Swiss Leading House
Economics of Education • Firm Behaviour • Training Policies

Working Paper No. 170

High-Ability Influencers? The Heterogeneous Effects of Gifted Classmates

Simone Balestra, Aurélien Sallin and Stefan C. Wolter

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 170

High-Ability Influencers? The Heterogeneous Effects of Gifted Classmates

Simone Balestra, Aurélien Sallin and Stefan C. Wolter

May 2020

The Swiss Leading House on Economics of Education, Firm Behavior and Training Policies is a Research Program of the Swiss State Secretariat for Education, Research, and Innovation (SERI).

www.economics-of-education.ch

High-Ability Influencers?

The Heterogeneous Effects of Gifted Classmates*

Simone Balestra[1], Aurélien Sallin[1], and Stefan Wolter[2,3,4]

[1]*University of St. Gallen, Switzerland*

[2]*University of Bern, Switzerland*

[3]*CESifo and IZA, Germany*

[4]*Swiss Coordination Centre for Research in Education, Switzerland*

Abstract

This paper examines how exposure to students identified as gifted ($IQ \geq 130$) affects achievement in secondary school and enrollment in post-compulsory education. By using unique student-level administrative data on achievement combined with psychological examination records, we are able to study the causal impact of gifted students on their classmates in unprecedented detail. We find a positive and significant effect of the exposure to gifted students on school achievement in both math and language. The impact of gifted students is, however, highly heterogeneous along three dimensions. First, we observe the strongest effects among male students and high achievers. Second, we show that male students benefit from the presence of gifted peers in all subjects regardless of their gender, whereas female students seem to benefit exclusively from the presence of female gifted students. Third, we find that gifted students diagnosed with emotional or behavioral disorders have zero-to-negative effects on their classmates' performance, a detrimental effect more pronounced for female students. After compulsory schooling, the results show that exposure to gifted classmates increases the likelihood of choosing a selective academic track. This effect, however, is entirely driven by male students.

Keywords: gifted students, peer quality, gender, math, peer effects

JEL Classification: I21, I24, I26, J24

*We are grateful to Mirjam Bächli, Uschi Backes-Gellner, Eric Bettinger, Anne Brenøe, Beatrix Eugster, Guido Heineck, Immanuel Lampe, Edward Lazear, Helge Liebert, Joan Llull, Fanny Puljic, Ulrike Unterhofer, Martina Viarengo, and Ulf Zölitz for their constructive comments. The paper benefited from valuable feedback at the following meetings: Bildungsökonomischer Ausschuss im Verein für Socialpolitik, European Association of Labour Economists, Ski & Labor Seminar, Swiss Society of Economics and Statistics, Verein für Socialpolitik, Young Swiss Economists Meeting, and departmental seminars at the universities of St. Gallen and Zurich. For the data, we thank the School Psychological Service, the Ministry of Education of the canton of St. Gallen, and the Stellwerk test service provider. We acknowledge financing from the Swiss National Science Foundation (grant no. 176381). Address for correspondence: Aurélien Sallin, University of St. Gallen, Rosenbergstrasse 51, CH-9000 St. Gallen. Email: aurelien.sallin@unisg.ch. Telephone: +4171 224 31 93.

1 Introduction

In a context where the inclusion of special-needs students in the main classroom ("mainstreaming") is becoming the norm and where special education programs are increasingly being abandoned, evidence on the effects of inclusion on students' wellbeing, achievement, and post-education opportunities is more needed than ever. One particular population, which is traditionally segregated into special education classes, needs to be thoroughly investigated in a mainstreaming context: *gifted students* – i.e., students with an intellectual ability significantly higher than average.[1] It is a priori unclear whether and where such students exert a positive influence on their classmates, where – feeling bored and not fitting in – they are perceived as disruptive elements, and where they have no discernible effect on their peers. The aim of the present paper is to resolve this question by examining if and how gifted students affect their non-gifted classmates' achievement in secondary school and enrollment in post-compulsory education. Given the heterogeneous nature of peer effects in the classroom[2], our research emphasizes how the influence of gifted students differs for their male and female, high-achieving and low-achieving peers in math and non-math classes.

We analyze the impact of gifted students on their classroom peers in an inclusive academic setting, which offers ideal conditions for the identification of ability peer effects. One such feature is that no gifted students are segregated into gifted programs and that they are all included in regular schools, even though they may receive additional services or activities outside of the classroom. A second feature is that the status of gifted students is assessed and determined by the school psychological service, an independent and centralized institution that provides students and their families with diagnosis and counseling for school-related issues. This practice ensures that professional psychologists (and not parents, teachers or school administrators) diagnose students as gifted, and allows us to differentiate gifted students from simply high-achieving students. We carry out the analyses in the context of the Swiss education system, and we use student-level administrative data on achievement combined with detailed psychological examination records, uniquely linking students' school performance in a compulsory standardized test and administrative records from the school psychological service for ten consecutive cohorts of eight graders. To investigate career trajectories, we merge our data with administrative records containing detailed information on students' post-compulsory education choices.

[1] We understand gifted children or students as "children, students or youth who give evidence of high performance capability in areas such as intellectual, creative, artistic, or leadership capacity, or in specific academic fields, and who require services or activities not ordinarily provided by the school in order to fully develop such capabilities." (US Federal government statutory definition of gifted students, P.L. 103-382, Title XIV, p. 388).

[2] As pointed out by Booij, Leuven, and Oosterbeek (2017), see Black, Devereux, and Salvanes (2013); Burke and Sass (2013); Hoxby (2000); Lavy, Paserman, and Schlosser (2011); Lavy, Silva, and Weinhardt (2012).

For identification, we rely on the variation in classroom composition arising from within-school assignment of gifted students to classes when students transition from primary school (grades one through six) to secondary school (grades seven through nine). When transitioning from primary school to secondary school, students are assigned to new classes in their new school and will remain in the same class for the rest of their mandatory education. For equity reasons and to avoid stigma, information on students' psychological profiles are usually not shared between primary and secondary schools. This practice implies that gifted students can be neither identified nor assigned to specific classrooms or teachers as students enter secondary school. We demonstrate, using several tests, that the observed within-school, between-class variation in the proportion of gifted classmates is consistent with variation generated from a random process. We also find no systematic assignment of gifted students to a specific class or teacher. Causal conclusions are further motivated by the fact that parents do not have a free choice of school in Switzerland: nearly all pupils attend public schools, and the geographical distribution of the population in the individual communities shows no regularities that would correlate with the distribution of talented pupils. Finally, the low prevalence of intellectual giftedness (1-2% of the population) is useful for the estimation of peer effects, allowing us to conduct the analysis exclusively on non-gifted students without losing a significant portion of the sample. By excluding gifted students from the analysis, we explicitly distinguish between the subjects of a peer effects investigation (regular students) and the peers who potentially provide the mechanism for causal effects on these subjects (gifted students).

The results are the following: we document a positive effect of exposure to students identified as gifted in all school subjects. Our results indicate that exposure to gifted students raises math and language achievement of the other students by 8.3% of a standard deviation on average. When looking at who benefits the most from the presence of gifted students in the classroom, we observe the strongest effect for male students and for high achievers. Low-achieving students do not react to the presence of gifted classmates. In addition, we uncover a clear effect heterogeneity across gender and school subject. In math, male students profit significantly more than female students from gifted classmates, further amplifying the gender gap in math achievement. In contrast, we find no significant gender difference in the ability peer effect for language. This gender-subject heterogeneity is quite striking because classrooms (and thus peer composition) remain the same for all subjects. Moreover, we detect no other significant effect heterogeneity for characteristics like student's age, their native speaker status, class size, or teacher's gender.

Which gifted students generate the positive externalities? We provide compelling evidence that both the gender and the behavior of the gifted students matter, and that they matter even more to female students. We show that male students benefit from the presence of gifted peers in all subjects regardless of the gender of the gifted, whereas female students benefit exclusively from gifted female students. This pattern is more apparent in math and suggests that exposure to high-ability female peers may provide female students with a role model in quantitative fields, alleviating the negative effects of gender stereotypes. Not every gifted student is, however, a good peer. By distinguishing between gifted students who suffer from behavioral, emotional, or social problems from the other gifted students, we are able to isolate gifted students who do not exhibit disruptive behavior in the classroom. We find that female students are negatively affected by the presence of classmates who are gifted but disruptive. The evidence suggests that well-behaved gifted students improve their classmates' performance through both ability spillovers and reduced classroom disruption.

In terms of human capital investment, we augment our main findings by analyzing students' educational career trajectories after compulsory education. By looking at whether students choose an academic track or a vocational track, we find that being exposed to gifted classmates in secondary school significantly increases the likelihood of choosing the academic track. This effect is entirely driven by male students who enter the academic track instead of the vocational track, which reflects the main findings and may offer an additional explanation to the persistent under-representation of women in math-intensive careers (e.g., in STEM fields).

The present paper contributes to and brings together three strands of literature in economics. First, we contribute to the under-investigated field of research on gifted students. Rather than looking at how gifted students perform when they are segregated into talented programs (e.g., Booij, Haan, and Plug, 2016; Bui, Craig, and Imberman, 2014), we bring evidence on the situation of gifted students in an inclusive education system, and we propose a new approach to identify high-ability students in general. The literature so far has used previous achievement (e.g., Booij, Leuven, and Oosterbeek, 2017), individual fixed effects (e.g., Burke and Sass, 2013), socioeconomic background (e.g., Black, Devereux, and Salvanes, 2013), and parents' education (e.g., Cools, Fernández, and Patacchini, 2019) to determine student ability. Instead, we use formal assessments by external specialists (school psychologists) to identify gifted students. These external assessments are reliable assessment of students' cognitive abilities, extend beyond pure school performance and are less prone to biases arising from parents, teachers, or developmental factors. Given that being diagnosed as a "gifted" student does not automatically determine eligibility to targeted academic programs in Switzerland, our hybrid measure based on specialists' assessment and IQ tests is less likely to be manipulated.

Second, we contribute to the extensive literature on peer effects in education. This literature has provided quantified evidence that educational success cannot be explained only by students' own characteristics, parental background, and school environment, but that peers and the interactions between peers matter. One novel feature of our study is that, among the many peer dynamics occurring in the classroom documented so far, the heterogeneous influence of the population of gifted students has never been investigated. In addition, we are able to observe the classroom environment, where teaching occurs and students presumably affect directly their peers' learning. Although many scholars agree that classroom interactions play an important role in determining students' academic achievement and in shaping students' educational choices, most studies define peers at the school or cohort level. This definition of peer group may miss important interactions within classroom groups, because the estimation of spillover effects differs depending on the accuracy with which one identifies the set of relevant peers (Carrell, Fullerton, and West, 2009; Carrell, Sacerdote, and West, 2013). Finally, by presenting evidence on both school performance and career trajectory, we complement an emerging literature examining how peer characteristics during adolescence influence later career choices (Anelli and Peri, 2017; Black, Devereux, and Salvanes, 2013; Card and Payne, 2017; Carrell, Hoekstra, and Kuka, 2018; Mouganie and Wang, forthcoming; Zölitz and Feld, 2019).

Third, we contribute to a growing strand of literature that aims to understand the roots of the persistent gender gap in math (for a recent review, see Buckles, 2019). Although the gender gap in education enrollment and labor market participation has dramatically narrowed over the past 50 years, the gender gap in math achievement still persists in most developed countries (Ellison and Swanson, 2010).[3] The reasons for this persistence are still not totally understood: recent research shows that the gender gap in math achievement does not exist upon entry to school, supporting the idea that nurture (e.g., gender stereotypes, culture) rather than nature (e.g., innate biological differences between sexes) determines gender differences in achievement (Hyde and Mertz, 2009; Nosek et al., 2009; Pope and Sydnor, 2010). However, the gap appears to be large and significant in the middle school years and beyond (Fryer and Levitt, 2010), and is in turn mirrored in the education and career choices of young women (Brenøe and Zölitz, forthcoming; Card and Payne, 2017; Carrell, Page, and West, 2010). It is therefore crucial to understand what are the factors in the school environment that originate and widen the gender gap in math achievement, especially for students at the age of choosing their first important career direction. This study offers both new evidence on the formation of the gender gap in math and the likely mechanisms behind such gap.

[3] Data from the 2015 PISA study reveal that Switzerland has one of the highest gender gaps in math performance (2.3%), alongside with other countries exhibiting above-average gender gaps like the United States (1.9%), the United Kingdom (2.2%), and Germany (3.1%).

