristics include dummies for being woman, Swiss citizen, and immigrated to Switzerland after birth, respectively. Regional characteristics include dummies for the area type (urban, intermediate, rural) and the language region (German, French, Italian) the student's lower secondary school is located at. Sample restricted to students leaving school in 2012-2017. Students that enter the secondary track more than one year after completion of compulsory schooling are considered as no entries. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. LHS: left-hand side, RHS: right-hand side.

Table A.9: Descriptives – instrument for mediator

	Same municipality shares entering the following lower secondary tracks		
	Basic track	No level distinction	Advanced track
Male (yes=1)	0.303	0.031	0.645
Female (yes=1)	0.302	0.032	0.646
Swiss citizen (yes=1)	0.302	0.030	0.648
Foreign citizen (yes=1)	0.308	0.039	0.631
In grade six in 2011	0.309	0.029	0.642
In grade six in 2012	0.297	0.034	0.650
Urban resident in grade six	0.303	0.037	0.638
Intermediate resident in grade six	0.303	0.012	0.666
Rural resident in grade six	0.302	0.036	0.643
German-speaking resident in grade six	0.321	0.040	0.619
French-speaking resident in grade six	0.252	0.008	0.717
Italian-speaking resident in grade six	0.308	0.013	0.671
Romansh-speaking resident in grade six	0.347	0.012	0.628
Zurich	0.338	0.012	0.630
Bern	0.355	0.017	0.605
Luzern	0.256	0.060	0.673
Uri	0.304	0.011	0.668
Schwyz	0.292	0.002	0.681
Nidwalden	0.091	0.535	0.345
Glarus	0.421	0.003	0.566
Zug	0.251	0.002	0.727
Fribourg	0.198	0.001	0.789
Solothurn	0.344	0.013	0.626
Basel-Stadt	0.007	0.510	0.462
Schaffhausen	0.371	0.010	0.582
Appenzell-Ausserhoden	0.351	0.006	0.629
Appenzell-Innerhoden	0.321	0.000	0.679
Graubunden	0.320	0.004	0.665
Aargau	0.343	0.015	0.607
Thurgau	0.332	0.081	0.572
Vaud	0.227	0.014	0.737
Neuchatel	0.299	0.009	0.670
Geneva	0.328	0.002	0.637
Jura	0.209	0.005	0.771
N (students)	5,024	411	12,530

Notes: Sample restricted to students being in grade six in 2011-2012, leaving school 2017 or earlier, and being born within 30 days of the cutoff date. The correlation within muncipality over the years is 0.956.
Source: LABB data.

Table A.10: Transition probabilities between lower and upper secondary programs

	Row total					
	No entry	Apprenticeship	Bacc. school	Specialized school	Total	N
Special needs	57.25	42.44	0.18	0.13	100.00	17,173
Basic requirements	14.00	83.48	0.23	2.28	100.00	121,316
No level distinction	10.17	71.69	9.15	9.00	100.00	19,806
Advanced requirements	3.60	52.95	37.03	6.41	100.00	289,108
Total	8.77	61.66	24.40	5.16	100.00	–
N	39,252	275,855	109,188	23,108	–	447,403

	Column total					
	No entry	Apprenticeship	Bacc. school	Specialized school	Total	
Special needs	25.05	2.64	0.03	0.10	3.84	17,173
Basic requirements	43.28	36.71	0.26	11.99	27.12	121,316
No level distinction	5.13	5.15	1.66	7.71	4.43	19,806
Advanced requirements	26.55	55.50	98.05	80.20	64.62	289,108
Total	100.00	100.00	100.00	100.00	100.00	–
N	39,252	275,855	109,188	23,108	–	447,403

Notes: All students leaving compulsory schooling 2012-2017. The orginal names in the LABB data are: Separativer Unterricht allgemein, Sonderschule, Sonderklasse, ausländische Programme (all "Special needs"), Sek. I Grundansprüche ("Basic requirements"), Sek. I ohne Niveauunterscheidung ("No level distinction"), Sek. I erweiterte Ansprüche ("Advanced requirements").
Source: LABB data.

