Summer-born struggle: the effect of school starting age on health, education, and work*

Simone Balestra, Beatrix Eugster, and Helge Liebert

University of St. Gallen, Switzerland

Abstract

This paper offers a comprehensive analysis of the impact of school starting age (SSA) from childhood through the labor market. We first study the effect of SSA on a child's probability of developing special educational needs in early grades. Children with a higher SSA are less likely to develop behavioral problems and speech impediments, whereas learning disabilities, ADHD, and dyslexia/dyscalculia remain unaffected. The SSA-effect persists throughout compulsory schooling, resulting in higher test scores in grade eight and better-quality vocational training contracts. However, we find no effect on earnings and employment.

Keywords: school starting age, special needs, school performance.

JEL Classification: I14, I21, J13.

* We are grateful to Christina Felfe, Michael Lechner, Hannes Schwandt, Anthony Strittmatter, and Stefan Wolter for their constructive comments. For the data, we thank the School Psychological Service, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office. The paper benefited from valuable feedback at the following meetings: European Association of Labor Economists, Ski and Labor seminar, Swiss Society of Economics and Statistics, and departmental seminars at the universities of St. Gallen, Wuppertal, and Zürich. We acknowledge financing from the Swiss National Science Foundation (grant no. 100018_176381). Corresponding author: Simone Balestra, University of St. Gallen, Rosenbergstrasse 51, CH-9000 St. Gallen, simone.balestra@unisg.ch. This paper reflects the views of the authors alone. The usual disclaimer applies.

1 Introduction

Virtually all education systems have a single cutoff date that determines when children become eligible for compulsory schooling. This cutoff rule creates a continuum of ages at school entry, whereby the oldest child is up to one year older than his or her youngest classmates. Research has shown that age of school entry is an important determinant of early student achievement (Black, Devereux, and Salvanes, 2011; Elder and Lubotsky, 2009; Fredriksson and Öckert, 2014). Although this pattern is consistent across countries (Bedard and Dhuey, 2006), the underlying mechanism that supports this empirical regularity and the long-term effects of school starting age remain unclear.

One prominent explanation from developmental psychology is maturity (Whitebread, 2012). While children are ready to learn at all ages, young children are usually less prepared to engage in academic work than their older peers (Morrison, Alberts, and Griffith, 1997; Stipek and Ryan, 1997) and more vulnerable to external influences (Datar and Gottfried, 2015). This developmental disadvantage might trigger special educational needs in children, e.g., due to increased incidence of learning impairments or behavioral problems. Following ICD-10 diagnosis guidelines, we define special needs as an umbrella term for special educational requirements resulting from learning disabilities, communication disorders, emotional and behavioral disorders, physical disabilities, and developmental disabilities.

Individuals who develop special needs during childhood have a higher risk of subsequent history of unsuccessful education, difficulties in labor market integration, and lower earnings during adulthood (Hanushek, Kain, and Rivkin, 2002; Wagner and Blackorby, 1996). In addition, educating children with special needs is considerably more costly than educating children without special needs (Duncombe and Yinger, 2005). It is thus imperative for policy makers to understand the role of school starting age (SSA) in explaining child development, with respect to not only achievement but also to special needs demand.

In this paper, we study the causal effect of SSA on a child's probability of developing special educational needs in early grades. We then assess the persistence of SSA effects in terms of scholastic achievement at the end of compulsory schooling and on labor market entry. We are able to credibly identify the effects of SSA through a regression discontinuity design based on the exact day of birth. In Switzerland, children are supposed to enter compulsory school[1] in the fall if they have reached age four before August 1 of the same year. This institutional rule allows us to compare children born around the cutoff, children who are observationally similar but who enter school at different ages.

We base our analyses on the complete school cohorts in the Swiss canton of St. Gallen born from 1992 to 2003.[2] For these children we have administrative information from the School Psychological Service (SPS). The SPS is a canton-wide, centralized service that provides children and their families with counseling and diagnosis for school-related problems. For each child, we observe whether the child has ever been to the SPS, when the SPS first registered the child, when and how often the child visited the SPS, and whether the SPS dismissed the child. Furthermore, for those children who were sent to the SPS, we have detailed information on the reason for registration, school psychologists assessment, and diagnoses.

For all children we observe test scores in Math and German in grade eight, when students perform a compulsory standardized exam. We further merge the data with administrative records on vocational education and training (VET) and post-compulsory secondary education required for higher education. In Switzerland, roughly 65% of each cohort chooses the VET track after compulsory schooling, directly entering the labor market. About 20% of students continue on the academic track and enter higher education. Finally, we link our data to administrative records on earnings and employment from the Central Compensation

[1] In Switzerland, compulsory school consists of two years of kindergarten, six years of primary school, and three years of secondary school.

[2] Switzerland is a federal republic comprising 26 member states called cantons. St. Gallen is the fifth largest canton in Switzerland with a population of about 500,000.

Office, the social security administration in Switzerland.

Prior research shows some evidence that increases in SSA reduce children's risk of disability classification (Dhuey and Lipscomb, 2010; Elder, 2010) and improve measures of mental health (Dee and Sievertsen, 2018; Mühlenweg et al., 2012). We contribute to this literature along three dimensions. First, we present novel evidence on the relationship between SSA and special needs incidence. With respect to reliability and wealth of information, our data constitute a major improvement on existing data sets, which are primarily retrospective non-expert surveys.

Second, we shed light on the interplay between experts' evaluation and educators' behavior towards special education classification. Dhuey and Lipscomb (2010) and Schwandt and Wuppermann (2016) show that, in some cases, educators use special needs classification as a supplemental service that targets additional resources at younger students. In the presence of this over-referring, experts' evaluations of special needs become crucial. When specialists – as compared to teachers or parents – perform the diagnoses, the risk of applying relative standards is much smaller (Dalsgaard et al., 2012). We deal with this issue by distinguishing between the decision to refer a child for special needs evaluation (made by non-experts such as teachers or parents) and the results of an expert evaluation, which is made by SPS professionals.

Third, the data allow us to perform a comprehensive assessment of the medium- and long-run effects of SSA within a uniform institutional framework. The current literature on the long-term effects of SSA is mixed, finding small or zero effects on employment and earnings (Dustmann, Puhani, and Schönberg, 2017; Fredriksson and Öckert, 2014). In the medium run, we examine the effect of SSA on grade retention and test scores at the end of compulsory education. Moving further towards long-run persistence, we estimate the effects on post-compulsory education and we pose the question of whether SSA effects still matter at the age of labor market entry and investigate employment and earnings in the early years after compulsory schooling.

The results indicate that children entering school at a younger age have a higher risk of developing special needs than children starting school one year later. Being born shortly before the cutoff date increases the probability of developing special needs by five percentage points or about 14%. Importantly, this effect is entirely driven by special needs developed after kindergarten, not due to pre-existing health conditions. We also show that the effect does not simply stem from younger children's being over-referred for special needs evaluation. By distinguishing by type of special needs, we find that entering school at a younger age mostly increases behavioral problems and speech impediments. The incidence of learning disabilities, dyslexia/dyscalculia, and ADHD (Attention Deficit Hyperactivity Disorder) are unaffected by SSA. At the intensive margin, we find that younger children on average have nearly one more consultation with the SPS – a measure of SN severity – than older children.

At the end of compulsory education, differences in SSA still affect students' school outcomes. We find that younger students perform worse on standardized tests. While no effects are found for grade repetition during primary and secondary school, we find that younger students are more likely to be allocated to a bridge year between kindergarten and primary school. These results suggest that many children starting school at a younger age will eventually prolong their compulsory education by one year.

The differences at the end of compulsory schooling disappear in the long run, when adolescents choose their post-compulsory education track. Children starting school younger are as likely to find an apprenticeship position as children entering school later. However, Switzerland offers two quality-types of apprenticeships, characterized most notably by their different durations (2 vs. 3 or 4 years) and curricula contents. We find that children starting school younger are significantly less likely to enter the high-quality vocational track. On the other hand, children with lower SSA are equally likely to enter the academic track as their peers with higher SSA. For early labor market outcomes, we find no effect of SSA on earnings and employment for the first seven years after compulsory schooling.

While the reduced form estimates are informative, it is important to acknowledge that the RDD design is fuzzy because of individual decisions to delay school start. We estimate local average treatment effects by applying two-sample instrumental variable estimation, because we do not observe actual school starting age for the children in our main data. To construct the first stage, we use the official education statistics for the cohorts born between 2006 and 2011. Results suggest that redshirting is a common practice in Switzerland, but only for children born in June or July, i.e., shortly before the cutoff. While almost 100% of children born between August and May start school on time, only 63% of children born in the week before the cutoff actually start school in the regular year. As expected, the two-sample IV estimations are somewhat larger in size and less precisely estimated than our reduced-form estimates, but overall they confirm our previous results.

In sum, we find that a lower SSA triggers the development of special needs upon school entry and that this age effect persists until the end of compulsory education. However, we find no significant medium- and long-term effects in terms of post-compulsory education choices and labor market outcomes. Our results thus suggest that the negative effects of going to school at a younger age are limited within the domain of compulsory education.

The remainder of this study proceeds as follows. Section 2 gives an overview of the data and the institutional background. Section 3 introduces and explains our empirical strategy. Section 4 presents and discusses the results, along with the sensitivity analyses in section 5. Section 6 relates our findings to the existing literature and section 7 concludes.

2 Data and institutional background

In St. Gallen, children enter compulsory schooling in the fall if they have reached age four before August 1 of the same year. The typical school curriculum consists of two years of kindergarten, six years of primary school, and three years of secondary school. Ability tracking occurs after primary school, with children entering either a higher-ranked (*Sekundarschule*) or a lower-ranked (*Realschule*) track of secondary school. After finishing secondary school, children typically enter either an academic preparation track (i.e., high school) or vocational education.

Mainstreaming is a common practice in Swiss schools.[3] However, about 2% of children – those who have severe physical or mental handicaps that would not allow them to follow a regular curriculum – are educated in special education institutions. We do not have any information about these children in our data. The regular public schools, however, offer institutionalized support services for children who develop special educational needs. Special educational needs result from diagnoses associated either with behavioral problems or learning impairments (e.g., ADHD, dyslexia/dyscalculia or speech impediments). In such cases, the teacher notifies both the School Psychological Service and the parents.[4] The SPS then initiates contact with the family and schedules an assessment meeting. During this meeting, the SPS staff performs a diagnostic evaluation, provides a diagnosis, and recommends therapy if necessary. After the assessment session, the SPS keeps track of the child's progress and is the liaison between involved parties.

Our analysis focuses on the student population of the canton of St. Gallen. We combine different data sources for all school cohorts born from 1992 to 2003. These sources include data on special needs from the SPS, data on achievement test scores from "Stellwerk8", data

[3] In Switzerland, 95% of schools are public. We exclude the four private schools in St. Gallen from our sample.

[4] The SPS is organized at the cantonal level and operates through its eight regional offices, one for each school district.

on post-compulsory education, and register data on employment and earnings. We observe for the entire population of students the exact date of birth, gender, and an indicator of whether they are native German speakers.