2 Background and Data

2.1 Institutional Background

The education system in Switzerland has a federal structure and gives the cantons – similar to the states in the U.S., the countries in Germany, or the provinces in Canada – great freedom in educational policy decision-making. In contrast to the other three federal states, however, the degree of coordination between the cantons is relatively high and, depending on the language region (German, French, or Italian), the cantons now apply the same common curriculum in all subjects. In the Intercantonal Agreement on the Harmonization of Compulsory Education, which the majority of the cantons – including the one we consider in the present paper – have joined, equal school structures were established. These include a two-year entry level (kindergarten) and nine years of compulsory schooling, of which the first six years are allocated to the primary level and the last three to the lower secondary level. Pupils change schools and classes when moving from primary to lower secondary education.

Within each canton, schools are organized at the municipality level, and children are assigned to schools on the pure basis of their location of residence. This strict assignment procedure is thoroughly implemented, such that parents have no say about their child's school other than moving permanently to a different municipality or enrolling their children in a private school. Despite this rule, private schooling remains very rare in Switzerland. As the 2018 Education Report by the Swiss Coordination Center for Research in Education shows, more than 95% of children in 2016 attend public-funded schools in their municipality of residence.

The present analysis focuses on all students enrolled in the secondary schools of the Canton of St. Gallen (around 500,000 inhabitants). In this state, children are required to undergo eleven years of compulsory education, divided into kindergarten (two years), primary school (six years), and secondary school (three years). In most cases, secondary schooling takes place in larger schools administered by associations of municipalities (districts). Tracking occurs at the secondary school level and is based on students' academic performance in primary school as well as their teacher's recommendation. Students are either sent to a high-track secondary school (*Sekundarschule*) or to a lower, more practice-oriented secondary school track (*Realschule*). Once allocated to one of the two secondary school tracks, students are assigned to classes within each school-track. The administrative staff of each school has no prior knowledge about the students other than administrative data on gender, place of residence, primary school attended, and nationality. For equity reasons, information on students' disabilities, special needs, or high ability status is usually not shared between primary schools and secondary schools. This practice potentially creates a situation where class composition is quasi-random with respect to students' psychological profiles, a situation that we evaluate in the empirical strategy section. Once assigned to a class, students share the same peers for all the lectures and subjects, and classes remain unchanged for the three years of secondary school.[4]

At the end of their eighth year of compulsory schooling (second year of secondary school), all students are subjected to a mandatory standardized test (the so-called "Stellwerk 8" test). This computer-based adaptive test automatically adapts the difficulty of questions to the ability and knowledge revealed by the student in the previous questions (in the same fashion as the GMAT or GRE test). It tests core knowledge of mathematics, language (German), and, depending on the track, foreign languages (usually English) as well as natural sciences (including biology, chemistry, or physics). The correction of the test is computer-based, which eliminates concerns of teachers' bias or stereotyping in the results. The results are important both for students, who will use the test scores when choosing their post-compulsory education, and teachers, whose relative performance can be reflected by the rate of success of their students.

The Canton of St. Gallen bears the responsibility for the inclusion and education of children with high abilities and must guarantee the fulfillment of their educational needs. In this regard, emphasis is put on inclusion of students identified as gifted in regular classrooms (mainstreaming). Requests to send gifted children to special schools are accepted only under strict conditions: the child must have already skipped a grade, justify why the classroom environment is not adequate, and undergo a psychological evaluation. In all other cases (the vast majority), special activities, additional support, and enrichment programs are offered outside of class, depending on the school and upon request by parents and teachers. Acceleration (skipping a grade) or school start at a younger age are also possible in some rare cases.

The task of identifying and providing psychological support to children with high abilities is carried by the School Psychological Service (SPS), a centralized and independent office. It provides diagnoses of learning disabilities, behavioral difficulties and developmental deficiencies, assigns therapies and treatments to children, and offers counseling to children, parents, and teachers. For most students (about nine out of ten), services of the SPS are requested directly by the teacher, but some requests are also filed by the parents or the child's doctor. The referring party needs to justify its request by pointing out the reason for the child's registration with the SPS. The reasons most commonly brought up are learning disabilities, social or emotional problems, difficulties with the family and the parents, or challenging relationships with the teacher. After a request has been made, children and their parents are directly

[4]Grade repetition is not common in St. Gallen, as only about 1.5% of students in a cohort ever repeated a grade. Within-state and out-of-state mobility in St. Gallen are also low at about 2-3% (data are from the Swiss Federal Statistical Office for the years 2008-2017).

contacted by a caseworker from the SPS for an assessment of the situation and a health diagnosis. As part of the diagnosis, an intelligence test is often administered to children.

2.2 Data Sources

We use information on classroom composition, characteristics of students, and individual academic achievement from the Stellwerk test taken in eighth grade by the entire population of students in the Canton of St. Gallen.[5] To this, we add information from the administrative records of the SPS on individual psychological profile, giftedness status, and learning disabilities of each child. After merging these two sources, we observe the academic achievement, psychological profile, and peer group composition for each student enrolled in eight grade in the Canton of St. Gallen for ten consecutive school cohorts (2008 to 2017).

More precisely, the test score data allow us to observe the following for each cohort: composition of secondary school classes (with the school, the track, and the classroom as well as the teacher ID), basic characteristics of each student (birth date, gender, and whether the student is a native German speaker) and student academic achievement on the Stellwerk test (scores for all examined materials). In this analysis, we focus primarily on the scores in math, language (German), and a composite of the two, which are compulsory subjects for all students in all tracks, and standardize them with mean zero and standard deviation one. As we mentioned in the previous section, the classroom composition we observe in the Stellwerk data is the classroom composition that remains fixed over the whole three years of secondary school in all subjects.

Information on the gifted status of students is given by the administrative records of the SPS. In these records, we find information on each child who has had contact with the SPS at any point in his or her school years. They contain the reason of registration, the therapies assigned to the child, the number of visits to the SPS, the date of each visit, and all the notes left by the caseworker about the child. These notes give a very detailed source of information about the child's situation, the topics discussed during each interview, and an overall idea about the diagnosis. Important for our study, we observe the IQ score for many of the children registered at the SPS. Most of the requests to the SPS are made when the child is between six and nine years old, and first contacts with the SPS in primary school often coincide with the time when children start receiving school grades (second semester of second grade).

Four restrictions are imposed on the data, reported in detail in Appendix Table A.1. First, we restrict our data set to students enrolled in the higher track (*Sekundarschule*, 62% of the original sample) and

[5]This section follows Balestra, Eugster, and Liebert (2019) closely, who use an extended version of the same data set.

discard those in the lower track (*Realschule*). The reason for this is that the majority of gifted children (93%) pursue their education in the Sekundarschule. Second, we focus only on students who were actually required to take the Stellwerk test. This leaves out students from special education institutions, for which we do not observe complete classes. Third, we exclude segregated classes that are composed only of students with special needs. Finally, we remove classes and students with missing or implausible values (e.g., test scores exceeding the possible range, classes which are too small or large, or negative age at test). We are left with a final sample of 31,636 students in 1,593 classes from 80 schools.

2.3 Definition of the Key Variables

While remaining aware that intellectual giftedness is a multi-faceted concept whose definition has never been generally agreed upon (Sternberg, Jarvin, and Grigorenko, 2010), we understand intellectual giftedness as an intellectual ability significantly higher than average. Intellectual giftedness is believed to persist as a trait into adult life, with various consequences studied in longitudinal studies of giftedness over the last century (Gottfried et al., 1994). Albeit no generally agreed definition of giftedness for either children or adults has been reached, most school placement decisions and longitudinal studies over the course of individual lives have followed people with IQs in the top two percent of the population (Newman, 2008) – that is, IQ scores above 130 (two standard deviations above the mean). However, there is substantial variation in the threshold used across theories of intelligence, intelligence scales, and individual psychologists.[6]

IQ scores are known to mildly predict academic achievement (Deary et al., 2007; Neisser et al., 1996), since school success is also strongly determined by dedication, motivation, and parental background and investment. Criticism within the psychological community has raised doubts on the validity of the IQ score (see discussion in Sternberg, Jarvin, and Grigorenko, 2010): the IQ score, which maps intelligence unidimensionally, might miss other cognitive dimensions relevant to intelligence, such as emotional intelligence (Mayer et al., 2001; Zeidner et al., 2005), creativity (although much debated, see Make and Plucker, 2018), or domain-specific abilities. Nonetheless, the advantages of measuring cognitive ability with a uniform, normalized IQ scale and tying the definition of "giftedness" to a particular threshold score on the IQ scale are manifold, such as psychometric advantages (easy quantification of intelligence, reliability, internal and test-retest consistency), transparency (the concept of IQ is widely known), predictive accuracy for intelligence in general (e.g., Der, Batty, and Deary, 2009; Leon et al.,

[6]For instance, Silverman (2018) uses the threshold of 120 to identify "mildly gifted," 130 for "moderately gifted," 145 for "highly gifted," 160 for "exceptionally gifted" and 175 for "profoundly gifted." Moreover, the scale must be adapted for students with a native language different than the one of their main environment. Some researchers also use an IQ threshold of 116 for non-native speakers (Card and Giuliano, 2016).

2009), and external validity (measures of IQ exist for all ages, and have been normed across cultures and countries, e.g., Lynn and Meisenberg, 2010; Lynn and Vanhanen, 2012). As a consequence, in the US for instance, individual IQ testing is becoming less commonly used for identification of the gifted and a more holistic identification procedure is preferred.[7]

Our data allow us to mitigate the potential limits of IQ as a unique and reliable measure of giftedness: since the SPS records not only report the children's IQ scores, but also qualitative assessments of cognitive ability (as obtained from the diagnoses and comments of the caseworker), we are able to enhance the IQ scores with qualitative assessments. Qualitative assessments reliably complement quantitative scores, take into account other dimensions of intelligence not assessed by the IQ test, and allow for discarding false positives (Silverman, 2018).[8] They also allow us to differentiate between high achievers and gifted students, as the two attributes do not overlap in many cases. In comparison to the previous literature, we can integrate a richer notion of high intelligence in our analysis.

We then proceed in the following three steps to construct our indicator of giftedness. First, we select one IQ score per child. For many children, we observe different scores, either taken at different points in time, and/or estimated with different intelligence tests. For each child with more than one score, we take the highest score reached by the child.[9] With respect to the chosen intelligence test, we know from the SPS that each child is given the test that suits his or her situation the best.[10] We classify the child as gifted if his or her IQ score is equal to or above 130. We conduct our main analysis with a threshold of 130, and we show that our results are robust when applying more or less restrictive definitions of giftedness. In a second step, we code the written diagnosis of the caseworker – a trained psychologist – and assign the gifted status to the children who are diagnosed as such by the caseworker. In a final step, we remove false positives, i.e., children with a high IQ score but whose assessment does not diagnose high ability. For example, we discard cases in which the child reaches a high IQ score but the caseworker writes that the child had learned how to perform well on the test from his or her siblings. It is important to mention that we focus exclusively on gifted students who are identified prior to secondary school entry. By doing so, we make sure that gifted status does not depend on class composition in secondary school.

In sum, of the 361 students classified as gifted, the majority (256 students) were identified with

[7] See the report from the National Association for Gifted Children, 2015, `www.nagc.org`, and Peters and Matthews (2016).

[8] We are aware of the existence of bias in qualitative assessments, such as documented in McDermott, Watkins, and Rhoad (2014). Unfortunately, we do not observe the caseworker ID so we cannot take bias into account in our analysis.

[9] After discussion with the head of the SPS, we decided to take the highest score since many children need more than one attempt to be fully concentrated during the test. Using either the first IQ administered or the average of all IQ tests performed has no substantial impact on the results.

[10] The available intelligence tests are Snijders-Oomen nonverbal intelligence tests (SON), Kaufman Assessment Battery for Children I and II (K-ABC and K-ABCII), the Wechsler Intelligence Scale for Children (WISC-Hawik), Raven's Progressive Matrices (Raven), Kramer-Test, and Culture Fair Intelligence Test (CFT).

both IQ measure and diagnosis, whereas 107 students were identified only with the diagnosis and seven exclusively with the IQ measure. Finally, we found nine false positives consisting of children who had an IQ score above 130 but clearly were not gifted, according to the psychologist's assessment. Our measure of giftedness has been never used in the literature on ability peer effects and we argue that using a metric based on experts' diagnoses is less prone to measurement error, context-specific factors (e.g., school or class composition), and external influences (e.g., parents or teachers).