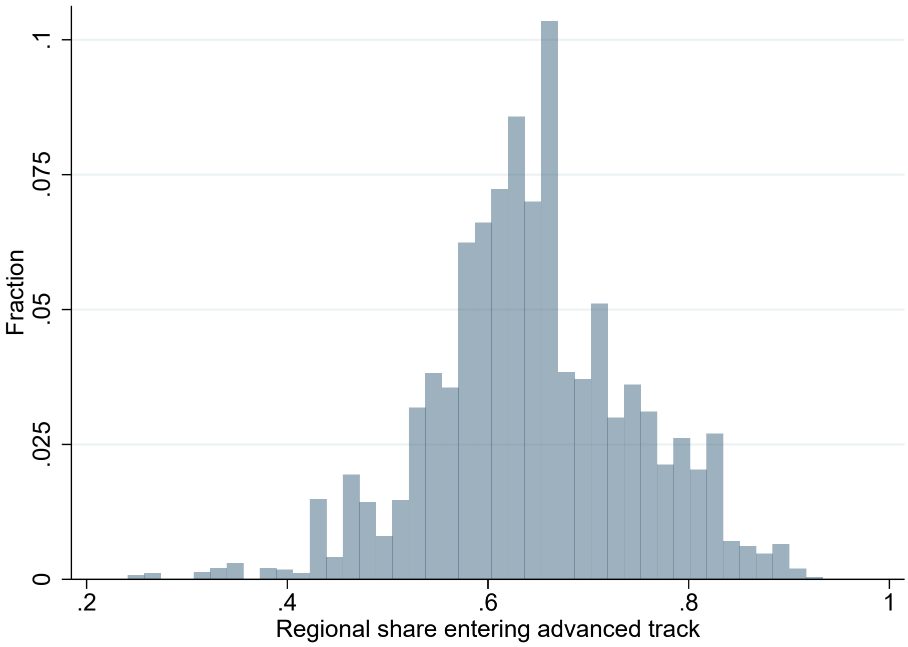
Table A.11: Transition probabilities btw. lower and upper secondary programs (w/o long bacc. school)

	Row total					
	No entry	Apprenticeship	Bacc. school	Specialized school	Total	N
Special needs	59.52	40.11	0.22	0.15	100.00	11,607
Basic requirements	15.89	80.06	0.37	3.68	100.00	74,390
No level distinction	10.59	68.19	10.59	10.62	100.00	15,972
Advanced requirements	3.70	48.43	40.01	7.87	100.00	192,023
Total	9.36	57.17	26.81	6.66	100.00	–
N	27,520	168,089	78,813	19,570	–	293,992

	Column total					
	No entry	Apprenticeship	Bacc. school	Specialized school	Total	
Special needs	25.11	2.77	0.03	0.09	3.95	11,607
Basic requirements	42.96	35.43	0.35	14.00	25.30	74,390
No level distinction	6.15	6.48	2.15	8.67	5.43	15,972
Advanced requirements	25.79	55.32	97.47	77.24	65.32	192,023
Total	100.00	100.00	100.00	100.00	100.00	–
N	27,520	168,089	78,813	19,570	–	293,992

Notes: All students leaving compulsory schooling 2012-2017. The orginal names in the LABB data are: Separativer Unterricht allgemein, Sonderschule, Sonderklasse, ausländische Programme (all "Special needs"), Sek. I Grundansprüche ("Basic requirements"), Sek. I ohne Niveauunterscheidung ("No level distinction"), Sek. I erweiterte Ansprüche ("Advanced requirements").
Source: LABB data.

Figure A.1: Proportion of students entering the advanced lower secondary track per municipality (Instrument for mediator)



Notes: The figure shows the distribution of the proportions of students entering the advanced track on the lower secondary level, calculatedd for every student in the sample for his or her muncipality (where student i attended primary school). These values are used to instrument the mediator (students i lower secondary schooling track: basic or advanced). Sample restricted to students leaving compulsory schooling 2015-2017.
Source: LABB data.