For every child who has ever been in contact with the SPS, we observe the age at and reason for registration (e.g., learning difficulties, disruptive behavior, family problems), the number of consultations, and comprehensive information on the diagnoses and the suggested treatments. This data set allows us to construct measures of the onset and severity of special needs. Furthermore, we can distinguish between non-expert and expert assessment by comparing the teacher- or parents-initiated registration at the SPS with the SPS staff assessment and, if appropriate, a further diagnosis.

The achievement data are based on a compulsory standardized test taken in eighth grade, called "Stellwerk8." Stellwerk8 is a norm-referenced, self-scoring, adaptive, computer-based exam similar in spirit to the Graduate Record Examination. All students in grade eight – except those enrolled in special education schools – are tested. The test is administered between February and April, towards the end of the school year. The test results are important for students. After the test, students receive a certificate with their Stellwerk8 results. This certificate is usually provided to potential employers when students apply for apprenticeship positions during ninth grade (the last year of compulsory education).

In addition to the data on school performance, we add data about students' career path after compulsory education. Upon finishing compulsory schooling, most Swiss students typically enter VET by applying for VET positions and signing a training contract with a firm. VET combines part-time formal education with training and experience at the workplace.[5] We link our data to information about all VET contracts signed in the canton of St. Gallen

[5]Students who attend a VET program study part-time at school for 1 to 1.5 weekdays. For the remaining time (3.5 to 4 weekdays), students work as apprentices in host companies with whom they have an employment contract for their entire three- to four-year training period. See Oswald and Backes-Gellner (2014) for an overview of the Swiss VET system. For a broader perspective on VET across countries, see Wolter and Ryan (2011).

in 2008-2016. A smaller percentage of each cohort enters the academic preparation track to obtain a higher education entrance qualification. We link the children in our data to the administrative high school records to track those who enter the academic track after compulsory schooling. For this analysis, we lose the last cohort of the data because of poor match overlap in the VET and high school data, due to children not yet having left school.

The last piece of information we add to our data set is provided by the Swiss Federal Statistical Office and comes from the Swiss Central Compensation Office (CCO). The CCO is the central body for first-pillar social security in Switzerland, which covers old-age and survivors' insurance, disability insurance, and compensation for loss of earnings. To qualify for these benefits, workers pay social security taxes on their earnings, and the claimant's benefits are based on the wage earner's contributions. The CCO thus possesses monthly employment and earnings data for all registered individuals. For individuals in our data who ever worked between 2007 and 2016, we observe how many months per year they worked and their respective (net) earnings.

Table 1 shows the descriptive statistics of our sample. About 33% of children are referred to special needs services at some point during their school career. However, 6% are dismissed without a diagnosis after the initial screening and do not receive further support. The remaining 27% are diagnosed primarily with learning impairments (19%) and behavioral problems (7%). While special educational needs comprise a variety of learning disabilities, behavioral problems consist of disruptive behavior almost exclusively. These proportions match the incidence reported in aggregate statistics at the federal level and are also in line with figures reported from other OECD countries (OECD, 2008). For about 70% of children with special needs the first onset occurs within three years after entering school, and for more than 90% by the end of primary school.

While we have information on school performance for nearly all children (99.2%, panel

Table 1: Descriptive statistics

	(1) Mean	(2) SD	(3) Min	(4) Max	(5) N
A1. Special needs: Incidence, onset, severity					
Special needs (SN)	0.337	0.473	0.00	1.00	51,138
SN: onset before primary school	0.067	0.249	0.00	1.00	51,138
SN: onset during primary school	0.271	0.444	0.00	1.00	51,138
SN: dismissed after initial examination	0.058	0.235	0.00	1.00	51,138
SN: positive diagnosis	0.279	0.448	0.00	1.00	51,138
Consultations	2.968	6.489	0.00	204	51,138
A2. Special needs: Type					
Learning impairment	0.193	0.395	0.00	1.00	51,138
Behavioral problem	0.067	0.250	0.00	1.00	51,138
Speech impediment	0.118	0.322	0.00	1.00	51,138
Dyslexia/dyscalculia	0.107	0.309	0.00	1.00	51,138
ADHD	0.030	0.171	0.00	1.00	51,138
B. School performance, 8th grade					
Test score: Composite (standardized)	0.000	1.000	-4.27	4.08	50,717
Test score: Math (standardized)	0.000	1.000	-4.24	3.70	50,717
Test score: German (standardized)	0.000	1.000	-4.20	3.88	50,717
Bridge year between KG and PS	0.050	0.219	0.00	1.00	50,717
Grade repetition by 8th grade	0.017	0.131	0.00	1.00	50,717
C. Later outcomes					
Vocational education	0.645	0.478	0.00	1.00	41,642
Academic preparation track	0.178	0.382	0.00	1.00	41,642
Employment (months per year)	10.62	2.280	1.00	12.0	28,152
Monthly wage	1,872	964.0	300	4,825	28,152
D. Covariates					
Female	0.484	0.500	0.00	1.00	51,138
Non-native speaker	0.158	0.364	0.00	1.00	51,138
Born after July 31	0.410	0.492	0.00	1.00	51,138
Age at SPS registration	8.918	2.383	3.18	16.4	17,241

Notes: Descriptive statistics for the main estimation sample. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

b of Table 1), we have information on students' post-compulsory education choice for approximately 80% of the sample. For labor market outcomes, we have employment and earnings data for 55% of the sample.[6] The higher attrition rates for later outcomes are primarily due to many students not having finished yet compulsory school. For these students we observe neither their post-compulsory education choice nor their labor market outcomes.

The average worker in our data earns 1,900 Swiss Francs[7] per month and is employed for approximately ten and a half months per year. The earnings we observe are lower than the Swiss average, because we observe individuals in their early career. The wages reported in the present study are earned when the workers are between 17 and 24. Most of the earnings we observe are apprenticeship wages, not only because 65% of each Swiss cohort follows the vocational track but also because very few university graduates enter the labor market before 24.

3 Empirical strategy

The school starting cutoff date causes some children to be older than others when entering compulsory education. We adopt a regression discontinuity (RD) design around the August 1 cutoff to study the effect of school starting age on the development of special needs in early childhood.

Let X_i denote the forcing variable and $\bar{x}$ the known cutoff, normalized to $\bar{x} = 0$. In our case, X_i represents the exact date of birth and $\bar{x}$ is the cutoff date August 1. The cutoff date determines whether child i enters compulsory school in the relevant year ($X_i < 0$) or one year later ($X_i \geq 0$). Finally, let $Y_i(1)$ and $Y_i(0)$ denote the potential outcomes of the

[6]Missing post-compulsory data arises for a variety of reasons. First, some children are too young. Note that the CCO has a time-lag for data provision of around two years. Second, because individuals are mobile across cantons. We observe VET contracts and tertiary education only within the canton of St. Gallen. Third, some students do not enroll in further education and have no earnings, so they do not appear in any register.

[7]In 2018, 1 Swiss Franc buys 1.01 USD.

children entering compulsory school in the relevant year or not, respectively. The parameter of interest is the local average treatment effect at the cutoff τ_{RD}:

$$\tau_{\text{RD}} = \text{E}[\, Y_i(1) - Y_i(0) \mid X_i = \bar{x} \,] \,. \tag{1}$$

Under a testable continuity assumption, Hahn, Todd, and Van der Klaauw (2001) show that τ_{RD} is nonparametrically identifiable as the difference of two conditional expectations evaluated at the cutoff $\bar{x} = 0$:

$$\begin{aligned} \tau_{\text{RD}} &= \mu^{+} - \mu^{-} = \lim_{x \to 0^{+}} \mu(x) - \lim_{x \to 0^{-}} \mu(x) \,, \\ \mu(x) &= \text{E}[\, Y_i \mid X_i = x \,] \end{aligned} \tag{2}$$

The objective is to estimate a flexible approximation near the cutoff of the regression functions $\mu^{-}(x) = \text{E}[\, Y_i(0) \mid X_i = \bar{x} \,]$ and $\mu^{+}(x) = \text{E}[\, Y_i(1) \mid X_i = \bar{x} \,]$. To do so, we approximate the regression functions above and below the cutoff by means of local linear regressions, with weights computed by applying a kernel function on the distance of each observation's score to the cutoff. This nonparamentric local polynomial approach has become the standard choice for estimation of RD treatment effects (Gelman and Imbens, 2018).

In implementing the RD approach, we need to choose the kernel function for weighting the observations and the bandwidth for determining the sample size around the cutoff. The choice of the kernel function makes little difference in practice. For our main specifications, we rely on a triangular kernel, which is better suited for estimating a function at boundary points than the epanechnikov kernel (Cheng, Fan, and Marron, 1997).

For the choice of bandwidth, we follow a recent approach developed in Calonico, Cattaneo, and Titiunik (2014) and Calonico et al. (2018). They show that commonly used bandwidth selectors tend to yield bandwidths that are too large to ensure the validity of the underlying distributional approximations, potentially leading to non-negligible bias. They propose an alternative method, with the RD point estimate corrected by an estimated bias term; the standard error estimates are then adjusted for the additional variability resulting from the estimation of the bias correction term. Throughout the paper, we present the bias-corrected estimates, and we select the bandwidth such that the point estimator for the bias-corrected estimate is mean square error optimal (see Calonico et al., 2018). We also test the sensitivity of our results to different bandwidth choices.

Although predetermined individual characteristics are not required for identification, their inclusion may improve precision. Identification is valid if the conditional expectation functions of the covariates are continuous at the cutoff (Calonico et al., 2018). In some regressions, we include gender, an indicator for non-native speaker, birth-year cohort fixed effects, postal code, and test year fixed effects. However, including covariates does not change the results qualitatively, and in our main specification we only include gender, non-native speaker, and birth-year cohort fixed effects.

As outlined above, identification in the RD design relies on the idea of local randomization around the threshold. We present some evidence for the validity of the identifying assumption. One main concern is that individuals manipulate the running variable by systematically timing birth in consideration of the school starting threshold. Manipulation typically leads to asymmetric sample selection and sorting on either side of the cutoff, which is often indicated by bunching in the distribution on one side of the assignment threshold. Figure A1 shows the distribution of date of birth in our sample. There is no visible change in births before August 1. More formally, using the test outlined in McCrary (2008), we cannot reject the null hypothesis of a zero discontinuity at the threshold (p-value = 0.273).

In addition, if observations are locally randomized at the threshold, any pre-determined characteristics should be balanced at the threshold. Panels (a) and (b) in Figure A2 show discontinuity graphs for the probability of being female and the probability of being a non-native speaker, respectively. Both characteristics are balanced at the threshold (p-values for gender and non-native speaker are 0.987 and 0.857, respectively). To alleviate residual concerns about the covariate balance, the main regression results control for gender, non-native speaker, and birth cohort fixed effects. The evidence presented in Figures A1 and A2 supports the internal validity of the RD design.

We complement our reduced form estimates with a two-sample instrumental variable strategy (TSIV). This approach is a fuzzy RD design, with the difference that first stage and reduced form are calculated from two separate samples. We base our approach on theoretical work by Angrist and Krueger (1992) and Inoue and Solon (2010), along with the applied studies conducted by Dee and Evans (2003), Devereux and Hart (2010), and Van den Berg, Pinger, and Schoch (2016).