2.4 Summary Statistics

Table 1 reports the summary statistics for our final sample. The typical eighth grade class consists of 20 students and there are 0.2 gifted classmates per class. Despite this low prevalence, 18% of all students are exposed to at least one gifted classmate in eight grade. Every second student is a female student, one in ten students is non-native German speaker, and the average age at which students take the Stellwerk test is approximately 15.

The subsample of gifted students is of particular interest, as shown in Table 1, Panel D. Around 1.15% of the sample is identified as gifted, which is around one percentage point lower than the percentage of individuals with $IQ \geq 130$ under the normal curve ($\approx 2.1\%$). This discrepancy is likely explained by the fact that some gifted individuals undergo primary school undetected, thus never entering in contact with the SPS before reaching secondary school. Consequently, as we only observe *identified* gifted students, our measurements of being exposed to gifted students on other students will underestimate the true effect (attenuation bias). In this sense, our findings can be interpreted only as the effect of being exposed to students identified as gifted.

Approximately 31% of the identified gifted students in our sample are female, which is similar in proportion to other European countries.[11] The under-representation of gifted female students, documented in many different contexts (Petersen, 2013), might be explained by teachers' gender bias in referrals for giftedness assessment (which has been documented by Bianco et al., 2011) or by the fact that gifted female students are less likely to be identified by means of IQ testing and standardized tests (Petersen, 2013). It can also be the consequence of a higher prevalence of disruptive behaviors among male students in general, which makes female students pass unnoticed by the teachers and not assessed by the psychologist: in our data set, while 25% of male students have been referred to the SPS, only 16.5% of female students were referred.

[Table 1 here]

[11] Data are from the EASIE 2014 Dataset Cross-Country Report, retrieved from `https://www.european-agency.org/data`.

Appendix Figure A.1 shows the distribution of class size and gifted classmates in absolute and relative terms. The figure indicates that while most classes have no gifted student, nearly all classes exposed to gifted students contain exactly one gifted student. Figure 1 exhibits the distribution of test scores by student type, showing that gifted students perform on average better than regular students – almost a standard deviation better. However, not all gifted students are high achievers and, at the same time, not every high achiever is classified as a gifted student. This finding is confirmed in Appendix Figure A.2, which shows the distribution of gifted students' classroom ranks on the Stellwerk test and reveals that almost 60% of gifted perform in the top five of their classroom, while the rest might perform even in the lowest ranks. This is likely because our definition of gifted transcends the test score dimension, featuring a measure of intellectual ability based on psychological examinations rather than previous achievement.

[Figure 1 here]

3 Empirical Strategy

The aim of this paper is to evaluate the impact of exposure to identified gifted peers on student achievement. Empirically, we estimate the following linear model:

$$Y_{icst} = \alpha + \beta \,\text{Exposure}_{cst} + \gamma X_{icst} + \delta \overline{X}_{(-i)cst} + \varepsilon_{icst} \tag{1}$$

where Y_{icst} is the outcome of interest, such as the math test score of student i in classroom c, school s, and year t. X_{icst} is a vector of individual characteristics that include age at test, an indicator for gender, and an indicator for native German speaker. $\overline{X}_{(-i)cst}$ is a vector of average characteristics of i's class (class size, proportion of female students, proportion of native German speakers, and mean age). The variable of interest is Exposure_{cst}, a measure of exposure to *identified* gifted students in a given class. We parametrize such measure for each student as an indicator being exposed to at least one gifted classmate in grade eight, but the results are consistent to alternative specifications (e.g., the proportion of gifted students per class). The peer spillover parameter is β, which represents the impact of being exposed to a gifted student on i's outcome. The error term ε_{icst} is assumed to consist of two components: a school-by-year fixed effect and an idiosyncratic error term (i.e., $\varepsilon_{icst} = \mu_{st} + e_{icst}$). Finally, standard errors are clustered at the classroom level throughout the paper.

The estimation of the interest parameter β suffers from three main identification problems. First, Manski (1993)'s well-known reflection problem states that all behaviors in a peer group are affected by the behaviors of the other members of the group. Namely, a student simultaneously influences the outcome of the group and the group influences the outcome of the student. We tackle this problem in two ways. First, all variables in Equation (1) are determined before secondary school, including the status of gifted student. This strategy ensures that neither the gifted status nor other individual characteristics are influenced by contemporary class composition. Second, we exclude gifted students from all regressions in order to separate the subjects of our investigation (regular students) and the peers who potentially provide the mechanism for causal effects on these subjects (gifted students). As discussed by Angrist (2014), this distinction eliminates mechanical links between own and peer characteristics, making it easier to isolate variation in peer characteristics that is independent of subjects' own characteristics.

The second main identification problem stems from common unobserved shocks at the group level. These shocks at the class and school levels could tamper with the identification of peer effects of gifted students on their classmates. For instance, the outbreak of an epidemic or the introduction of new pedagogical methodologies for a lesson could impact the overall academic performance of a classroom or a school, which would confound the peer effects estimation if correlated with the proportion of gifted peers. To resolve this issue, we introduce a series of fixed effects that control for unobserved heterogeneity at multiple levels (namely the school-by-year level).

The third identification problem is endogenous peer selection. If individuals are systematically assigned to groups according to a specific characteristic, the researcher cannot determine whether a difference in outcome is a causal peer effect or simply the reason that individuals joined the specific group of study. We take care of this problem by ensuring that gifted students are quasi-randomly assigned to classes at the secondary school level (i.e., identification between classes within the same school-year). As we already mentioned previously, the transfer from primary to secondary school is regulated in such a way that students are assigned to schools based on their place of residence. In each school, students from different places of residence (and consequently different primary schools) are mixed. Importantly, students' psychological profiles are unknown to secondary school administrators such that equity among students is guaranteed and stigma when transitioning between schools is avoided. We exploit this policy rule for identification and formally test the validity of the strategy with two types of balancing tests. First, we test whether individual and group characteristics predict exposure to gifted peers. The aim of this test is to detect potential selection into classrooms. Second, we regress the indicator for gifted status on classroom fixed effects (controlling for school-year fixed effects), which should be jointly insignificant if assignment to class is random with respect to gifted status. We repeat this exercise for teacher fixed effects (for a detailed explanation of this type of approaches see Chetty et al., 2011).

Table 2 shows the results for the first balancing test, where each regression includes class size and school-year fixed effects. None of the coefficients in Table 2 are statistically significant and, in addition, the size of the coefficients is very small. In column 4 we also test for joint significance of the individual and group characteristics. We cannot reject the null hypothesis that the coefficients on gender, native speaker, and age are jointly zero, with a p-value of 0.592. Neither can we reject the null hypothesis that the coefficients on the class-level characteristics are jointly zero (p-value of 0.332). We repeat this randomization test using the proportion of gifted peers instead of the exposure indicator. The results are reported in Appendix Table A.2 and show virtually the same findings as in Table 2: no individual or class characteristic significantly predicts the proportion of gifted classmates.

[Table 2 here]

Although we find no evidence for selection into classrooms according to observable characteristics, we might suspect selection into classrooms based on unobservables. For this reason, in the second set of balancing tests, we test whether either the classroom indicator or the teacher indicator predict gifted status. To do so, we proceed as follows: in a first step, we regress gifted status on the school-by-year fixed effects and retrieve the residuals from such regression. Then, we regress the residuals from the first step on classroom (teacher) fixed effects and test whether these fixed effects are jointly significant. The resulting p-values of these F-tests are 1.00 for the classroom fixed effects and 0.77 for the teacher fixed effects, well above any conventional level of significance.

Our identification strategy relies on variation between classes within school-years.[12] While families can potentially choose their district of residence thereby influencing schooling options for their children, possible selection into schools does not confound the results. In any case, mobility in Switzerland is generally low: approximately 80% of people do not move within five years and moving for school choice alone is likely to be a rare occurrence. In addition, municipalities within the canton of St. Gallen are very homogeneous in terms of demographics, indicating that such strategic behavior is most likely limited. The between-municipality variation in the unemployment rate (coefficient of variation: 0.42), the share of rich (0.19) and poor taxpayers (0.69), and the share with secondary (0.19), higher secondary (0.07), and tertiary education (0.22) is small.[13] Also, we find low geographical variation when examining the prevalence of gifted students across municipalities, as Appendix Figure A.3 shows.

[12] Further evidence on the validity of our identification strategy is presented by Vardardottir (2015). She uses PISA data from secondary schools in Switzerland and shows that track-by-school fixed effects render peer group composition conditionally uncorrelated with a large set of students' characteristics, while track fixed effects and school fixed effects separately do not. The approach we follow in the present paper is even more conservative, as we exploit variation within school-(track)-year.

[13] Data are from Eugster and Parchet (2019).

4 Results

In this section, we present and discuss the results in three parts. First, we introduce the main results on the effect of exposure to gifted students on test scores and perform a comprehensive sensitivity analysis. Second, we proceed to investigate potential heterogeneous effects and mechanisms driving the main results. Third, we examine longer-term outcomes by estimating the effect of exposure to gifted peers on post-compulsory education trajectories.

4.1 Main Results

The main results are presented in Table 3. Specifications 1 to 3 show the effect of exposure to peers identified as gifted on the composite test score (math and language) for all students with different sets of added regression controls, whereas specifications 4 and 5 consider math and language test scores separately. As regression controls, we include student-level controls (column 2) and classroom-level controls (column 3). Crucial for our identification strategy, school-year fixed effects are added to all specifications. Standard errors are clustered at the classroom level.

The estimated coefficients of exposure to identified gifted peers consistently reveal a positive effect on students' own academic performance. The most conservative specification (with student- and classroom-level controls in column 3) indicates that exposure to gifted students raises the achievement of the other students by 8.3% of a standard deviation on average. All effects are statistically significant at the 1% level, and adding covariates does not substantially change the estimates. Note that because we exclude the gifted from the analysis, these results do not reflect the effect of giftedness on the gifted students themselves. Interestingly, class size and peers' characteristics such as the proportion of native speakers and mean age have no influence on achievement. The sole class-level characteristic that has a significant effect is the proportion of female students in the class: having more female classmates increases student achievement in all subjects. This finding corroborates previous studies conducted in the United States (Hoxby, 2000; Whitmore, 2005) and Israel (Lavy and Schlosser, 2011).

[Table 3 here]

As expected, our main specifications in Table 3 confirm the existence of a consistent gender gap in overall score of around 0.2 standard deviations. On average, female students perform less well on the composite test score. However, when we separate the test scores into math and language, we observe that the gender gap in performance is solely driven by math scores, as female students perform on average better than male students in language (column 5).

In order to assess whether the exposure to gifted peers impacts the gender gap, we examine potential gender-related heterogeneous effects in Table 4. We estimate the effect of the exposure to gifted peers including an interaction between exposure and own gender. For both math and overall score, we reject the null hypothesis that the effect is similar for both female students and male students. The estimated coefficients show that not only female students perform on average less well than male students on the composite score, but also that the presence of gifted peers in the classroom exacerbates this difference. This is particularly striking for math scores, for which the positive impact of the exposure to a gifted peer almost completely disappears for female students. Whereas gifted peers have a deterrent effect on female students' performance in math, female students do benefit from gifted peers as much as male students when it comes to performance in language.

[Table 4 here]

Quantifying our results in terms of the gender gap, we find that female students exposed to gifted students are better off than female students without such exposure. This is true for both math and language. However, compared to male students, our results suggest that the exposure to gifted students increases the gender gap in math performance by 27.7%. This increase is due to a disproportional increase in male students' performance. In contrast, we observe a decrease in the language gender gap by 18.4% caused by the exposure to gifted students.

The results are not sensitive to alternative specifications, treatment or outcome definitions, or identification strategies. More specifically, we conduct four sets of robustness checks. First, to make sure that the results are not driven by our definition of exposure to identified gifted students, we conduct the same analysis while defining exposure as the proportion of gifted classmates. Table A.3 in the Appendix shows that our main results hold despite being slightly smaller in magnitude. We also explore potential nonlinearities in the share of gifted students per class by adding quadratic and cubic transformations of the share of gifted students per class. We do not find any nonlinearities in the relationship between average test scores and the share of gifted peers.

Second, we check that our results are not sensitive to a specific definition of giftedness. As mentioned previously, the cutoff in IQ score that determines whether a student is classified as gifted is debated in the literature. For this reason, we conduct the analysis by considering other IQ thresholds. Figure A.4 in the Appendix displays the results when IQs of 135, 130, 125, 120, 115, and 110 are used as thresholds for classifying a child as gifted. In addition, we follow Card and Giuliano (2016) and use the threshold of 116 for non-native speakers (130 for native speakers). We find that all the alternative thresholds are within the 95%-confidence interval of the main estimate (IQ threshold of 130).[14] The results are also robust to using alternative definitions of giftedness, namely gifted students identified only by means of IQ testing (quantitative assessment) and gifted students identified only by means of qualitative assessment.