Because our model is exactly identified, it is straightforward to show that dividing the reduced-form parameter τ_{RD} by the first stage is equivalent to the TSIV estimator. Let γ_{RD} denote the first stage effect, i.e., the effect of date of birth on the actual SSA. The following expression provides an estimate for the effect of SSA on the outcome of interest Y_i:

$$\hat{\beta}_{\text{RD}} = \hat{\tau}_{\text{RD}} / \hat{\gamma}_{\text{RD}} \tag{3}$$

Standard errors for $\hat{\beta}_{\text{RD}}$ are computed using bootstrap, under the assumption of zero covariance between the first-stage and reduced-form estimates. This assumption is plausible in our application, because reduced-form and first-stage estimates stem from different samples (different cohorts).

4 Results

4.1 Graphical evidence

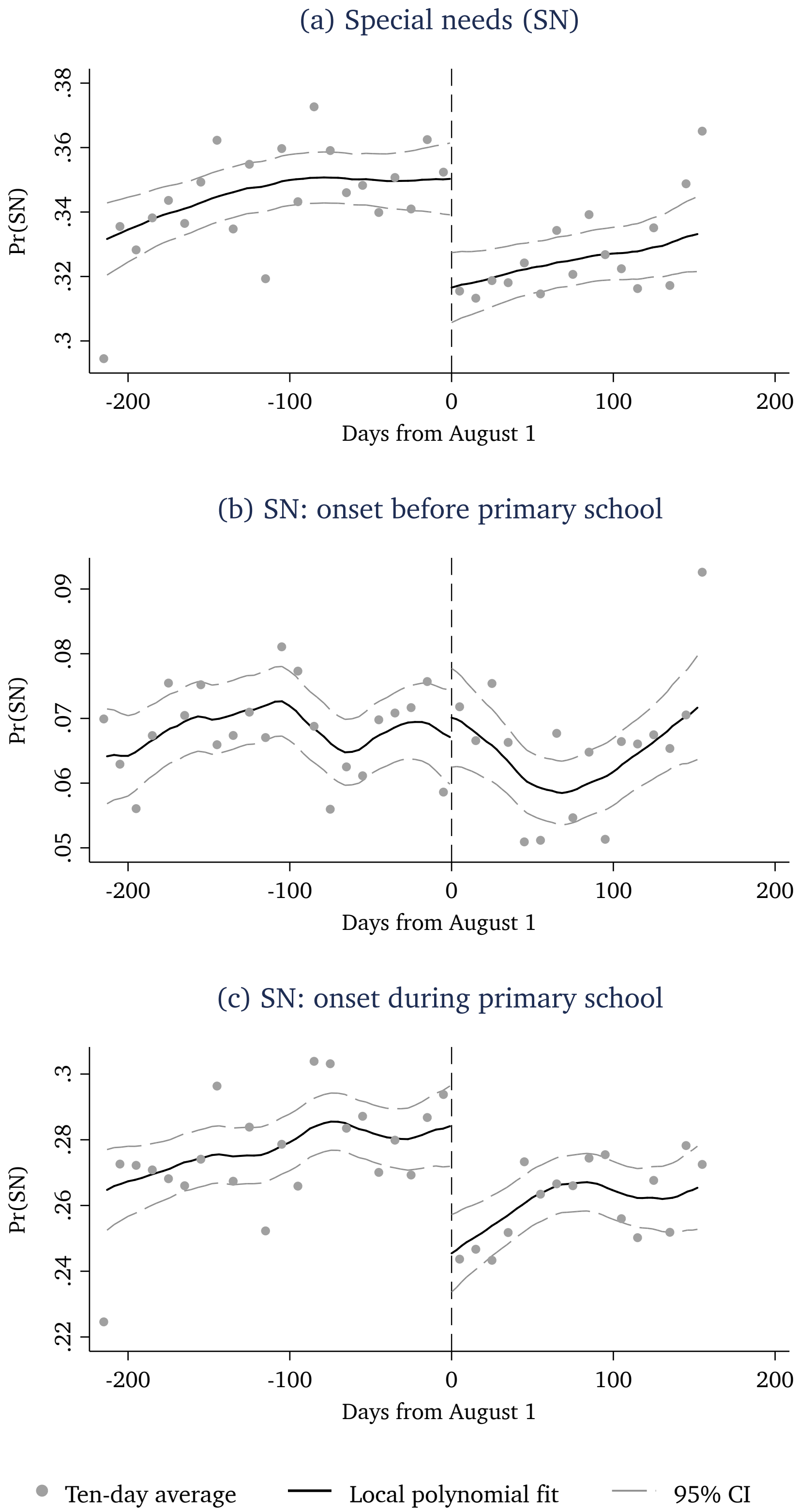
In this section, we present graphical evidence on the incidence of special educational needs among children starting school at relatively younger ages. Panel (a) in Figure 1 shows the probability of developing special needs by date of birth, relative to the school starting threshold of August 1. A child is classified as having special needs if he or she came into contact with the SPS before or during school time. On average, one third of all children are in contact with the SPS at least once. The graph in panel (a) shows that this fraction is about 31% for children born on or after August 1. In contrast, children born on July 31 or earlier have a four percentage points higher probability of becoming in contact with the SPS.

This result could be driven by several mechanisms. First, as in many other school systems, the possibility of redshirting a child (i.e., delaying school entry by one year) exists in Switzerland. Typically, because redshirting occurs before the child enters kindergarten, parents requesting redshirting would send their child to the SPS for examination. Therefore, the higher proportion of special needs among children born at or before July 31 could be driven by parents wishing to delay the school entry of their children and not by the occurrence of special needs. Note, however, that parents do not always have to send their child for examination to the SPS for redshirting but the local school authorities, usually together with a pediatrician, can also approve the request.[8] Panel (c) of Figure 1 shows no discontinuity in the incindece of special needs occuring before or during kindergarten. This result suggests that redshirting is not a main driver of special needs incidence on the left side of the cutoff.

Another explanation is that children develop special needs when subjected to the pressure of grading and peer performance in primary school. Panel (b) of Figure 1 shows the incidence of special needs occurring after entering primary school. Interestingly, the main

[8]Note that only about 6% of cases at the SPS were referred by their parents.

Figure 1: Special needs, incidence and onset



Notes: Discontinuity graphs plotting ten-day averages and cutoff-dependent local linear regression estimates of school starting age on measures of special needs incidence, onset, and severity. Local linear regression results are based on triangular kernel weights and a bandwidth choice following Calonico et al. (2018). Confidence limits (95%) are plotted in gray. Data are from the Stellwerk-Test service provider and the School Psychological Service St. Gallen.

effect is entirely driven by onset of special needs upon primary school entry.

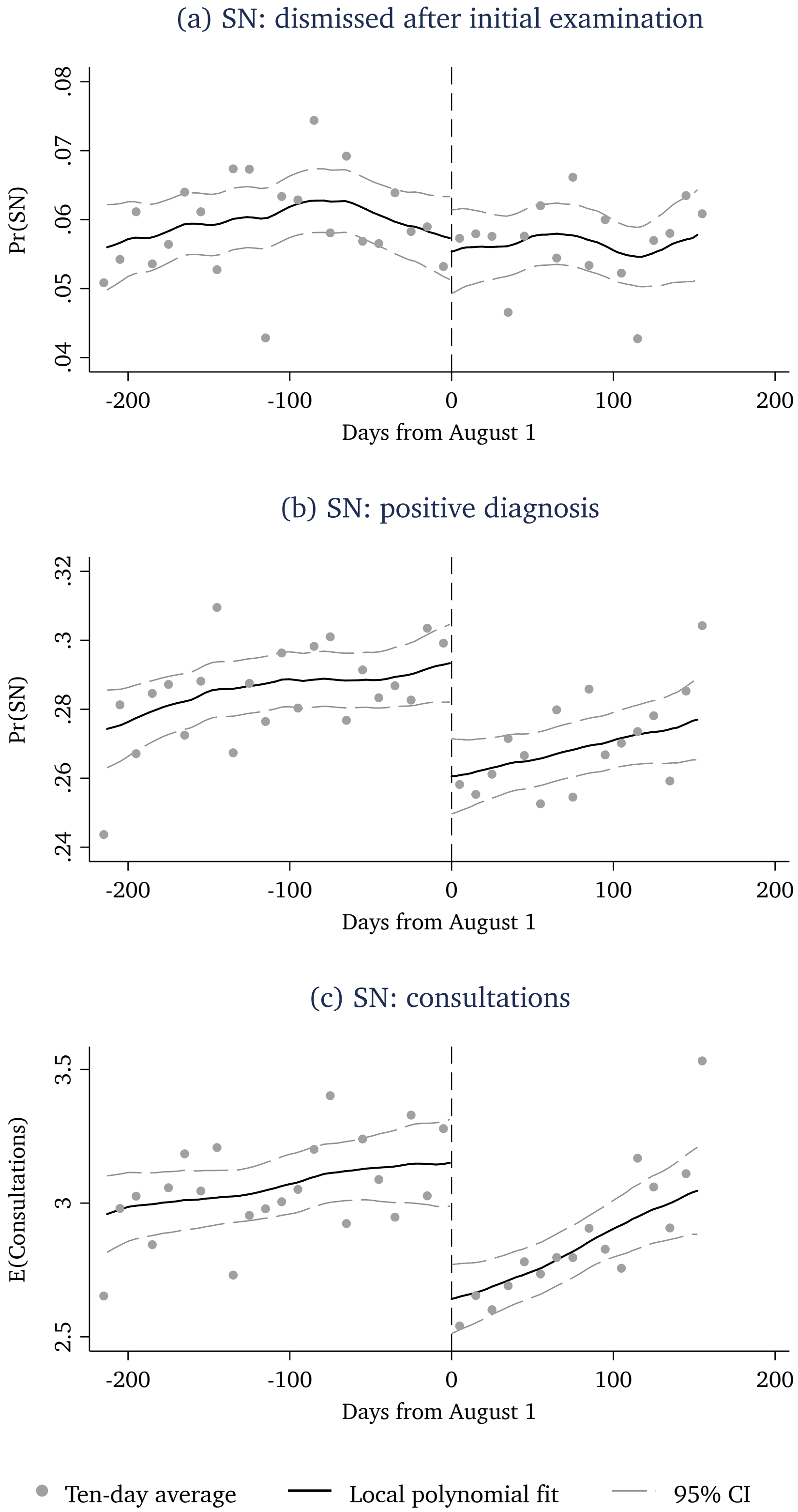
In Figure 2, we investigate whether the effect is driven by purely superficial increases in screening or by actual changes in true special needs diagnoses. This mechanism would be possible if teachers and parents were inclined to refer children for special needs evaluation because of their younger age rather than because of any actual need for special education. We distinguish between cases which are initially registered with the SPS but dismissed after initial examination by SPS professionals (panel a) and cases with positive diagnoses requiring further treatment (panel b). While there may be a slightly higher – but not statistically significant – rate of registered cases without a diagnosis at a younger age, we find that most of the effect is driven by cases diagnosed positively and receiving further treatment. In addition, we find a large discontinuity in the number of consultations received (panel c). Children born just before August 1 and entering school early require about 0.6 consultations more than children from the same cohort entering school one year later.

4.2 Special needs incidence, onset, and type

In this section we present the reduced form estimation results of the effect of school starting age on the development of special needs. Table 2 shows the estimated threshold effects using local linear regressions. We control in all regressions for gender, non-native speaker, and year-of-birth fixed effects. Observations are weighted using a triangular kernel, and the bandwidth is symmetric around the threshold and chosen by minimizing the regression discontinuity mean squared error.

Panel (a) in Table 2 presents the point estimates corresponding to the graphs in the previous section. Being born on August 1 instead of July 31 decreases the probability of developing special needs by 5.1 percentage points, a finding corresponding approximately to a 14% reduction. This effect is entirely driven by special needs developed after entering

Figure 2: Special needs, dismissals and consultations



Notes: Discontinuity graphs plotting ten-day averages and cutoff-dependent local linear regression estimates of school starting age on measures of special needs incidence, onset, and severity. Local linear regression results are based on triangular kernel weights and a bandwidth choice following Calonico et al. (2018). Confidence limits (95%) are plotted in gray. Data are from the Stellwerk-Test service provider and the School Psychological Service St. Gallen.

primary school, when children are in a more competitive school environment. We also find that the effect is not simply driven by younger children being over-referred for screening. Instead, we find that younger children are also diagnosed positively and receive treatment at higher rates, leading to 0.7 more consultations on average.

Panel (b) shows the results for common types of special needs, to understand what kind of special need is driving the overall results. We find that entering school at an older age decreases significantly behavioral problems by 3.5 percentage points and speech impediments by 3.4 percentage points. The other common types of diagnosis, such as learning impairments, dyslexia/dyscalculia, and ADHD, are not affected by the SSA. Given that learning disabilities, dyslexia/dyscalculia, and ADHD are mostly determined by genetic inheritance, it is unlikely that SSA could trigger these conditions (Elder, 2010; Mühlenweg et al., 2012). We also find no effect for domestic violence, an outcome determined outside of school and unrelated to SSA.

The effect of SSA on behavioral problems is the main driver of the overall results, because being born on August 1 instead of July 31 nearly halves the probability of developing behavioral problems upon school entry. As the administrative records of the SPS reveal, there are three main categories of behavioral problems: socio-emotional problems, disruptive behavior, and violent behavior. It appears thus that children starting school at a younger age tend to display a lack of maturity in their behavior, as suggested by developmental psychologist (Whitebread, 2012). Our results are also consistent with Dee and Sievertsen (2018), who show that a one-year delay in the start of school reduces inattention and hyperactivity at age seven.

Table 2: Reduced-form effect of school starting age on special needs

(A) Special needs: Incidence, onset and severity

	Special needs (SN)	SN, onset before school	SN, onset during school	No diagnosis, dismissal	Positive diagnosis	Nr. of consultations
τ_{RD}	-0.051***	0.018	-0.067***	0.005	-0.054***	-0.692***
	(0.019)	(0.012)	(0.018)	(0.010)	(0.020)	(0.256)
Bandwidth	56	40	53	50	49	55
$\bar{Y}_{-\text{bw}}$	0.353	0.069	0.284	0.058	0.295	3.221
$N_{-\text{bw}}$	8,056	5,604	7,607	7,179	6,898	7,901
$N_{+\text{bw}}$	8,137	5,533	7,653	7,169	6,863	7,961
N	51,138	51,138	51,138	51,138	51,138	51,138

(B) Special needs: Type

	Learning impairment	Behavioral problem	Speech impediment	Dyslexia/ dyscalculia	ADHD	Domestic violence
τ_{RD}	-0.028	-0.035***	-0.037**	-0.022	-0.009	-0.003
	(0.018)	(0.013)	(0.015)	(0.013)	(0.008)	(0.007)
Bandwidth	41	38	38	43	45	49
$\bar{Y}_{-\text{bw}}$	0.198	0.075	0.122	0.105	0.033	0.025
$N_{-\text{bw}}$	5,867	5,471	5,471	6,036	6,483	6,898
$N_{+\text{bw}}$	5,817	5,377	5,377	5,961	6,419	6,863
N	51,138	51,138	51,138	51,138	51,138	51,138

Notes: Estimates for τ_{RD} correspond to the treatment effect derived in section 3. All models include birth cohort specific effects and indicators for gender and non-native speaker. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

4.3 Medium-term outcomes

Early differences in the development of special needs can lead to persistent differences in achievement. However, children registered with the SPS typically receive support, and they may also repeat a grade if the gap in achievement is too large for them to bridge. Educational achievement differences between younger and older children may thus fade away over time (Crawford, Dearden, and Greaves, 2014).

In Table 3, we test whether differences still persist at the age when children are about to finish secondary school and take a compulsory standardized test. To remove possible grading effects, the regressions control additionally for the year of the test. We also investigate two distinct grade retention outcomes. First, whether a child was assigned to a bridge year between kindergarten and primary school. The purpose of this bridge year is to smooth the transition from the kindergarten environment to the school environment. Second, whether a child repeated a grade by the end of secondary school.

The first set of outcomes focuses on educational achievement in eighth grade, towards the end of compulsory schooling (around age 15). We find that children born on August 1 or after consistently outperform younger students. Their test scores in Math, German, and a composite score[9] of both are on average 0.1 standard deviations better than those of children born on July 31 or before. Although we cannot separate the SSA effect from the age-at-test effect, our results corroborate the findings in the literature (Bedard and Dhuey, 2006; Black, Devereux, and Salvanes, 2011; Fredriksson and Öckert, 2014).

While we do not find that younger children have a higher risk of failing a grade, younger children are 45% more likely to do a bridge year between kindergarten and primary school. This result indicates that many children starting school at a younger age will eventually prolong their compulsory education by one year. A flexible school system that allows children

[9] Using a composite score has the advantage of increasing precision by reducing measurement error in the dependent variable (West and Peterson, 2006).

to do so might be key to alleviate disadvantages of younger children when entering the labor market. Adding this additional year early in the school career might be preferable to later grade repetition, since it avoids early disadvantages that might carry forward through several grades or any stigma effect associated with failing a grade.

4.4 Long-term outcomes

The last set of outcomes of interest is post-compulsory education decisions, employment, and earnings. We investigate whether the differences in SSA have a lasting impact on life transitions and labor market entry for young adults.

Table 4 presents the results for post-compulsory education choices. The table focuses on three outcomes defined as follows. First, a summary measure of whether a student entered any post-compulsory education track. Second, for those who started post-compulsory education, whether a student entered the VET track or the academic preparation track. Third, for those who choose VET, whether they selected the high-quality track.[10]

Overall, we do not find that achievement differences persist into differences in success at transition to post-compulsory education. Children entering school early are just as likely to enter VET by signing a training contract as children entering school later. Similarly, younger children are also equally likely to enter the academic preparation track. However, we find that among students entering the vocational track, being born just after the cutoff date increases the probability of choosing a high-quality vocational track. This finding suggests that although SSA does not influence the choice between academic or vocational education, SSA affects the choice of the type of VET.

For individuals who entered the labor market between 2007 and 2016, we have infor-

[10]Switzerland has two tracks within the vocational system. The higher quality track lasts three to four years and is dedicated to jobs that require extensive skills and competences (e.g., computer scientist or nurse). The lower quality track lasts two years and is devolved to jobs that require basic skills (e.g., painter or plumber). For many jobs, both a high- and a low-quality track is available.

Table 3: Reduced-form effect of school starting age on educational outcomes

	STANDARDIZED TEST IN GRADE 8			GRADE RETENTION	
	Composite	Math	German	Bridge year between KG and PS	Grade repetition by 8th grade
τ_{RD}	0.099**	0.075*	0.110***	-0.027***	-0.003
	(0.042)	(0.044)	(0.041)	(0.008)	(0.005)
Bandwidth	44	44	44	61	72
$\bar{Y}_{-\text{bw}}$	0.004	0.004	0.004	0.059	0.018
$N_{-\text{bw}}$	6,295	6,295	6,295	8,652	10,104
$N_{+\text{bw}}$	6,240	6,240	6,240	8,836	10,344
N	50,717	50,717	50,717	50,717	50,717

Notes: Estimates for τ_{RD} correspond to the treatment effect derived in section 3. All models include birth cohort specific effects and indicators for gender and non-native speaker. Models in columns 1, 2, and 3 also control for year of test. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

Table 4: Reduced-form effect of school starting age on post-compulsory education choice

	Post-compulsory education started	Vocational track (vs. academic)	High-quality vocational track
τ_{RD}	0.017	0.037	0.061**
	(0.020)	(0.027)	(0.024)
Bandwidth	40	29	33
$\bar{Y}_{-\text{bw}}$	0.826	0.788	0.857
$N_{-\text{bw}}$	4,652	2,771	2,527
$N_{+\text{bw}}$	4,647	2,769	2,424
N	41,642	34,249	27,896

Notes: Estimates for τ_{RD} correspond to the treatment effect derived in section 3. All models include birth cohort specific effects and indicators for gender and non-native speaker. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

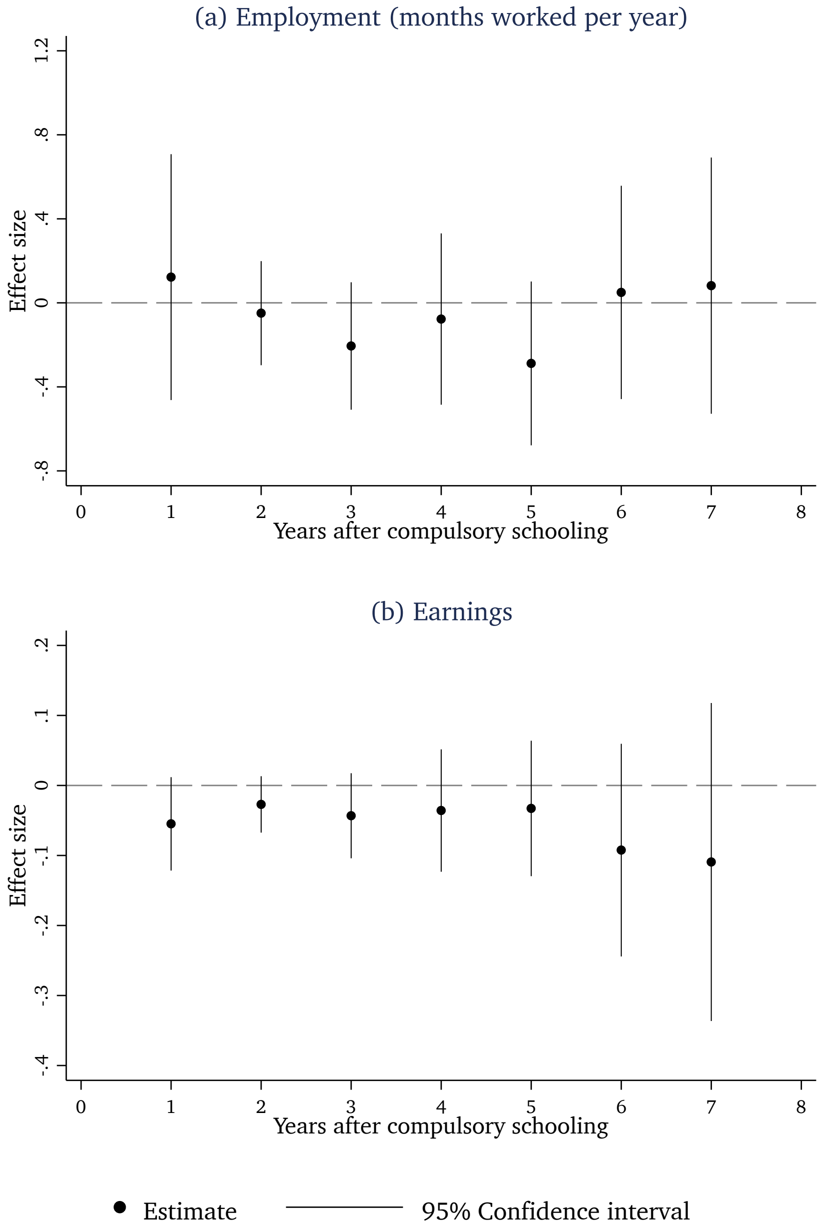
mation about their earnings and employment. To make labor market outcomes more comparable, we divide these outcomes according to when they are measured relative to compulsory school end. In detail, we examine employment and earnings one year after compulsory school, two years after compulsory school, and so on, up to seven years after compulsory school end (the maximum we observe in the data).