Third, we check whether the same patterns occur for academic performance in other subjects. Appendix Figure A.5 presents estimates of the main specification for performance in natural sciences (biology, chemistry, and physics) and foreign language (English) as outcomes. These findings reinforce our conclusion that a gender gap in performance for STEM-related fields exists: similar to math, female students perform relatively less well than male students in natural sciences and do not benefit as much from gifted peers for science-related subjects. In contrast, and similar to the pattern we found for language, female students do perform relatively better than male students in foreign languages. The evidence thus suggests that the pattern we uncover in the main analysis for math and (first) language also extends to other STEM and non-STEM subjects.

Fourth, teachers are often mentioned as crucial determinants of students' performance and preferences for particular subjects.[15] To test whether the gender heterogeneity documented in the main specification arises from teachers' own characteristics (either observed or unobserved), we repeat the main analysis by adding teacher fixed-effects. By doing so, we change our strategy from within-school-year, between-classes identification to within-teacher, over-time identification. This change is imposed by the data because no teacher holds two classes in the same year.[16] The results, presented in Appendix Table A.4, show the peer effect estimates net of teachers' time-constant characteristics. We bring evidence that the results do not change when within-teacher estimations are conducted: both point estimates and significance remain very much alike to those presented in the main analysis. From these findings we surmise that teacher characteristics do not explain the gender heterogeneity documented in the main analysis.

One might be concerned that teachers adapt their (instructional) behavior depending on the presence or absence of gifted students in their class. If so, teacher fixed effects would not totally account for teachers' adaptation in behavior, teaching style, or teaching goals induced by the presence of a gifted

[14] To allow for comparability across different IQ thresholds, we keep the estimation sample equal by running all regressions in Appendix Figure A.4 on the sample of students with IQ below 110 (29,905 observations).

[15] For instance, Dee (2007) shows that assignment of children to a same-gender teacher improves their school achievement significantly (in terms of scores, student's engagement with the subject as well as teachers' perception of student performance). Carrell, Page, and West (2010) exploit random assignment to teachers in the U.S. Air Force Academy to document that, while males are not sensitive to the teacher's gender, having a female instructor increases females' performance in STEM and likelihood to choose a STEM-related career path (see also Mansour et al., 2018). Focusing on stereotypes, Carlana (2019) shows that teachers holding implicit negative stereotypes about female students' ability to excel in math have a negative and quantitatively significant influence on female students' performance and career choices. Similar evidence is documented by Alan, Ertac, and Mumcu (2018).

[16] In practice, we specify the error term as $\varepsilon_{icst} = \pi_s + \phi_t + \eta_{cs} + u_{icst}$, where η_{cs} represents the teacher fixed effect (π_s and ϕ_t are school and time fixed effects, respectively).

student. Teachers in the presence of a gifted student might either adapt their teaching style towards the whole classroom, or provide gifted students with personalized teaching. In the first case, teachers must weigh the interests of the gifted students with the interest of the other students and the composition of the classroom, which we control for in our estimation. In the second case, our estimates are immune to the effect of gifted students on themselves. In both cases, teachers do not know *ex ante* the gifted status of their students when they are assigned to a class, which would make any adaptation slow and more costly.

Finally, we investigate whether other students' characteristics could alternatively explain heterogeneous effects of exposure to gifted students in the classroom. As presented in Appendix Table A.5, we document that relatively young students, students in small classrooms, students with a foreign language as mother tongue, and students with school teacher of the same gender do not react differently to gifted peers. In conclusion, we are confident that these channels do not tamper with our main results and that gender is the main driving force behind our findings.

4.2 Heterogeneity and Mechanisms

Provided that we uncover no source of heterogeneity other than the gender-subject result we presented previously, what exactly drives the effect heterogeneity across gender and why does the gender gap in math achievement widen when students are exposed to gifted peers? In this section, we explore several possible explanations for the peculiar gender-subject heterogeneity in the ability peer effect proposed by the scientific literature in economics, psychology, and education science.

The first step is understanding which students are affected the most by the presence of gifted classmates. We look at quantile effects of the exposure to a gifted peer. In Figure 2, we estimate the treatment effect of the exposure to gifted peers for classmates belonging to a given percentile of the test score distribution. The figure plots (unconditional) quantile treatment effects (following Firpo, Fortin, and Lemieux, 2009) and the respective 95%-confidence intervals for different percentiles of the achievement distribution. We find that students in the lower tail of the distribution do not benefit from their gifted peers. The peer effect becomes statistically significant from the fourth decile on, reaching a peak around the eighth decile. The effect heterogeneity over the achievement distribution indicates that high-achieving but non-gifted students react the most to the presence of gifted peers. These findings point towards the conclusions drawn by Card and Giuliano (2014) and Duflo, Dupas, and Kremer (2011), who document that mainstreaming of high ability peers in the classroom has positive peer effects only for children who are academically strong but not for children on the lower end of the test score distribution.

[Figure 2 here]

When we assess gender-specific impacts of being exposed to a gifted peer for children in different quantiles of the test scores distribution, we do not only find a gender penalty of having a gifted peer in math across the whole performance distribution, but also that this gender penalty is significantly larger for female students at the bottom of the score distribution. Appendix Figure A.6 displays this pattern clearly: the gender penalty is distinguishable from zero for the lowest four deciles of the test score distribution. Moreover, as suggested by our previous results, this penalty disappears in language achievement. In summary, both male and female peers in the upper tail of the achievement distribution are positively impacted by their exposure to gifted students. However, low-achieving female students are relatively more negatively impacted than their low-achieving male counterparts in math. These results thus suggest that the widening of the gender gap is driven primarily by decreased performance of low-achieving female students.

A possible – and much discussed – culprit for the reported gender difference in reaction to high ability peers is the fact that male students and female students perceive competition differently (Niederle and Vesterlund, 2011). On the one hand, female students have a higher tendency than male students to experience test pressure, which puts them at disadvantage in standardized tests (e.g., Gneezy, Niederle, and Rustichini, 2003; Montolio and Taberner, 2018; Saygin, forthcoming). On the other hand, female students shy away from competitive environments (e.g., Morin, 2015; Niederle and Vesterlund, 2007) because of lower self-confidence, lower academic self-concept, or lower willingness to compete against male students. This phenomenon is more prominent among high-ability students (Buser, Peter, and Wolter, 2017; Preckel et al., 2008).

A further mechanism that could possibly explain the heterogeneous impact of gifted peers on both male and female students is a "role model" mechanism. Gifted students may be seen as a source of inspiration by their peers and may influence their peers' achievement and motivation positively, especially if the gifted students are themselves exemplary peers. In recent years, the presence of female role models – or the lack thereof – has been shown to have a significant impact on behaviors, preferences and career choices of other female students and women, and it has emerged as a prominent explanation for the gender imbalance in STEM careers (Avilova and Goldin, 2018; Buckles, 2019; Porter and Serra, 2019).[17] In this spirit, we investigate whether female students react differently to gifted female peers than to gifted male peers (and inversely). In addition, we look at whether gifted female or male students

[17] Female role models are defined as "women who can influence role aspirants' achievements, motivation, and goals by acting as behavioral models, representations of the possible, and/or inspirations" (Morgenroth, Ryan, and Peters, 2015, p. 4).

who are exemplary students influence their female and male peers more positively than disruptive gifted students.

In Table 5, we decompose the effect of exposure to gifted peers into exposure to gifted female students and gifted male students and we estimate these effects on female students and male students separately. Note that we lower the IQ threshold to 115 for this analysis (one standard deviation above the mean instead of two), because the low prevalence of gifted students causes a loss of statistical power when using the 130-threshold. However, as presented previously in Figure A.4, the magnitude of the ability spillover is not greatly affected by the choice of the IQ threshold. We bring evidence that male students emulate their gifted peers irrespective of the gender of their gifted peers. As shown in Panel A of Table 5, coefficients of exposure to gifted male students are similar in magnitude to the coefficients of exposure to gifted female students. Interestingly, male students seem to be less affected by gifted peers when it comes to language: the size of the coefficient on language alone is twice as small as the coefficient on math, yet remaining marginally significant. We conduct a simple F-test to show that the two coefficients are not statistically different from each other, which reinforces our conclusion that male students are indifferent with respect to the gender of their high-ability peer.

Turning our attention to female students, we find that female students are far from indifferent to the gender of their gifted classmates. Panel B of Table 5 documents that female students do not react at all when they are exposed to gifted male students; however, they strongly react to the exposure to other gifted female students (around 10% of a standard deviation in overall test score). The only exception is in language, where female students are positively affected by gifted male students as much as by gifted female students. For math, we can conclude that female students are not indifferent to the gender of their high-ability peer, but that they experience a positive influence in the presence of gifted female students. Once again, we stress the fact that the presence of a gifted female student in a classroom is as good as random, allowing to draw causal conclusions from our quasi-experimental setting. Not only is this finding in line with the above-mentioned literature, but it also adds valuable understanding on the importance of same-gender role models.

[Table 5 here]

So far, we have implicitly assumed that all gifted students are "good peers", in the sense that they positively affect and influence their peers. However, this might not be true, especially if gifted students suffer from emotional, behavioral or social problems and disrupt the classroom. Evidence that psychological disorders such as ADHD negatively impact the academic performance, increase the tendency to underachieve and impair executive functioning of gifted individuals as much as non-gifted individuals is well-documented (Antshel et al., 2008, 2007; Brown, Reichel, and Quinlan, 2009; Gomez et al., 2019; Mahone et al., 2002).[18] It is therefore important to distinguish between gifted children who, by their disruptive behavior, are less likely to be seen as exemplary students by their peers, and gifted students who are more likely to foster productivity and generate positive externalities in the classroom.

Having information on the psychological profile of each child sent to the School Psychological Service, we identify gifted students who have been also referred for exhibiting behavioral, emotional or social difficulties (14.4% of all gifted students, see Table 1). Children with such difficulties are usually referred to the SPS for disrupting the classroom or for showing mental or emotional problems. As presented in Table 6, we find that gifted children with emotional issues have a statistically insignificant influence on their peers, despite the negative sign of the coefficient. When looking at the effect of the presence of gifted children without emotional difficulties, the effect is not only significantly positive, it is also 50% larger than the effect of exposure to all gifted children together.

[Table 6 here]

Our evidence suggests that only gifted peers who are not disturbing the classroom environment have positive impact on their peers; disruptive high ability peers at best have no impact and at worst put their peers' performance in jeopardy. This finding can be made even more salient when we analyze which students in the test score distribution are affected. Appendix Figure A.7 shows, on the one hand, that gifted children with emotional issues do not affect their peers except for negatively influencing peers in the 30th and 40th percentile. On the other hand, gifted children without disruptive tendencies positively inspire all their peers towards better performance, with the effect being slightly larger among high achievers – as in our main analysis.

In addition, we find that female students are more sensitive to the presence of disruptive gifted students than male students. Whereas male students are not affected by the presence of gifted peers with behavioral, emotional, or social issues, female students react strongly and negatively. Column 4 of Table 6 reports the interaction effects of being a female student with the exposure to an emotionally unstable gifted peer. Female students perform dramatically less well when they are put in the same class as an

[18]Gifted students who suffer from other mental health disorders are also referred to as "twice-exceptional." There is a debate in the literature in psychology on whether gifted individuals are relatively more likely as the general population to develop psychological disorders (such as ADHD, anxiety, and mood disorders) or depression. Some researchers suggest that gifted individuals are more inclined to developing such disorders and defend an "overexcitability perspective," such as Karpinski et al. (2018) who survey adult members of Mensa (the largest and oldest high IQ society in the world, which is open to people who score at the 98th percentile or higher on a standardized, supervised IQ or other approved intelligence test), whereas others argue that the prevalence of psychological disorders is not higher among the gifted (e.g., Peyre et al., 2016).

emotionally unstable gifted student (around 35% of a standard deviation less than male students), while male students do not react at all to the same gifted peers. However, both male students and female students react positively to the presence of gifted peers who do not have emotional issues, and there is no statistically significant gender difference in this respect.[19]

In sum, our results strongly suggest that both the quality of high-ability spillovers and the sensitivity to disruption play a role in the widening of the gender gap for students exposed to gifted peers. In fact, being exposed to high-ability classmates is no guarantee of improved school performance. Instead, the way high ability peers behave in class is a critical factor for the development of virtuous spillover effects, especially when it comes to spillovers on female students. Our empirical findings are consistent with both the role-model interpretation and previous research on the detrimental effects of disruptive students (e.g., Carrell, Hoekstra, and Kuka, 2018).