Figure 3 summarizes the results, divided into employment (panel a) and earnings (panel b). As for the main models, estimations control for birth year fixed-effects, gender, and non-native speaker, and are based on triangular kernel weights and a bandwidth choice following Calonico et al. (2018). We find that SSA has no effect on either the number of months per year worked or monthly earnings. The points estimates are all relatively small in size and similar to each other, suggesting that we are actually estimating a zero effect. These results confirm findings from Dustmann, Puhani, and Schönberg (2017), who find no effects of SSA on experience-adjusted wages and employment in Germany.[11]

In sum, our analysis of the effects of SSA from kindergarten through the labor market indicates that students starting school at a younger age are more likely to develop behavioral problems upon primary school entry. This effect diminishes over time and, although being present for test scores in eighth grade, disappears after compulsory schooling and is not found in the labor market.

[11] In their study, Dustmann, Puhani, and Schönberg (2017) use the school cutoff to instrument for track attendance during compulsory schooling. Performing their reduced form analysis in our data with a binary outcome for being in the high track in secondary school yields an insignificant effect of 0.006 (p-value = 0.714). The exclusion restriction in the paper of Dustmann, Puhani, and Schönberg (2017) states that SSA influences wages and employment exclusively through tracking. Our main findings on special needs outcomes challenge the validity of this exclusion restriction, since we show strong effects on the development of special needs arising before tracking takes place.

Figure 3: The effect of SSA on labor market outcomes



Notes: Local linear regression results are based on triangular kernel weights and a bandwidth choice following Calonico et al. (2018). Earnings are on a logarithmic scale. Sample sizes are as follows: 5,604 for the first year, 18,928 for the second year, 21,181 for the third year, 17,521 for the fourth year, 13,453 for the fifth year, 8,977 for the sixth year, and 4,393 for the seventh year. Data are from the School Psychological Service, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

4.5 Two-sample instrumental variable estimation

The results presented so far are reduced-form estimates, because in our main data set we do not observe the actual school starting age.[12] In this section, we follow a two-sample instrumental variable approach (TSIV) and instrument actual school starting age with the cutoff rule.

We obtain a first stage from the official education statistics (SDL, which stands for *Statistik der Lernenden* in German). The SDL collects the administrative records of the universe of students in Switzerland and the data are provided by the Swiss Federal Statistical Office. It includes all individuals who participate – part-time or full-time – in a program for a specific educational degree for at least half a year. The SDL has existed since 2010 and includes all levels of education from kindergarten to tertiary level (excluding universities).

We acquired the SDL data for the canton of St. Gallen. The data comprise the population of children who were born between 2006 and 2011 and started compulsory education between 2010 and 2016. The raw data contain 29,580 children, but we exclude those in special schools (820 children), those entering kindergarten before turning four (52 children), and those entering kindergarten after turning six (10 children). The final data set contains 28,698 observations, approximately 4,100 children entering compulsory education each year.

Note that the cohorts covered by the SDL (2006-2011) are not the same cohorts covered by our main data set (1992-2003). Although relying on two different cohorts makes the assumption of zero covariance between the first-stage and reduced-form estimates more plausible, we have to assume that the first stage effect from the SDL is generalizable to the older cohorts. To determine how credible this latter assumption is, we proceed as follows. First, we include cohort fixed-effects in the calculation of the first stage. This allows us to estimate a within-cohort first stage effect. Second, we estimate one first stage for each cohort

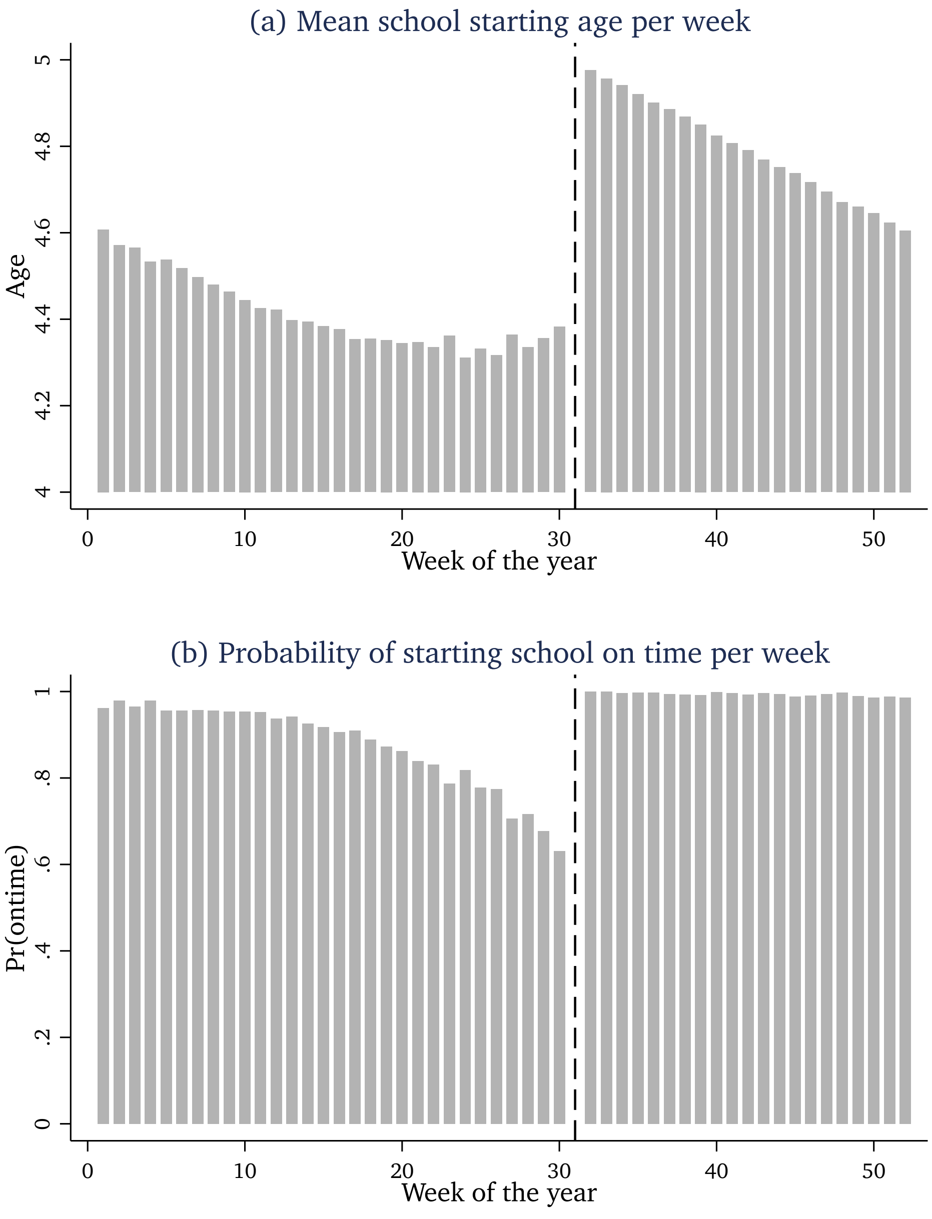
[12] In the SSA literature this is a common problem, see for example, Dalsgaard et al. (2012) and Evans, Morrill, and Parente (2010), who also apply a TSIV approach in a fuzzy RD context.

and examine whether the first stage effect is similar across different cohorts.

Before presenting the regression analyses, Figure 4 shows the average school starting age for each week of the year (panel a), along with the weekly average probability of starting school on time (panel b). Figure 4 has three findings. First, we observe a continuum of SSAs depending on children's date of birth, from children going to kindergarten when they are barely four years old to children who are already five as they start kindergarten. Second, we observe a clear discontinuity in SSA at the cutoff date. Children born in the week after the cutoff are approximately seven months older than children born in the week before the cutoff. Third, panel (b) provides some insights on redshirting practices in Switzerland. Redshirting is virtually non-existent for children born either after the cutoff or in the first three months of the year. From April on, the probability of starting school on time decreases from 90% to 63% in the week before the cutoff. The presence of redshirting underscores the importance of dividing the reduced-form estimates by the first stage, in order to estimate a local average treatment effect of SSA.

Appendix Table A1 presents the first stage estimates using the SDL data. Confirming the findings in Figure 4, the first stage effect is approximately 0.63, independent of the covariates included, the kernel function used, the inference method applied, and the bandwidth selected. The first stage effect is significant at the 1% level and corresponds to seven and a half months. In other words, children born shortly after the cutoff date start compulsory schooling seven months older than their peers born shortly before the cutoff. A comparison of children born in the months June and July and entering school on time with those born in the same months but entering school late shows, that red-shirting is significantly more common for boys (44%) and native speakers (43%) than for girls (28%) and non-native speakers (18%). Appendix Figure A3 shows that the first stage effect is very stable across cohorts, which mitigates the concern of using two different cohorts for reduced-form and first stage.

Figure 4: Graphical first stage



Notes: Each bin represents a calendar week. The vertical dashed line indicates the week of the cutoff. Data are from the Swiss Federal Statistical Office.

Table 5: Second stage effects of school starting age on special needs and educational outcomes

(A) Special needs: Incidence, onset and severity

	Special needs (SN)	SN, onset before school	SN, onset during school	No diagnosis, dismissal	Positive diagnosis	Nr. of consultations
β_{RD}	-0.081**	0.028	-0.107***	0.008	-0.086***	-1.103**
	(0.034)	(0.020)	(0.034)	(0.016)	(0.032)	(0.439)

(B) Special needs: Type

	Learning impairment	Behavioral problem	Speech impediment	Dyslexia/ dyscalculia	ADHD	Domestic violence
β_{RD}	-0.044	-0.056**	-0.059**	-0.035	-0.014	-0.004
	(0.028)	(0.022)	(0.025)	(0.022)	(0.012)	(0.015)

(C) Educational outcomes: test scores and grade retention

	Composite score	Math score	German score	Bridge year between KG and PS	Grade repetition by 8th grade
β_{RD}	0.158**	0.120	0.175**	-0.043***	-0.005
	(0.061)	(0.081)	(0.071)	(0.016)	(0.009)

(D) Post-compulsory education: track choice

	Post-compulsory education	VET track (vs. academic)	High-quality VET track
β_{RD}	0.027	0.059	0.097**
	(0.033)	(0.039)	(0.037)

Notes: Estimates for β_{RD} correspond to the treatment effect derived in equation 3. All models include birth cohort specific effects and indicators for gender and non-native speaker. Standard errors are computed using bootstrap (500 repetitions) and reported in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

Table 5 shows the second stage estimates for special needs (incidence, onset, severity, and type), educational outcomes (test scores and grade retention), and post-compulsory education track choice. We do not present second stage results for labor market outcomes because the reduced-form effect for these outcomes is not statistically significant. Overall, the second stage analysis confirms the reduced-form estimates. As expected, the second stage coefficients are larger in size and less precisely estimated. For example, being one year older at school start – and thus the oldest instead of the youngest student in class – reduces the probability of developing special needs during school by 10.7 percentage points, a large and economically relevant effect.