4.3 Trajectories after Compulsory Education

Although the results presented here already have far-reaching educational policy consequences, the question remains whether the impact that gifted students have on their peers is limited to academic achievement during the time spent together, or whether it has longer-term consequences beyond that time? Linking our data with administrative data of the educational system allows us to follow the educational career of around 80% the students beyond the period of compulsory schooling. Attrition is almost exclusively due to individuals not having yet completed compulsory education (i.e., the last two school cohorts in the data). Specifically, we are interested in whether the presence of a talented peer in the class influences the peers' educational choices. In Switzerland, tracking into a vocational track or academic track occurs after compulsory education, when students are about 16 years old. A minority (less than twenty percent) of the students opt for the selective academic track (baccalaureate schools) and the majority (two third of a cohort) choses a vocational track, which in the region under observation is usually offered in the form of apprenticeships.[20]

By looking at whether students choose an academic track, a vocational track, or no post-compulsory education at all, we examine whether being exposed to gifted students has longer-run effects. Results are presented in Table 7, divided into three binary outcomes as follows: no post-compulsory education started, vocational track started, and academic track started. In each regression, the reference category is always the other two trajectories combined, in order to avoid conditioning on downstream outcomes.

[19] The same patterns discussed here hold if we perform the analysis for math and language separately.
[20] A detailed overview of the Swiss education system can be found here: http://www.edk.ch/dyn/11586.php.

[Table 7 here]

We find that being exposed to gifted classmates in secondary school significantly increases the likelihood of choosing the academic track. Interestingly, this effect is entirely driven by male students who enter the academic track instead of the vocational one. No effect whatsoever is found for female students. Finally, we find that exposure to gifted peers during secondary school does not change the probability of pursuing any post-compulsory education degree. This result is expected, given our previous findings that low-achieving students remain unaffected by the presence of gifted students.

5 Conclusion

The present study sheds light on the relevance of gifted students and their heterogeneous spillovers effects in the classroom. Heterogeneity is observable in at least three dimensions. First, not all gifted students impact their peers in the same manner but the impact depends on the gender of the gifted student and also whether gifted students show behavioral problems or not. Second, not all peers are affected in the same way but effects differ by gender and ability of the peers and third, peers are not affected in the same way in all subjects. We find that while male students benefit from the presence of gifted peers in all subjects regardless of their gender, female students benefit exclusively from the presence of gifted female students. The nature of our data allows us to test a number of potential mechanisms. We show that neither teachers nor classroom composition are responsible for driving the heterogeneity in the ability spillovers. Instead, we present suggestive evidence consistent with the hypothesis that academic role models and classroom behavior are fundamental in determining the gender gap in math.

In general, we find that gifted students are influential in fostering emulation and impacting positively the academic achievement and the career choices of their peers. They are therefore fundamental forces in the classroom production function that should not be ignored in designing successful educational policies, especially when considering whether gifted students should be segregated in more "elite" schools or pull-out programs. We also show that gifted students are as vulnerable as their classmates to develop disruptive behaviors and psychological issues: gifted students who were diagnosed with socio-emotional problems generate null-to-negative spillovers on their peers. Therefore, further research calls for a better understanding of optimal classroom allocation, possibly balancing the positive spillovers from high-ability students with negative spillovers from disruptive students.

References

Alan, Sule, Seda Ertac, and Ipek Mumcu. 2018. "Gender stereotypes in the classroom and effects on achievement." *Review of Economics and Statistics* 100 (5):876–890.

Anelli, Massimo and Giovanni Peri. 2017. "The effects of high school peers' gender on college major, college performance and income." *Economic Journal* 129 (618):553–602.

Angrist, Joshua. 2014. "The perils of peer effects." *Labour Economics* 30:98–108.

Antshel, Kevin, Stephen Faraone, Katharine Maglione, Alysa Doyle, Ronna Fried, Larry Seidman, and Joseph Biederman. 2008. "Temporal stability of ADHD in the high-IQ population: results from the MGH Longitudinal Family Studies of ADHD." *Journal of the American Academy of Child & Adolescent Psychiatry* 47 (7):817–825.

Antshel, Kevin, Stephen Faraone, Kimberly Stallone, Andrea Nave, Felice Kaufmann, Alysa Doyle, Ronna Fried, Larry Seidman, and Joseph Biederman. 2007. "Is attention deficit hyperactivity disorder a valid diagnosis in the presence of high IQ? Results from the MGH Longitudinal Family Studies of ADHD." *Journal of Child Psychology and Psychiatry* 48 (7):687–694.

Avilova, Tatyana and Claudia Goldin. 2018. "What can UWE do for economics?" *AEA Papers & Proceedings* 108:186–190.

Balestra, Simone, Beatrix Eugster, and Helge Liebert. 2019. "Peers with special needs: effects and policies." Working Paper, University of St. Gallen.

Bianco, Margarita, Bryn Harris, Dorothy Garrison-Wade, and Nancy Leech. 2011. "Gifted girls: Gender bias in gifted referrals." *Roeper Review* 33 (3):170–181.

Black, Sandra, Paul Devereux, and Kjell Salvanes. 2013. "Under pressure? The effect of peers on outcomes of young adults." *Journal of Labor Economics* 31 (1):119–153.

Booij, Adam, Ferry Haan, and Erik Plug. 2016. "Enriching students pays off: Evidence from an individualized gifted and talented program in secondary education." IZA Discussion Paper No. 9757, IZA.

Booij, Adam, Edwin Leuven, and Hessel Oosterbeek. 2017. "Ability peer effects in university: evidence from a randomized experiment." *Review of Economic Studies* 84 (2):547–578.

Brenøe, Anne and Ulf Zölitz. forthcoming. "Exposure to more female peers widens the gender gap in STEM participation." *Journal of Labor Economics* .

Brown, Thomas, Philipp Reichel, and Donald Quinlan. 2009. "Executive function impairments in high IQ adults with ADHD." *Journal of Attention Disorders* 13 (2):161–167.

Buckles, Kasey. 2019. "Fixing the leaky pipeline: strategies for making economics work for women at every stage." *Journal of Economic Perspectives* 33 (1):43–60.

Bui, Sa, Steven Craig, and Scott Imberman. 2014. "Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students." *American Economic Journal: Economic Policy* 6 (3):30–62.

Burke, Mary and Tim Sass. 2013. "Classroom peer effects and student achievement." *Journal of Labor Economics* 31 (1):51–82.

Buser, Thomas, Noemi Peter, and Stefan Wolter. 2017. "Gender, willingness to compete and career choices along the whole ability distribution." IZA Discussion Paper No. 10976, IZA.

Card, David and Laura Giuliano. 2014. "Does gifted education work? For which students?" NBER Working Paper No. 20453, National Bureau of Economic Research.

———. 2016. "Universal screening increases the representation of low-income and minority students in gifted education." *Proceedings of the National Academy of Sciences* 113 (48):13678–13683.

Card, David and Abigail Payne. 2017. "High school choices and the gender gap in STEM." NBER Working Paper No. 23769, National Bureau of Economic Research.

Carlana, Michela. 2019. "Implicit stereotypes: Evidence from teachers' gender bias." *Quarterly Journal of Economics* 134 (3):1163–1224.

Carrell, Scott, Richard Fullerton, and James West. 2009. "Does your cohort matter? Measuring peer effects in college achievement." *Journal of Labor Economics* 27 (3):439–464.

Carrell, Scott, Mark Hoekstra, and Elira Kuka. 2018. "The long-run effects of disruptive peers." *American Economic Review* 108 (11):3377–3415.

Carrell, Scott, Marianne Page, and James West. 2010. "Sex and science: how professor gender perpetuates the gender gap." *Quarterly Journal of Economics* 125 (3):1101–1144.

Carrell, Scott, Bruce Sacerdote, and James West. 2013. "From natural variation to optimal policy? The importance of endogenous peer group formation." *Econometrica* 81 (3):855–882.

Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126 (4):1593–1660.

Cools, Angela, Raquel Fernández, and Eleonora Patacchini. 2019. "Girls, boys, and high achievers." NBER Working Paper No. 25763, National Bureau of Economic Research.

Deary, Ian, Steve Strand, Pauline Smith, and Cres Fernandes. 2007. "Intelligence and educational achievement." *Intelligence* 35 (1):13–21.

Dee, Thomas. 2007. "Teachers and the gender gaps in student achievement." *Journal of Human Resources* 42 (3):528–554.

Der, Geoff, David Batty, and Ian Deary. 2009. "The association between IQ in adolescence and a range of health outcomes at 40 in the 1979 US National Longitudinal Study of Youth." *Intelligence* 37 (6):573–580.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya." *American Economic Review* 101 (5):1739–1774.

Ellison, Glenn and Ashley Swanson. 2010. "The gender gap in secondary school mathematics at high achievement levels: evidence from the American Mathematics Competitions." *Journal of Economic Perspectives* 24 (2):109–128.

Eugster, Beatrix and Raphaël Parchet. 2019. "Culture and taxes." *Journal of Political Economy* 127 (1):296–337.

Firpo, Sergio, Nicole Fortin, and Thomas Lemieux. 2009. "Unconditional quantile regressions." *Econometrica* 77 (3):953–973.

Fryer, Roland and Steven Levitt. 2010. "An empirical analysis of the gender gap in mathematics." *American Economic Journal: Applied Economics* 2 (2):210–240.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in competitive environments: Gender differences." *Quarterly Journal of Economics* 118 (3):1049–1074.

Gomez, Rapson, Vasileios Stavropoulos, Alasdair Vance, and Mark Griffiths. 2019. "Gifted children with ADHD: how are they different from non-gifted children with ADHD?" *International Journal of Mental Health and Addiction* :1–15.

Gottfried, Allen, Adele Eskeles Gottfried, Kay Bathurst, and Diana Wright Guerin. 1994. *Gifted IQ: Early Developmental Aspects – The Fullerton Longitudinal Study*. Springer Science & Business Media.

Hoxby, Caroline. 2000. "Peer effects in the classroom: learning from gender and race variation." NBER Working Paper No. 7867, National Bureau of Economic Research.

Hyde, Janet and Janet Mertz. 2009. "Gender, culture, and mathematics performance." *Proceedings of the National Academy of Sciences* 106 (22):8801–8807.

Karpinski, Ruth, Audrey Kinase Kolb, Nicole Tetreault, and Thomas Borowski. 2018. "High intelligence: A risk factor for psychological and physiological overexcitabilities." *Intelligence* 66:8–23.

Lavy, Victor, Daniele Paserman, and Analia Schlosser. 2011. "Inside the black box of ability peer effects: evidence from variation in the proportion of low achievers in the classroom." *Economic Journal* 122 (559):208–237.

Lavy, Victor and Analia Schlosser. 2011. "Mechanisms and impacts of gender peer effects at school." *American Economic Journal: Applied Economics* 3 (2):1–33.

Lavy, Victor, Olmo Silva, and Felix Weinhardt. 2012. "The good, the bad, and the average: evidence on ability peer effects in schools." *Journal of Labor Economics* 30 (2):367–414.

Leon, David, Debbie Lawlor, Heather Clark, David Batty, and Sally Macintyre. 2009. "The association of childhood intelligence with mortality risk from adolescence to middle age: findings from the Aberdeen Children of the 1950s cohort study." *Intelligence* 37 (6):520–528.

Lynn, Richard and Gerhard Meisenberg. 2010. "National IQs calculated and validated for 108 nations." *Intelligence* 38 (4):353–360.

Lynn, Richard and Tatu Vanhanen. 2012. "National IQs: a review of their educational, cognitive, economic, political, demographic, sociological, epidemiological, geographic and climatic correlates." *Intelligence* 40 (2):226–234.

Mahone, Mark, Kathleen Hagelthorn, Laurie Cutting, Linda Schuerholz, Shelley Pelletier, Christine Rawlins, Harvey Singer, and Martha Denckla. 2002. "Effects of IQ on executive function measures in children with ADHD." *Child Neuropsychology* 8 (1):52–65.

Make, Matthew and Jonathan Plucker. 2018. "Creativity." In *Handbook of Giftedness in Children: Psychoeducational Theory, Research, and Best Practices*, edited by Steven Pfeiffer. Springer International Publishing, 247–270.

Manski, Charles. 1993. "Identification of endogenous social effects: the reflection problem." *Review of Economic Studies* 60 (3):531–542.

Mansour, Hani, Daniel Rees, Bryson Rintala, and Nathan Wozny. 2018. "The effects of professor gender on the post-graduation outcomes of female students." IZA Discussion Paper No. 11820, IZA.

Mayer, John, Donna Perkins, David Caruso, and Peter Salovey. 2001. "Emotional intelligence and giftedness." *Roeper Review* 23 (3):131–137.

McDermott, Paul, Marley Watkins, and Anna Rhoad. 2014. “Whose IQ is it? Assessor bias variance in high-stakes psychological assessment.” *Psychological Assessment* 26 (1):207.214.

Montolio, Daniel and Pere Taberner. 2018. “Gender differences under test pressure and their impact on academic performance: a quasi-experimental design.” IEB Working Paper 2018/21, IEB.

Morgenroth, Thekla, Michelle Ryan, and Kim Peters. 2015. “The motivational theory of role modeling: How role models influence role aspirants’ goals.” *Review of General Psychology* 19 (4):465–483.

Morin, Louis-Philippe. 2015. “Do men and women respond differently to competition? Evidence from a major education reform.” *Journal of Labor Economics* 33 (2):443–491.

Mouganie, Pierre and Yaojing Wang. forthcoming. “High performing peers and female STEM choices in school.” *Journal of Labor Economics* .

Neisser, Ulric, Gwyneth Boodoo, Thomas Bouchard, Wade Boykin, Nathan Brody, Stephen Ceci, Diane Halpern, John Loehlin, Robert Perloff, Robert Sternberg, and Susana Urbina. 1996. “Intelligence: knowns and unknowns.” *American Psychologist* 51 (2):77–101.

Newman, Tina. 2008. “Assessment of giftedness in school-age children using measures of intelligence or cognitive abilities.” In *Handbook of Giftedness in Children: Psychoeducational Theory, Research, and Best Practices*, edited by Steven Pfeiffer. Springer International Publishing, 161–176.

Niederle, Muriel and Lise Vesterlund. 2007. “Do women shy away from competition? Do men compete too much?” *Quarterly Journal of Economics* 122 (3):1067–1101.

———. 2011. “Gender and competition.” *Annual Review of Economics* 3 (1):601–630.

Nosek, Brian A, Frederick L Smyth, Natarajan Sriram, Nicole M Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan, Robin Bergh, Huajian Cai, Karen Gonsalkorale et al. 2009. “National differences in gender–science stereotypes predict national sex differences in science and math achievement.” *Proceedings of the National Academy of Sciences* 106 (26):10593–10597.

Peters, Scott J and Michael Stuart Matthews. 2016. “Gifted education research from the economists’ perspective: What have we learned?” *Journal of Advanced Academics* 27 (2):150–161.

Petersen, Jennifer. 2013. “Gender differences in identification of gifted youth and in gifted program participation: A meta-analysis.” *Contemporary Educational Psychology* 38 (4):342–348.

Peyre, Hugo, Franck Ramus, Maria Melchior, Anne Forhan, Barbara Heude, and Nicolas Gauvrit. 2016. “Emotional, behavioral and social difficulties among high-IQ children during the preschool period: Results of the EDEN mother-child cohort.” *Personality and Individual Differences* 94:366–371.

Pope, Devin and Justin Sydnor. 2010. “Geographic variation in the gender differences in test scores.” *Journal of Economic Perspectives* 24 (2):95–108.

Porter, Catherine and Danila Serra. 2019. “Gender differences in the choice of major: The importance of female role models.” *American Economic Journal: Applied Economics* .

Preckel, Franzis, Thomas Goetz, Reinhard Pekrun, and Michael Kleine. 2008. “Gender differences in gifted and average-ability students: Comparing girls’ and boys’ achievement, self-concept, interest, and motivation in mathematics.” *Gifted Child Quarterly* 52 (2):146–159.

Saygin, Perihan. forthcoming. “Gender bias in standardized tests: evidence from a centralized college admissions system.” *Empirical Economics* .

Silverman, Linda Kreger. 2018. “Assessment of giftedness.” In *Handbook of Giftedness in Children: Psychoeducational Theory, Research, and Best Practices*, edited by Steven Pfeiffer. Springer International Publishing, 183–207.

Sternberg, Robert, Linda Jarvin, and Elena Grigorenko. 2010. *Explorations in Giftedness*. Cambridge University Press.

Vardardottir, Arna. 2015. “The impact of classroom peers in a streaming system.” *Economics of Education Review* 49:110–128.

Whitmore, Diane. 2005. “Resource and peer impacts on girls’ academic achievement: evidence from a randomized experiment.” *AEA Papers & Proceedings* 95 (2):199–203.

Zeidner, Moshe, Inbal Shani-Zinovich, Gerald Matthews, and Richard Roberts. 2005. “Assessing emotional intelligence in gifted and non-gifted high school students: outcomes depend on the measure.” *Intelligence* 33 (4):369–391.

Zölitz, Ulf and Jan Feld. 2019. “The effect of peer gender on major choice in business school.” Department of Economics Working Paper No. 270, University of Zurich.

Tables and Figures

Table 1: Descriptive statistics

	(1) Sample mean	(2) Standard deviation
A. Outcome:		
Test score: Mathematics	0.000	1.000
Test score: Language (German)	0.000	1.000
Test score: Composite	0.000	1.000
B. Exposure to gifted children:		
Gifted student	0.011	0.106
Gifted classmates (in %)	1.145	2.722
Gifted classmates (number)	0.222	0.515
Exposure to gifted classmate	0.184	0.387
C. Classroom characteristics:		
Female	0.523	0.499
Native German speaker	0.912	0.284
Age at test	14.81	0.699
Class size	20.47	3.307
Male teacher	0.650	0.477
D. Characteristics of students identified as gifted ($N = 361$)		
Female	0.313	0.464
With emotional or social problems	0.144	0.351
Referred to SPS by parents	0.105	0.307
Referred to SPS by teachers	0.864	0.342
Within top 5 of their class	0.595	0.491

Notes: Descriptive statistics for the main estimation sample, based on 31,636 students in 1,593 classes from 80 schools. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table 2: Balancing tests

	(1) Exposure to gifted peers	(2) Exposure to gifted peers	(3) Exposure to gifted peers	(4) Exposure to gifted peers
Female	0.010			0.011
	(0.008)			(0.009)
Female classmates (%)	0.116			0.133
	(0.0158)			(0.0158)
Native speaker		0.001		0.002
		(0.007)		(0.0007)
Native speaker classmates (%)		0.080		0.100
		(0.126)		(0.128)
Age at test			0.001	0.001
			(0.002)	(0.002)
Classmates mean age at test			-0.039	-0.045
			(0.027)	(0.027)
Class size FE	Yes	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes	Yes
Observations	31,275	31,275	31,275	31,275

Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the school-year level (level of randomization). Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table 3: Spillovers from gifted classmates

	(1) Composite test score	(2) Composite test score	(3) Composite test score	(4) Math test score	(5) Language test score
Exposure to gifted classmates	0.083***	0.086***	0.083***	0.084***	0.059***
	(0.026)	(0.025)	(0.025)	(0.024)	(0.023)
Female		-0.206***	-0.189***	-0.359***	0.039***
		(0.012)	(0.015)	(0.015)	(0.014)
Native speaker		0.412***	0.412***	0.202***	0.517***
		(0.021)	(0.021)	(0.020)	(0.022)
Age at test		-0.192***	-0.191***	-0.167***	-0.164***
		(0.009)	(0.009)	(0.009)	(0.009)
Female classmates (%)			0.426**	0.390**	0.345**
			(0.172)	(0.164)	(0.149)
Native speaker classmates (%)			-0.127	0.028	-0.252*
			(0.143)	(0.137)	(0.129)
Classmates mean age at test			-0.007	-0.008	-0.003
			(0.038)	(0.039)	(0.032)
Class size			0.003	-0.002	0.007
			(0.009)	(0.009)	(0.008)
School-by-year FE	Yes	Yes	Yes	Yes	Yes
Observations	31,275	31,275	31,275	31,275	31,275

Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table 4: Spillovers from gifted classmates by subject and gender

	(1) Composite test score	(2) Math test score	(3) Language test score
Exposure to gifted classmates	0.117***	0.130***	0.072***
	(0.029)	(0.030)	(0.027)
Female	-0.177***	-0.343***	0.043***
	(0.015)	(0.015)	(0.015)
Exposure * Female	-0.067**	-0.089***	-0.025
	(0.032)	(0.033)	(0.030)
Individual characteristics	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes
Observations	31,275	31,275	31,275

Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. Individual characteristics include gender, native German speaker, and age at test. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table 5: Spillovers by gender of the gifted

	(1) Composite test score	(2) Math test score	(3) Language test score
	A. Boys		
Exposure to gifted boys	0.078***	0.089***	0.044*
	(0.027)	(0.027)	(0.026)
Exposure to gifted girls	0.082**	0.092***	0.049
	(0.035)	(0.032)	(0.033)
F-test for equality of coefficients	*0.925*	*0.955*	*0.908*
	B. Girls		
Exposure to gifted boys	0.039	0.015	0.053*
	(0.029)	(0.028)	(0.027)
Exposure to gifted girls	0.103***	0.121***	0.057**
	(0.030)	(0.031)	(0.026)
F-test for equality of coefficients	*0.121*	*0.011***	*0.913*
Individual characteristics	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. In panel A, $N = 14,281$; in panel B, $N = 16,128$. Individual characteristics include gender, native German speaker, and age at test. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table 6: Exposure to gifted with and without social, behavioral, or emotional difficulties

	(1) Composite test score (all)	(2) Composite test score (boys)	(3) Composite test score (girls)	(4) Composite test score (all)
Exposure to gifted with difficulties	-0.062	-0.015	-0.126*	0.021
	(0.051)	(0.067)	(0.067)	(0.061)
Exposure to gifted without difficulties	0.122***	0.178***	0.060*	0.144***
	(0.026)	(0.032)	(0.034)	(0.030)
Female	-0.190***			-0.178***
	(0.015)			(0.016)
(Exposure to gifted with difficulties)* (Female)				-0.177***
				(0.069)
(Exposure to gifted without difficulties)* (Female)				-0.044
				(0.034)
Individual characteristics	Yes	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes	Yes
Observations	31,275	14,832	16,443	31,275

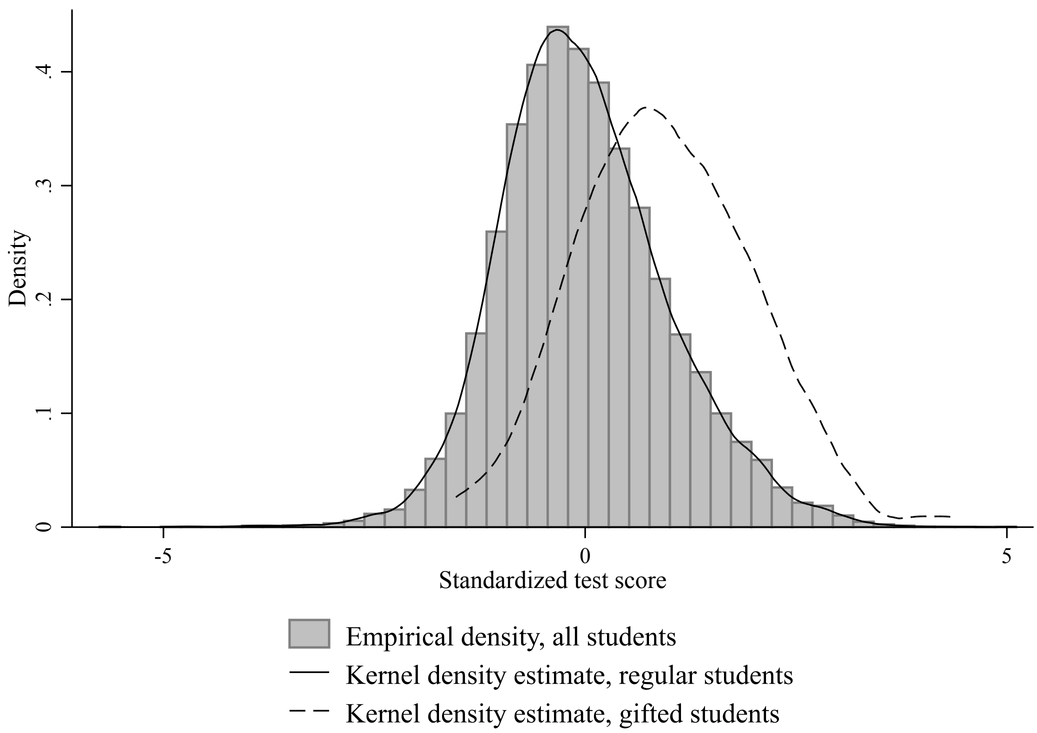
Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. Individual characteristics include gender, native German speaker, and age at test. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table 7: Post-compulsory education trajectories