5 Sensitivity and robustness checks

We perform a series of robustness checks to demonstrate that our results are stable and not driven by a spurious correlation in the data. First, we perform specification checks using our main outcome, special needs incidence. In Table A2, we perform the analysis without covariates (column 1), then add birth cohort fixed effects (2), individual covariates (column 3, our main specification) and postal code fixed-effects (4). We change the inference to clustering at the running variable (column 5), use an asymmetric optimal bandwidth selection (6), and change the kernel used for weighting (7). None of these modifications change the size or significance of our main estimate.

Second, we perform a set of placebo tests. In Table A3, we assume a placebo cutoff in the middle of the distribution of the running variable both left (panel a) and right (panel b) of the original cutoff. We repeat the analysis for all outcomes used in Table 2. One estimate out of 24 is significant at 10%, that is, no more than what we should expect by chance.

Third, we verify that our results are not driven by a specific bandwidth choice. Although such concerns are mitigated by our algorithm-based bandwidth selections, we test the stability of our estimates and the bias-variance trade-off inherent to bandwidth choice by repeating our analysis for a large set of outcomes and bandwidths. Results are plotted in Figure A4. The graphs indicate no particular pattern deviating from our main results.

6 Discussion

This paper adds to a growing literature on the impact of SSA on diverse outcomes over the life-cycle. Most of the existing studies can be divided into three groups, based on their main outcomes. The first group studies the effect of SSA on the development of special needs conditions, with most attention given to ADHD and hyperactivity. The second group investigates early educational and cognitive achievements, and the third group looks at mid- to long-term outcomes including earnings, mental health, and crime.

The first group of papers finds that a higher SSA leads to decreases in the probability of receiving special education services (Dhuey and Lipscomb, 2010) and in the incidence of symptoms of inattention and hyperactivity (Dee and Sievertsen, 2018; Elder, 2010; Mühlenweg et al., 2012). We complement the findings of these papers by showing that the effect on special education needs is driven by onset after school enrollment and not due to pre-existing conditions. Moreover, we find that the development of behavioral problems is the main driver of the increase in special needs incidence, whereas learning impairments, ADHD, and dyslexia/dyscalculia, are unaffected by SSA.

The absence of a significant effect of SSA on ADHD incidence is an important difference from the literature. This difference is likely due to who performs the ADHD diagnosis. Dalsgaard et al. (2012) discuss the role of specialist behavior in the effect of SSA on the incidence of ADHD. They suggest that the effect of SSA on ADHD is driven by non-specialist diagnoses or over-referral of young children to special education services. Our results corroborate their findings and show that no effect exists when specialists are performing the diagnosis.

Results from the second group of papers indicate that a higher SSA increases test scores in grades three through eight. This finding is not only consistent across different institutional settings (Bedard and Dhuey, 2006; Dhuey et al., 2017; McEwan and Shapiro, 2008) but also comparable to our estimates in terms of sign and magnitude. Similarly, McEwan and Shapiro (2008) and Dhuey et al. (2017) also report compensatory behavior towards younger children such as redshirting and grade repetition. We complement on these findings showing that even in a setting with strong compensatory behavior in the forms of redshirting and institutionalized bridge year before grade one, there is still a significant and large effect of school starting age on the development of special needs, which carries forward to mid-term outcomes such as test-scores and the quality of VET.

The main concern with research about the SSA effect on test scores is age at test. As Crawford, Dearden, and Greaves (2014) argue, a large portion of the SSA effect on test scores is driven by age at test. Given that we have data on a standardized test administered towards the end of grade eight, we cannot distinguish between SSA effects and age-at-test effects (beyond controlling for year of birth). While this caveat clearly applies to test scores, it is less of a concern for the other outcomes under study. Special needs are diagnosed by experts individually and at different points in time for each concerned child, which should alleviate any age effects. Also, long-term outcomes are measured over longer time spans and therefore show less age effects on average.

Finally, the studies on long-term impacts of SSA find negligible effects on IQ scores, mental health at age 18, and earnings (Black, Devereux, and Salvanes, 2011); but some significant effects on criminal behavior at young ages (Landerso, Skyt Nielsen, and Simonsen, 2017). Similar to these studies, our results show no significant effects on medium- to long-run outcomes such as the probability of entering VET or academic preparation, earnings, and employment.

In summary, our paper connects and complements the above-mentioned strands of literature, by investigating different outcomes and mechanisms from childhood through the labor market. In that sense, most closely related to our work is Fredriksson and Öckert (2014), who study the effects of SSA on educational attainment and earnings over the life-cycle. They find strong effects for test scores but no effects for earnings. We add to this study by shedding light on the potential mechanism through which these effects operate, i.e., the incidence of special educational needs upon compulsory school entry.

7 Conclusions

The results of this paper suggest that starting school at a relatively younger age can be an important factor in the onset of special needs during the early years of primary school. At the extensive margin, younger children are more likely to be diagnosed with behavioral problems. At the intensive margin, they receive more frequent examinations and counseling by the school psychologists. Although younger children are more likely to be assigned to a bridge year after kindergarten and children with special needs receive therapies and support, they still score lower than their older peers in standardized tests at the end of compulsory schooling. However, the age differences at school start do not translate into differences in post-compulsory education choices and labor market outcomes. Both younger and older school starters are equally likely to start vocational education or academic preparation, and no significant difference is found for earnings and employment. Thus educational achievement differences do not appear to jeopardize the transition after compulsory schooling.

Taking this result into account, we maintain that educational differences due to SSA matter and need consideration in their own right. Most school systems are characterized by a universal date threshold that determines school start. Simple measures may mitigate the vulnerability of relatively young children who are born just before the cutoff date. It seems, the flexibility of the school system is key. Postponing school entry by one year, or granting an extra year to bridge between kindergarten and primary school, avoids carrying on early developmental delays over to compulsory schooling. Although the practice of redshirting already occurs before kindergarten, it is currently initiated exclusively by the parents. Given that parents of high socio-economic status are more likely to redshirt (Bassok and Reardon, 2013), this practice creates disadvantages for younger children in families from a lower socio-economic background. Instead, redshirting should be subject to an institutionalized process and external evaluation. For example, to efficiently identify children at risk, one possibility would be improving information sharing between preschool and kindergarten educators. Currently, such information sharing practice is not institutionalized in Switzerland. Nonetheless, at a slightly later stage, the bridge year in the canton of St. Gallen is a good example for such an institutionalized process, because all children are screened for school readiness by professionals before being sent to primary school.

Since we cannot differentiate between relative and absolute school starting age, investigating whether a general increase in school starting age would lead to a reduction in special needs conditions is outside the scope of this paper. Resolving this issue – possibly by means of a reform of school starting age – would be a valuable complement to our results and should thus be the focus of future research.

References

Angrist, Joshua D. and Alan 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):328–336.

Bassok, Daphna and Sean F. Reardon. 2013. "Academic redshirting in kindergarten: Prevalence, patterns, and implications." *Educational Evaluation and Policy Analysis* 35 (3):283–297.

Bedard, Kelly and Elizabeth Dhuey. 2006. "The persistance of early childhood maturity: International evidence of long-run age effects." *The Quarterly Journal of Economics* 121 (4):1437–1472.

Black, Sandra E. Paul J. Devereux, and Kjell G. Salvanes. 2011. "Too young to leave the nest? The effects of school starting age." *The Review of Economics and Statistics* 93 (2):455–467.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik. 2018. "Regression discontinuity designs using covariates." *The Review of Economics and Statistics* .

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82 (6):2295–2326.

Cheng, Ming-Yen, Jianqing Fan, and James S. Marron. 1997. "On automatic boundary corrections." *The Annals of Statistics* 25 (4):1691–1708.

Crawford, Claire, Lorraine Dearden, and Ellen Greaves. 2014. "The drivers of month-of-birth differences in children's cognitive and non-cognitive skills." *Journal of the Royal Statistical Society: Series A* 177 (4):829–860.

Dalsgaard, Soren, Maria Knoth Humlum, Helena Skyt Nielsen, and Marianne Simonsen. 2012. "Relative standards in ADHD diagnoses: The role of specialist behavior." *Economic Letters* 117 (3):663–665.

Datar, Ashlesha and Michael A. Gottfried. 2015. "School entry age and children's social-behavioral skills: Evidence from a national longitudinal study of U.S. kindergartners." *Educational Evaluation and Policy Analysis* 37 (3):333–353.

Dee, Thomas S. and William N. Evans. 2003. "Teen drinking and educational attainment: evidence from two-sample instrumental variables estimates." *Journal of Labor Economics* 21 (1):178–209.

Dee, Thomas S. and Hans Henrik Sievertsen. 2018. "The gift of time? School starting age and mental health." *Health Economics* 27 (5):781–802.

Devereux, Paul J. and Robert A. Hart. 2010. "Forced to be rich? Returns to compulsory schooling in Britain." *The Economic Journal* 120 (549):1345–1364.

Dhuey, Elizabeth, David Figlio, Krzysztof Karbownik, and Jeffrey Roth. 2017. "School starting age and cognitive development." Working Paper 23660, National Bureau of Economic Research.

Dhuey, Elizabeth and Stephen Lipscomb. 2010. "Disabled or young? Relative age and special education diagnoses in schools." *Economics of Education Review* 29 (5):857–872.

Duncombe, William and John Yinger. 2005. “How much more does a disadvantaged student cost?” *Economics of Education Review* 24 (5):513–532.

Dustmann, Christian, Patrick Puhani, and Uta Schönberg. 2017. “The Long-term Effects of Early Track Choice.” *The Economic Journal* 127 (603):1348–1380.

Elder, Todd E. 2010. “The importance of relative standards in ADHD diagnoses: Evidence based on exact birth dates.” *Journal of Health Economics* 29 (5):641–656.

Elder, Todd E. and Darren H. Lubotsky. 2009. “Kindergarten entrance age and children’s achievement impacts of state policies, family background, and peers.” *Journal of Human Resources* 44 (3):641–683.

Evans, William N., Melinda S. Morrill, and Stephen T. Parente. 2010. “Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children.” *Journal of Health Economics* 29 (5):657–673.

Fredriksson, Peter and Björn Öckert. 2014. “Life-cycle effects of age at school start.” *The Economic Journal* 124 (579):977–1004.