	(1) No post-compulsory education started	(2) Vocational track started	(3) Academic track started
Panel A. Full sample ($N = 25,407$)			
Exposure to gifted classmates	0.006	-0.024**	0.018*
	(0.006)	(0.011)	(0.011)
Panel B. Boys ($N = 12,133$)			
Exposure to gifted classmates	-0.003	-0.038***	0.041***
	(0.009)	(0.013)	(0.012)
Panel C. Girls ($N = 13,274$)			
Exposure to gifted classmates	0.014	-0.010	-0.004
	(0.010)	(0.015)	(0.015)
Mean outcome	0.157	0.610	0.233
Individual characteristics	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes

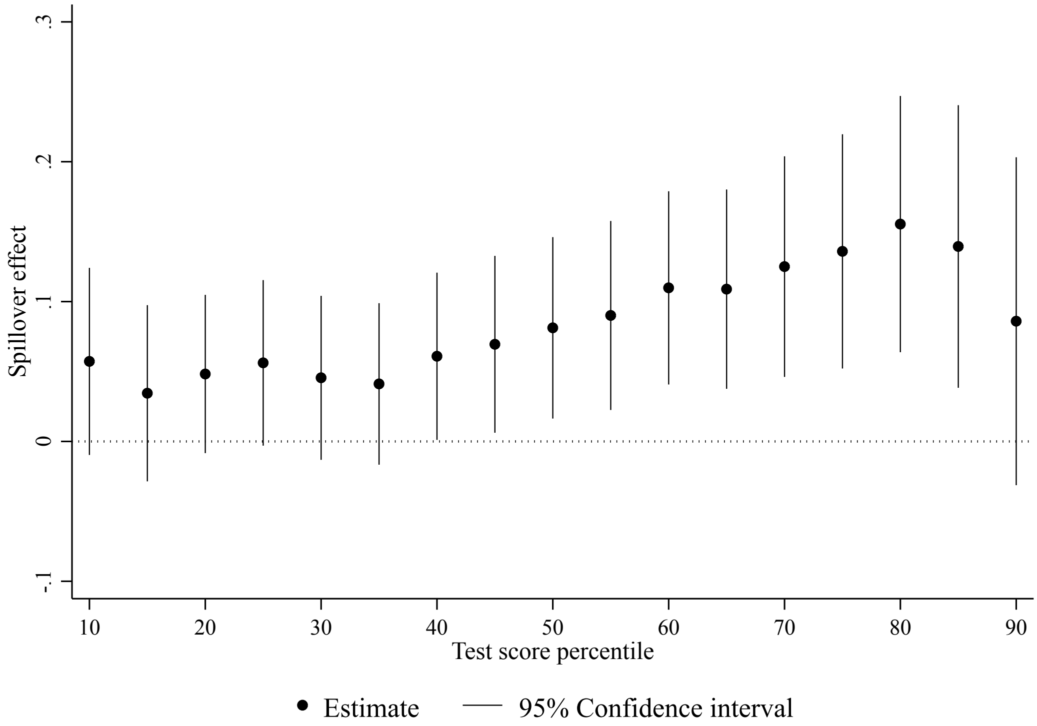
Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. Individual characteristics include gender, native German speaker, and year of birth. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen, the Ministry of Education of the canton of St. Gallen, and the Stellwerk test service provider.

Figure 1: Distribution of test scores



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

Figure 2: Quantile treatment effect of exposure to gifted classmates



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

Appendix: Supplementary Material

Table A.1: Construction of the sample

		Observations
Raw data		33,657
Segregated special schools	–	670
Missing/implausible covariates	–	15
Missing/implausible test scores	–	249
Missing/implausible class size	–	1,087
Final sample		31,636
Gifted students	–	361
Estimation sample		31,275

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

Table A.2: Balancing tests using the proportion of gifted peers as outcome

	(1) Proportion of gifted peers	(2) Proportion of gifted peers	(3) Proportion of gifted peers	(4) Proportion of gifted peers	(5) Proportion of gifted peers
Female	0.0007				0.0008
	(0.0006)				(0.0006)
Female classmates (%)	0.0085				0.0094
	(0.0113)				(0.0113)
Native speaker		-0.0002			-0.0001
		(0.0005)			(0.0005)
Native speaker classmates (%)		0.0045			0.0058
		(0.0086)			(0.0087)
Age at test			0.0001		0.0001
			(0.0002)		(0.0002)
Classmates mean age at test			-0.0021		-0.0024
			(0.0017)		(0.0017)
Class size				0.0003	0.0003
				(0.0006)	(0.0006)
School-by-year FE	Yes	Yes	Yes	Yes	Yes
Observations	31,275	31,275	31,275	31,275	31,275

Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors clustered at the level of randomization (school-year cells) are in parentheses below the coefficients. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table A.3: Sensitivity to the specification of the treatment variable

	(1) Composite test score	(2) Math test score	(3) Language test score	(4) Composite test score	(5) Math test score	(6) Language test score
Proportion of gifted classmates	1.384***	1.198***	1.198***	1.864***	1.753***	1.468***
	(0.309)	(0.297)	(0.321)	(0.374)	(0.363)	(0.400)
	[0.069]	[0.060]	[0.060]	[0.093]	[0.088]	[0.073]
Female	-0.189***	-0.360***	0.039***	-0.178***	-0.346***	0.045***
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)
Proportion * Female				-0.993**	-1.149***	-0.560
				(0.418)	(0.430)	(0.428)
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,275	31,275	31,275	31,275	31,275	31,275

Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. The marginal effect of adding one gifted peer to a class of 20 is shown in brackets, and is defined as the coefficient divided by 20. Individual characteristics include gender, native German speaker, and age at test. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table A.4: Within-teacher identification

	(1) Composite test score	(2) Math test score	(3) Language test score	(4) Composite test score	(5) Math test score	(6) Language test score
Exposure to gifted classmates	0.080***	0.056**	0.084***	0.112***	0.098***	0.096***
	(0.025)	(0.026)	(0.022)	(0.028)	(0.030)	(0.027)
Female	-0.206***	-0.372***	0.022*	-0.194***	-0.357***	0.026**
	(0.012)	(0.013)	(0.012)	(0.013)	(0.014)	(0.013)
Exposure * Female				-0.064**	-0.085***	-0.025
				(0.031)	(0.033)	(0.030)
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,275	31,275	31,275	31,275	31,275	31,275

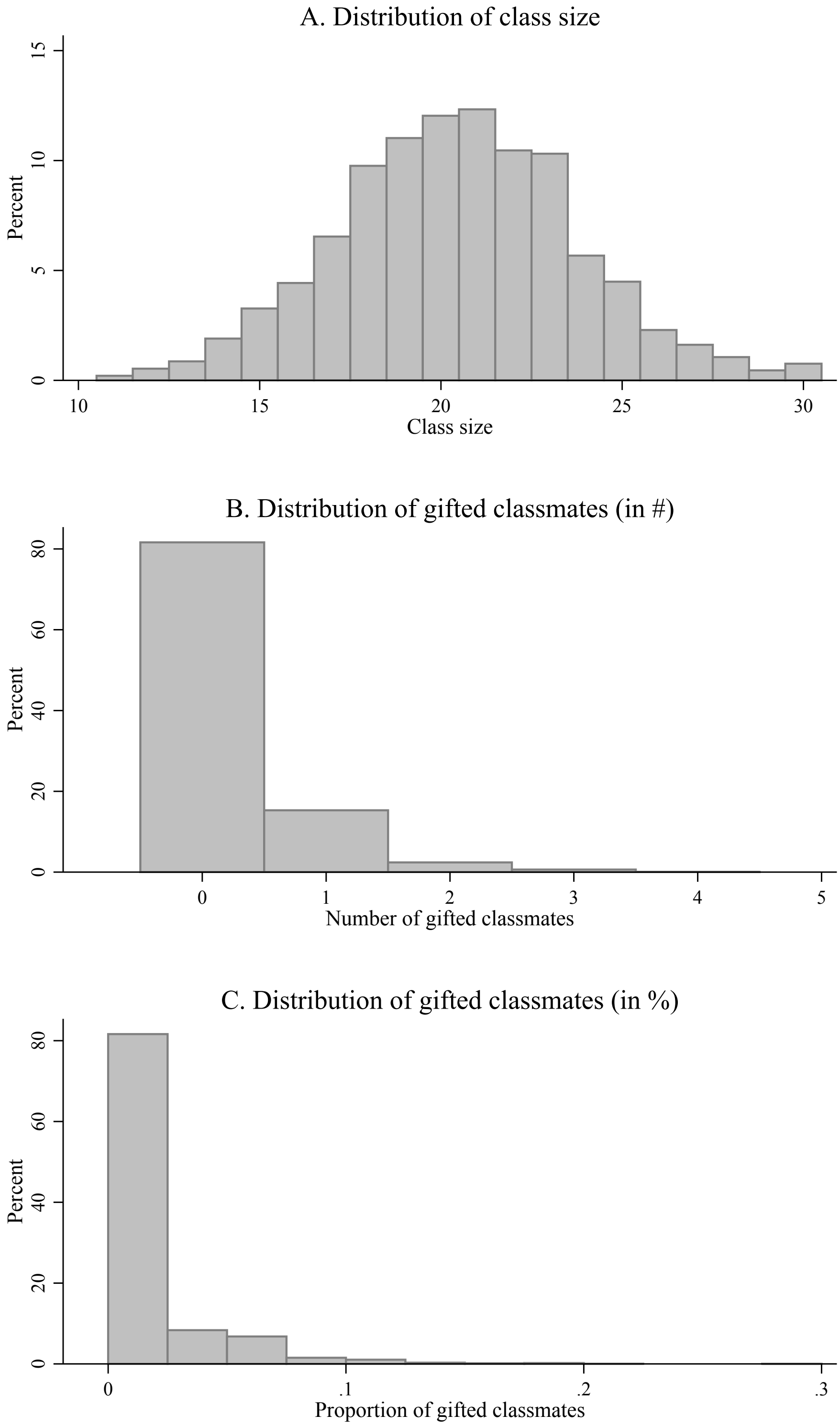
Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. Individual characteristics include gender, native German speaker, and age at test. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Table A.5: Heterogeneity analysis: relative age, native speaker, class size, and teacher's gender

	(1) Composite test score	(2) Composite test score	(3) Composite test score	(4) Composite test score
Exposure to gifted classmates	0.068**	0.045	0.055*	0.071**
	(0.027)	(0.053)	(0.032)	(0.030)
Relative young	-0.133***			
	(0.023)			
(Exposure to gifted classmates)*	0.041			
(Relative young)	(0.030)			
Native speaker		0.405***		
		(0.024)		
(Exposure to gifted classmates)*		0.041		
(Native speaker)		(0.048)		
Small class			-0.001	
			(0.046)	
(Exposure to gifted classmates)*			0.084	
(Small class)			(0.055)	
Student-teacher same gender				-0.007
				(0.014)
(Exposure to gifted classmates)*				0.024
(Student-teacher same gender)				(0.032)
Individual characteristics	Yes	Yes	Yes	Yes
Classroom characteristics	Yes	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes	Yes
Observations	31,275	31,275	31,275	31,275