Gelman, Andrew and Guido Imbens. 2018. “Why high-order polynomials should not be used in regression discontinuity designs.” *Journal of Business & Economic Statistics* .

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. “Identification and estimation of treatment effects with a regression-discontinuity design.” *Econometrica* 69 (1):201–209.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2002. “Inferring program effects for special populations: Does special education raise achievement for students with disabilities?” *The Review of Economics and Statistics* 84 (4):584–599.

Inoue, Atsushi and Gary Solon. 2010. “Two-sample instrumental variables estimators.” *The Review of Economics and Statistics* 92 (3):557–561.

Landerso, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen. 2017. “School starting age and the crime-age profile.” *The Economic Journal* 127 (602):1096–1118.

McCrary, Justin. 2008. “Manipulation of the running variable in the regression discontinuity design: A density test.” *Journal of Econometrics* 142 (2):698–714.

McEwan, Patrick and Joseph S. Shapiro. 2008. “The benefits of delayed primary school enrollment: Discontinuity estimates using exact birth dates.” *Journal of Human Resources* 43 (1).

Morrison, Frederick J., Denise M. Alberts, and Elizabeth M. Griffith. 1997. “Nature-nurture in the classroom: Entrance age, school readiness, and learning in children.” *Developmental Psychology* 33 (2):254–262.

Mühlenweg, Andrea, Dorothea Blomeyer, Holger Stichnoth, and Manfred Laucht. 2012. “Effects of age at school entry (ASE) on the development of non-cognitive skills: Evidence from psychometric data.” *Economics of Education Review* 31 (3):68–76.

OECD. 2008. *Students with Disabilities, Learning Difficulties and Disadvantages – Policies, Statistics and Indicators*. Paris: OECD Publishing.

Oswald, Yvonne and Uschi Backes-Gellner. 2014. “Learning for a bonus: How financial incentives interact with preferences.” *Journal of Public Economics* 118:52–61.

Schwandt, Hannes and Amelie Wuppermann. 2016. "The youngest get the pill: ADHD misdiagnosis in Germany, its regional correlates and international comparison." *Labour Economics* 43:72–86.

Stipek, Deborah J. and Rosaleen H. Ryan. 1997. "Economically disadvantaged preschoolers: Ready to learn but further to go." *Developmental Psychology* 33 (4):711–723.

Van den Berg, Gerard, Pia Pinger, and Johannes Schoch. 2016. "Instrumental variable estimation of the causal effect of hunger early in life on health later in life." *The Economic Journal* 126 (591):465–506.

Wagner, Mary M. and Jose Blackorby. 1996. "Transition from high school to work or college: How special education students fare." *The Future of Children* 6 (1):103–120.

West, Martin R. and Paul E. Peterson. 2006. "The efficacy of choice threats within school accountability systems: Results from legislatively induced experiments." *The Economic Journal* 116 (510):C46–C62.

Whitebread, David. 2012. *Developmental psychology and early childhood education: a guide for students and practitioners*. Sage Publications, London (U.K.).

Wolter, Stefan C. and Paul Ryan. 2011. "Apprenticeship." In *Handbook of the Economics of Education*, vol. 3, edited by Erik Hanushek, Stephen Machin, and Ludger Woessmann, chap. 11. Elsevier, 521–576.

Appendix: Tables and Figures

Table A1: First stage analysis

	(1)	(2)	(3)	(4)	(5)	(6)
γ_{RD}	0.622***	0.623***	0.627***	0.626***	0.626***	0.627***
	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)	(0.017)
Cohort FE		✓	✓	✓	✓	✓
Indiv. covariates			✓	✓	✓	✓
Kernel	Triangular	Triangular	Triangular	Triangular	Epanechnikov	Triangular
Inference method	Robust	Robust	Robust	Robust	Robust	Cluster
BW selection	Symmetric	Symmetric	Symmetric	Asymmetric	Symmetric	Symmetric
bw_-	63	63	69	66	64	68
bw_+	63	63	69	17	64	68
$\bar{Y}_{-\text{bw}_-}$	4.358	4.358	4.358	4.358	4.358	4.358
$N_{-\text{bw}_-}$	5,301	5,301	5,787	5,507	5,370	5,587
$N_{+\text{bw}_+}$	5,052	5,052	5,552	1,395	5,113	5,385
N	28,698	28,698	28,698	28,698	28,698	28,698

Notes: Estimates for γ_{RD} are derived by local linear regression with bandwidth choice following Calonico et al. (2018). *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the Swiss Federal Statistical Office.

Table A2: Specification checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
τ_{RD}	-0.047***	-0.050***	-0.051***	-0.049***	-0.051**	-0.050***	-0.050***
	(0.018)	(0.018)	(0.019)	(0.019)	(0.022)	(0.018)	(0.018)
Cohort FE		✓	✓	✓	✓	✓	✓
Indiv. covariates			✓	✓			
Postcode FE				✓			
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Epanechnikov
Inference method	Robust	Robust	Robust	Robust	Cluster	Robust	Robust
BW selection	Symmetric	Symmetric	Symmetric	Symmetric	Symmetric	Asymmetric	Symmetric
bw_-	65	67	56	56	55	88	58
bw_+	65	67	56	56	55	52	58
$\bar{Y}_{-\text{bw}_-}$	0.353	0.353	0.353	0.353	0.353	0.353	0.353
$N_{-\text{bw}_-}$	9,256	9,380	8,056	8,056	7,745	12,353	8,317
$N_{+\text{bw}_+}$	9,458	9,602	8,137	8,137	7,817	7,340	8,424
N	51,138	51,138	51,138	51,138	51,138	51,138	51,138

Notes: Estimates for τ_{RD} correspond to the treatment effect derived in section 3. *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

Table A3: Placebo cutoffs

(A) PLACEBO: DAYS BEFORE AUGUST 1, ARTIFICIAL CUTOFF DATE ON APRIL 18

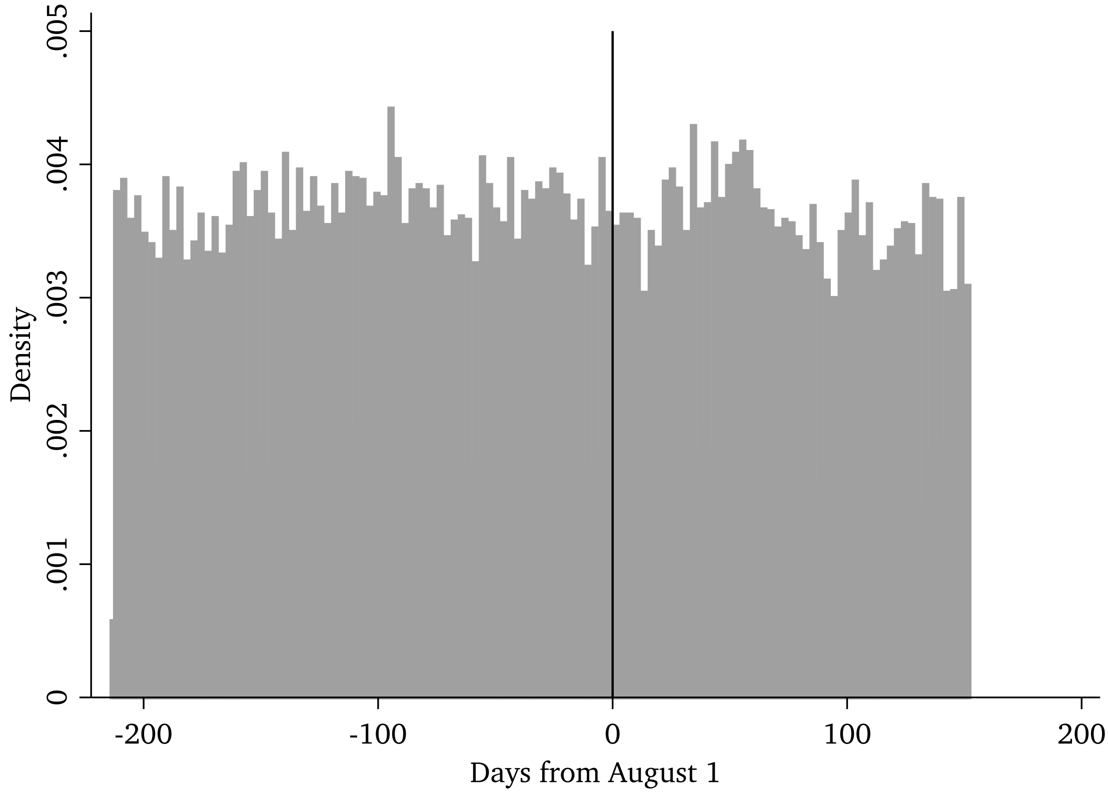
	ONSET AND SEVERITY						DIAGNOSIS					
	Special needs (SN)	SN, onset before school	SN, onset during school	No diagnosis, dismissal	Positive diagnosis	Nr. of consultations	Learning impairment	Behavioral problem	Speech impediment	Dyslexia/ dyscalculia	ADHD	Domestic violence
τ_{RD}	0.038	0.005	0.031	0.020*	0.023	0.255	0.037	0.006	-0.010	0.012	0.005	0.003
	(0.026)	(0.014)	(0.024)	(0.012)	(0.026)	(0.321)	(0.024)	(0.014)	(0.015)	(0.014)	(0.008)	(0.009)
BW	30	35	32	37	29	45	26	26	41	38	24	29
$N_{-\mathrm{bw}}$	4,279	5,029	4,547	5,179	3,984	6,302	3,700	3,700	5,737	5,442	3,421	3,984
$N_{+\mathrm{bw}}$	4,604	5,289	4,898	5,423	4,327	6,501	4,038	4,038	5,952	5,697	3,746	4,327
N	30,021	30,021	30,021	30,021	30,021	30,021	30,021	30,021	30,021	30,021	30,021	30,021

(B) PLACEBO: DAYS AFTER AUGUST 1, ARTIFICIAL CUTOFF DATE ON OCTOBER 15

	ONSET AND SEVERITY						DIAGNOSIS					
	Special needs (SN)	SN, onset before school	SN, onset during school	No diagnosis, dismissal	Positive diagnosis	Nr. of consultations	Learning impairment	Behavioral problem	Speech impediment	Dyslexia/ dyscalculia	ADHD	Domestic violence
τ_{RD}	0.010	0.003	0.005	0.001	0.023	-0.066	0.002	0.015	-0.032	-0.001	-0.010	0.012
	(0.031)	(0.017)	(0.026)	(0.014)	(0.030)	(0.335)	(0.022)	(0.017)	(0.021)	(0.017)	(0.011)	(0.009)
BW	22	18	26	31	19	31	29	20	19	31	19	29
$N_{-\mathrm{bw}}$	3,042	2,558	3,832	4,573	2,734	4,573	4,276	2,878	2,734	4,573	2,734	4,276
$N_{+\mathrm{bw}}$	2,913	2,529	3,576	4,253	2,652	4,253	3,978	2,759	2,652	4,253	2,652	3,978
N	20,722	20,722	20,722	20,722	20,722	20,722	20,722	20,722	20,722	20,722	20,722	20,722

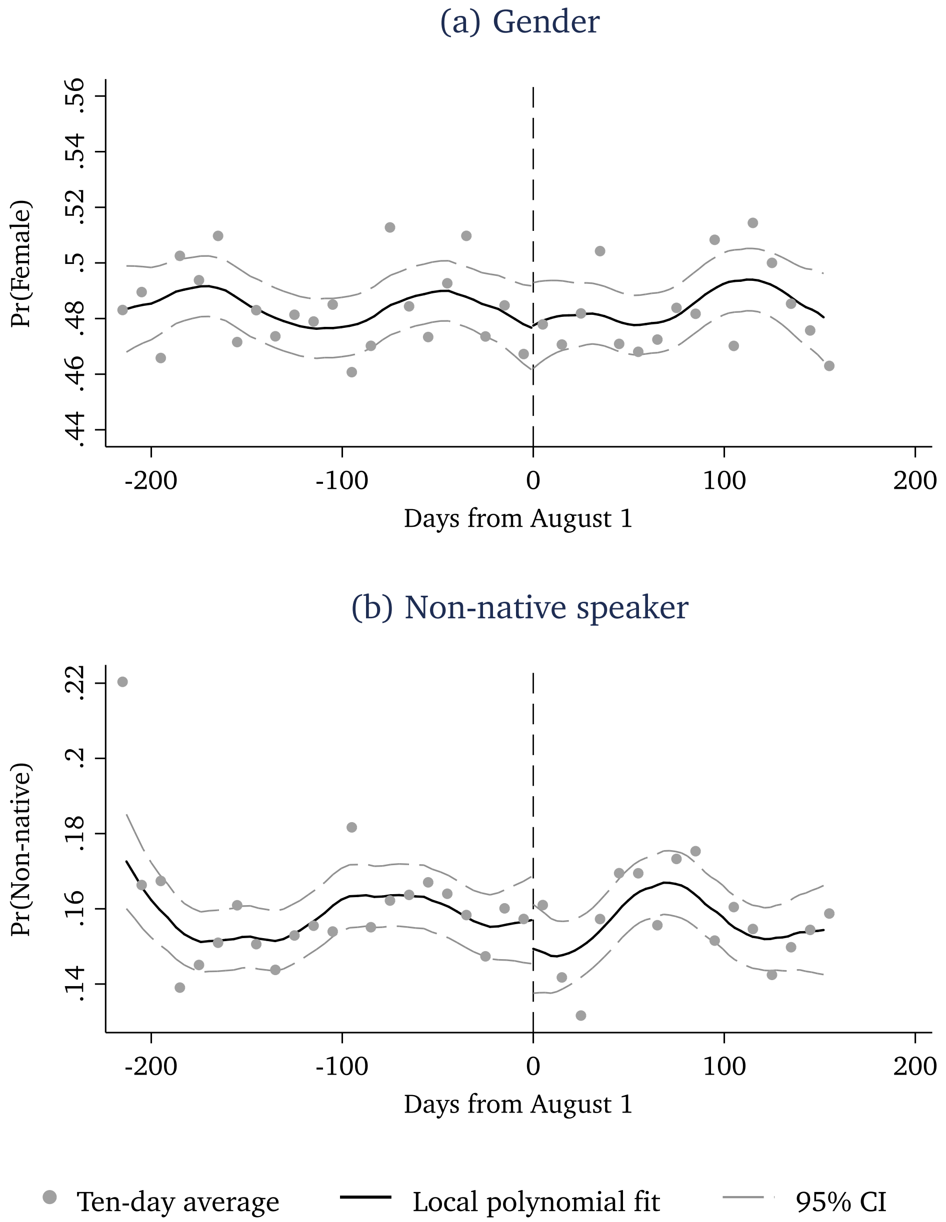
Notes: Estimates for τ_{RD} correspond to the treatment effect derived in section 3. All models include birth cohort specific effects and indicators for gender and non-native speaker. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% level, respectively. Data are from the School Psychological Service St. Gallen, the ministries of education canton of St. Gallen, the Stellwerk-Test service provider, the Swiss Federal Statistical Office, and the Swiss Central Compensation Office.

Figure A1: Distribution of birth dates



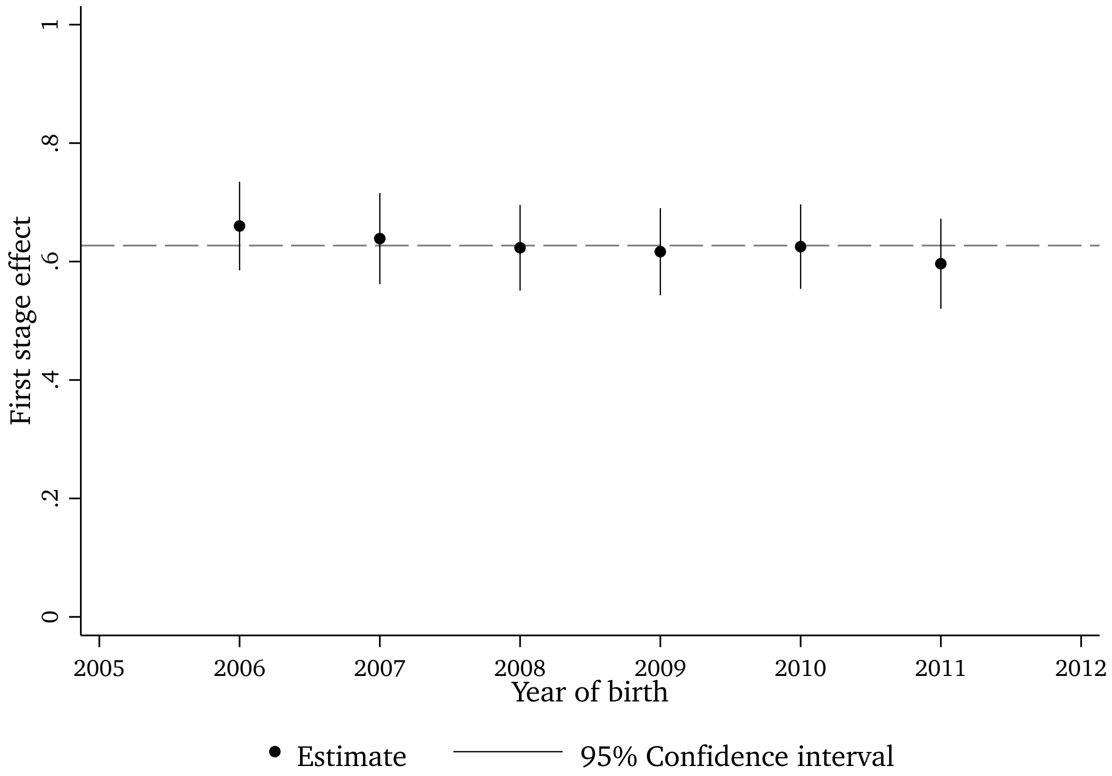
Notes: Data are from the Stellwerk-Test service provider.

Figure A2: Covariate balance



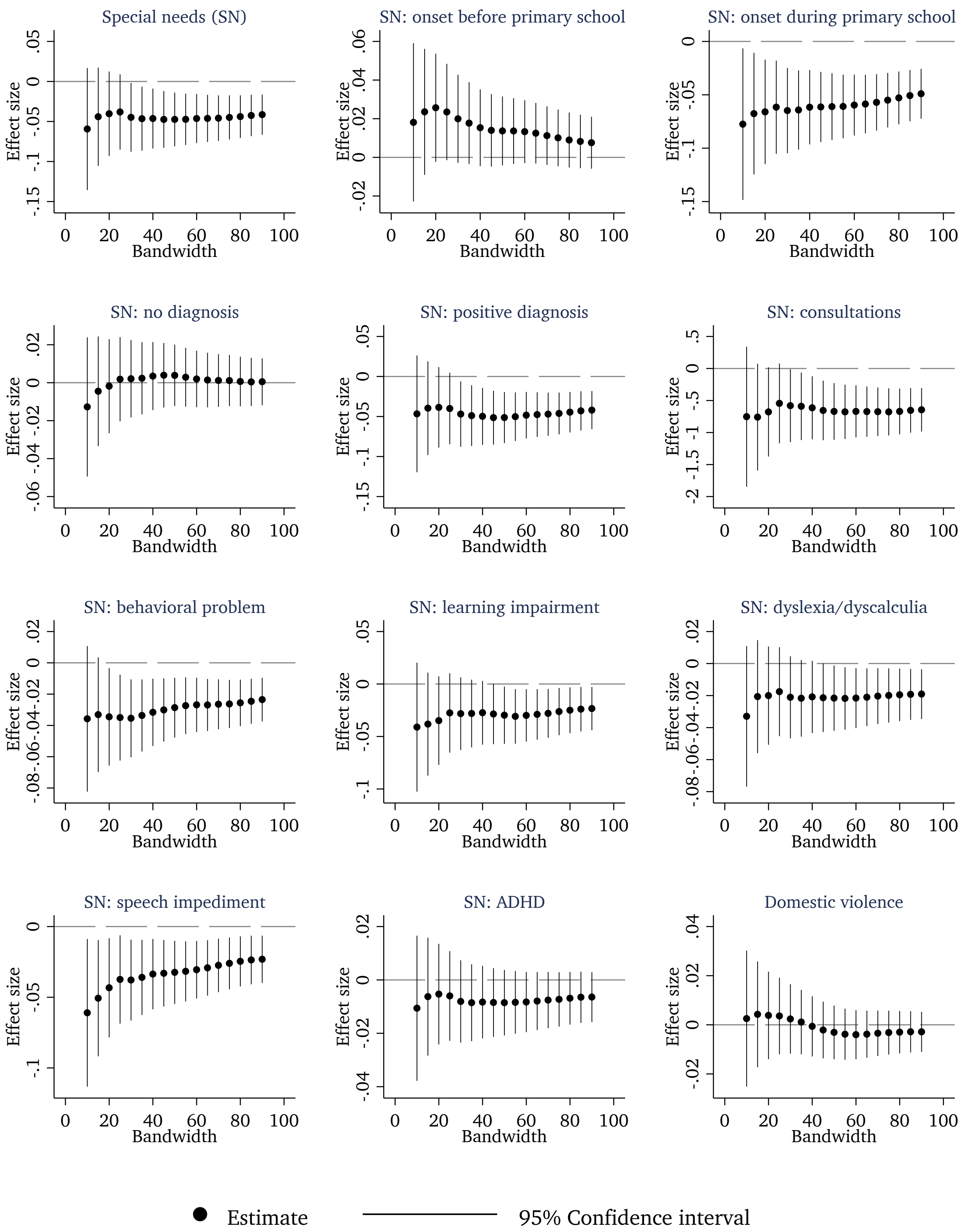
Notes: Data are from the Stellwerk-Test service provider.

Figure A3: First stage effect by birth year



Notes: Local linear regression results are based on triangular kernel weights and a bandwidth choice following Calonico et al. (2018). Sample sizes are as follows: 4,474 for 2006, 4,616 for 2007, 4,817 for 2008, 4,926 for 2009, 4,926 for 2010, and 4,939 for 2011. The dashed line indicates the overall first stage effect. Data are from the Swiss Federal Statistical Office.

Figure A4: Bandwidth variations



Notes: Data are from Stellwerk-Test service provider and the School Psychological Service St. Gallen.