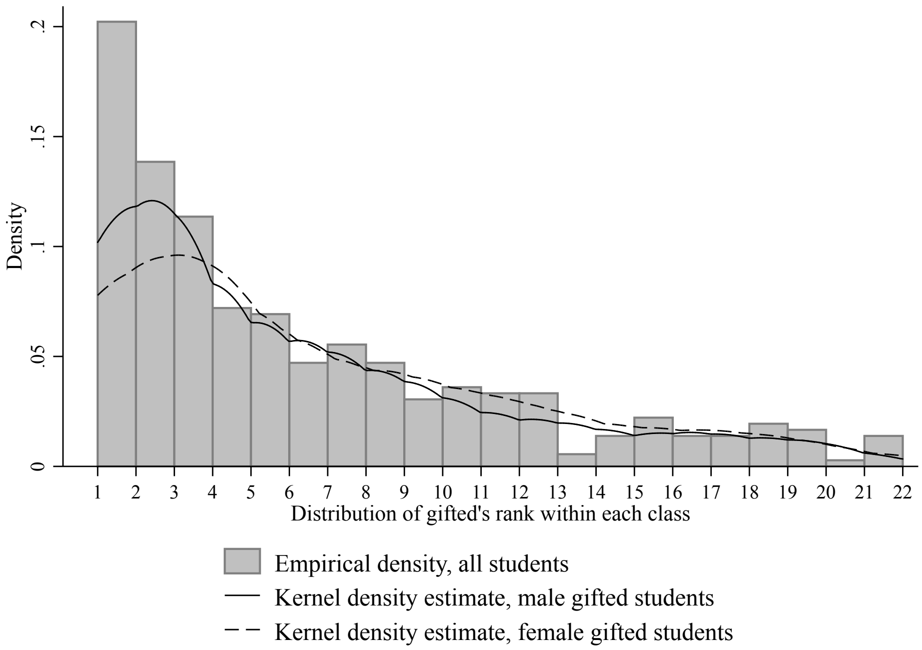
Notes: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Standard errors, shown in parentheses, are clustered at the classroom level. Individual characteristics include gender, native German speaker, and age at test. Classroom characteristics include class size, share of females, share of native German speakers, and average age at test. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Figure A.1: Distribution of class size and gifted classmates



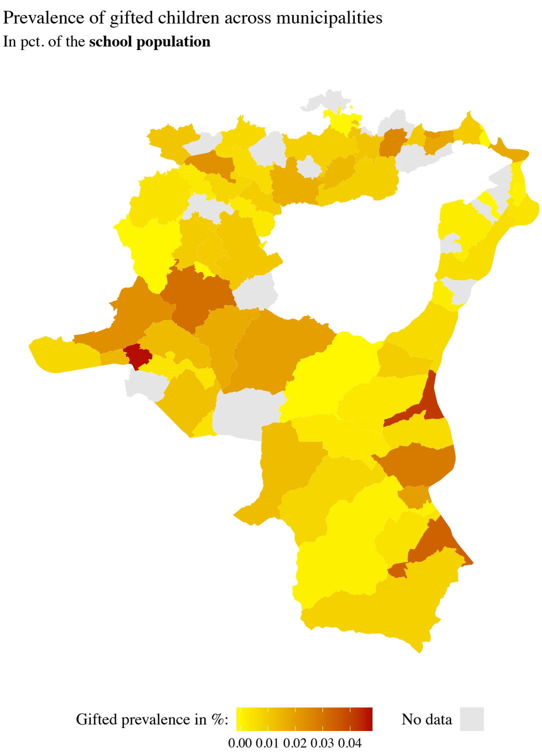
Notes: Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

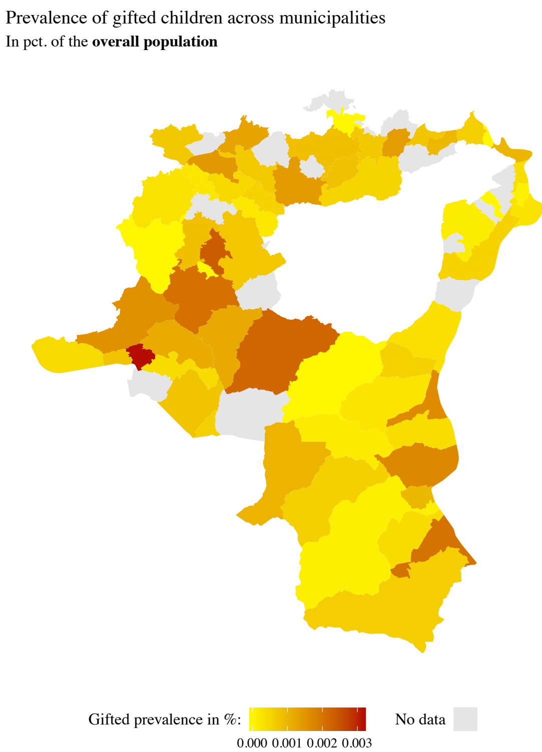
Figure A.2: Distribution of outcome ranks of gifted children in their classroom



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

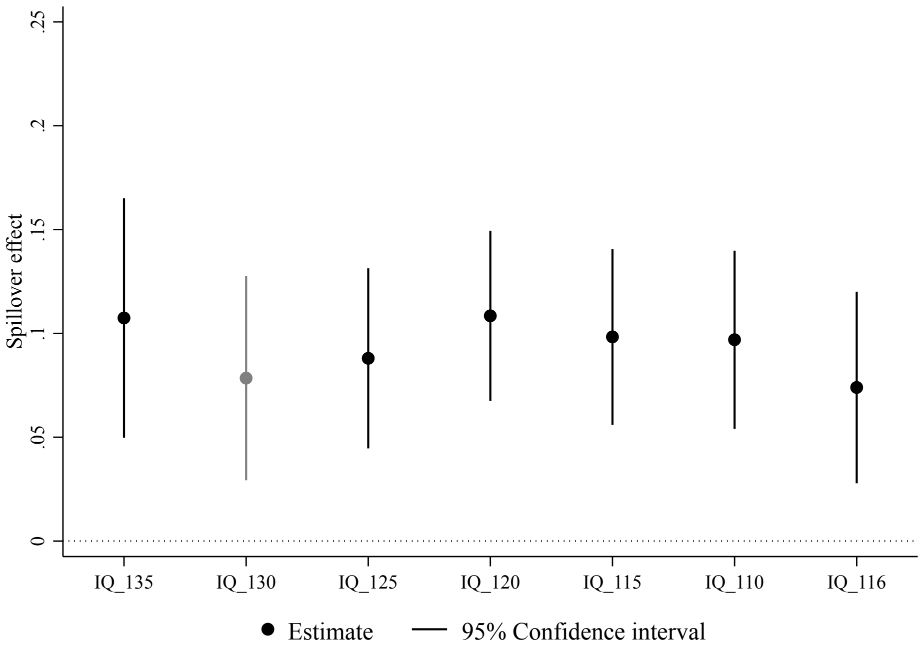
Figure A.3: Prevalence of gifted students by municipality





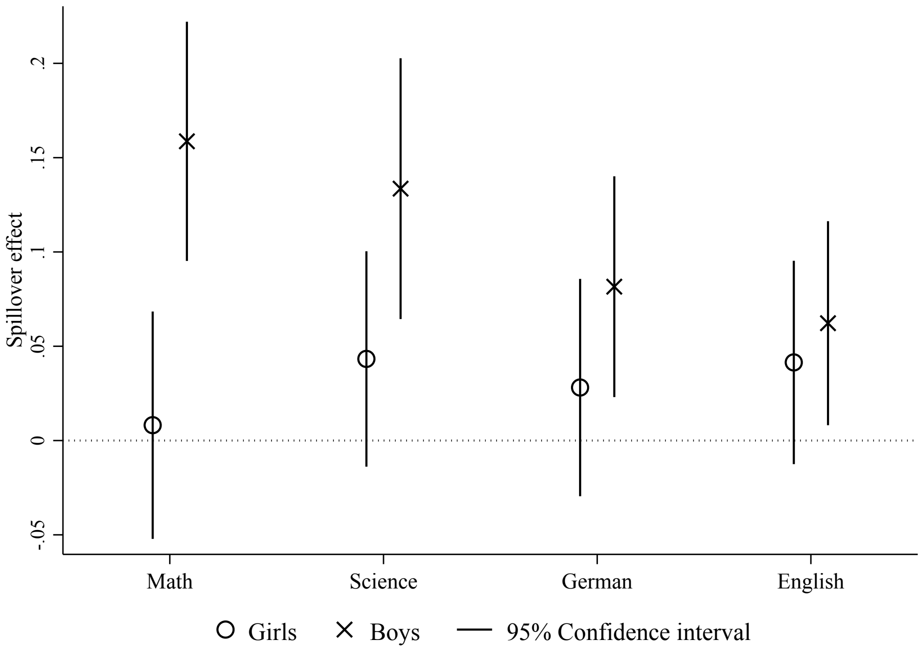
Notes: Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Figure A.4: Sensitivity to the IQ threshold for classifying a student as gifted



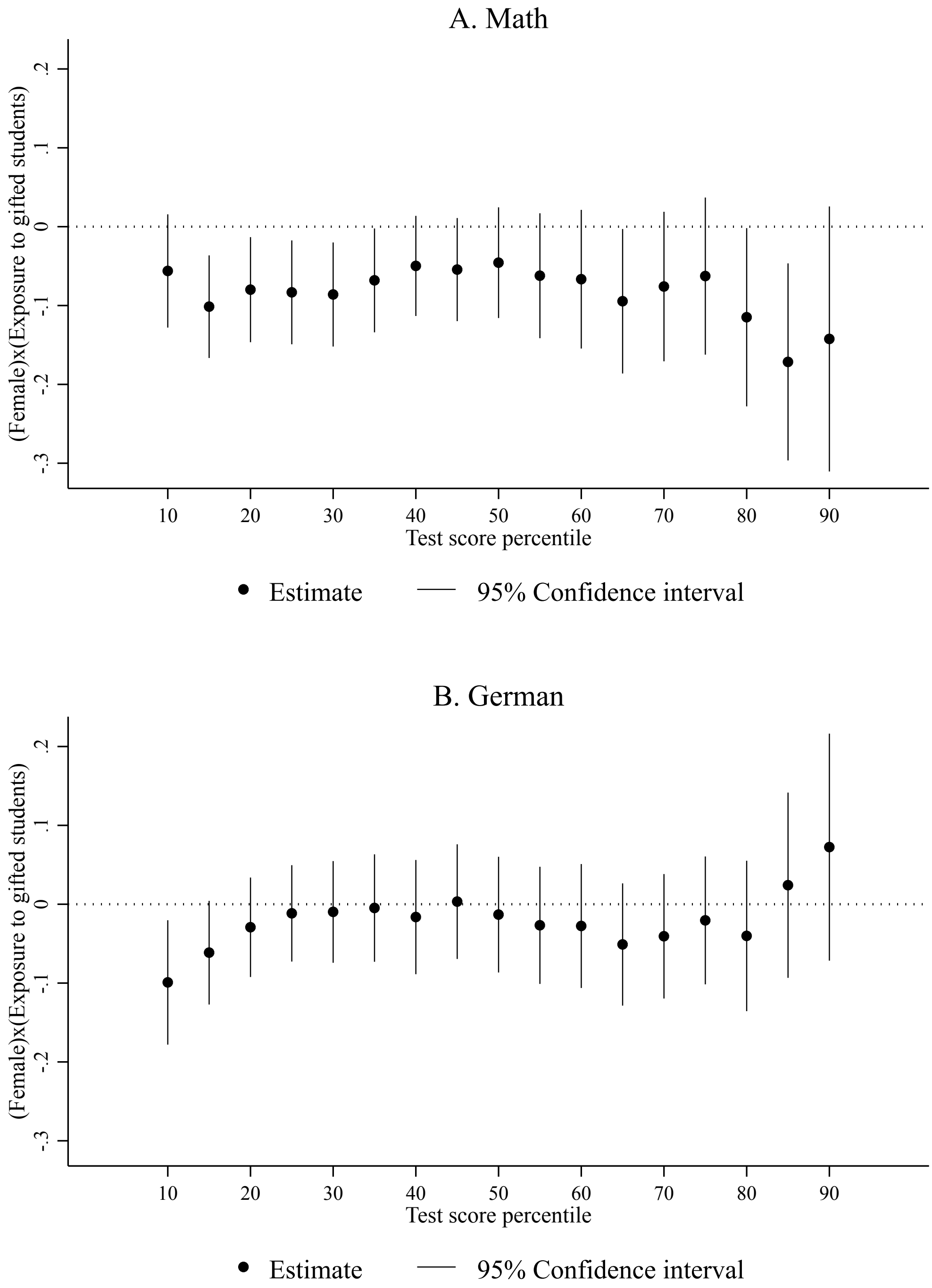
Notes: Results are based on separate regressions on the same estimation sample, which comprises students with IQ below 110 (29,905 observations). The gray estimate is based on the the main IQ threshold of the paper. The last estimate (IQ-116) uses the alternative IQ threshold for non-native speakers of 116 points. Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Figure A.5: Spillovers by school subjects



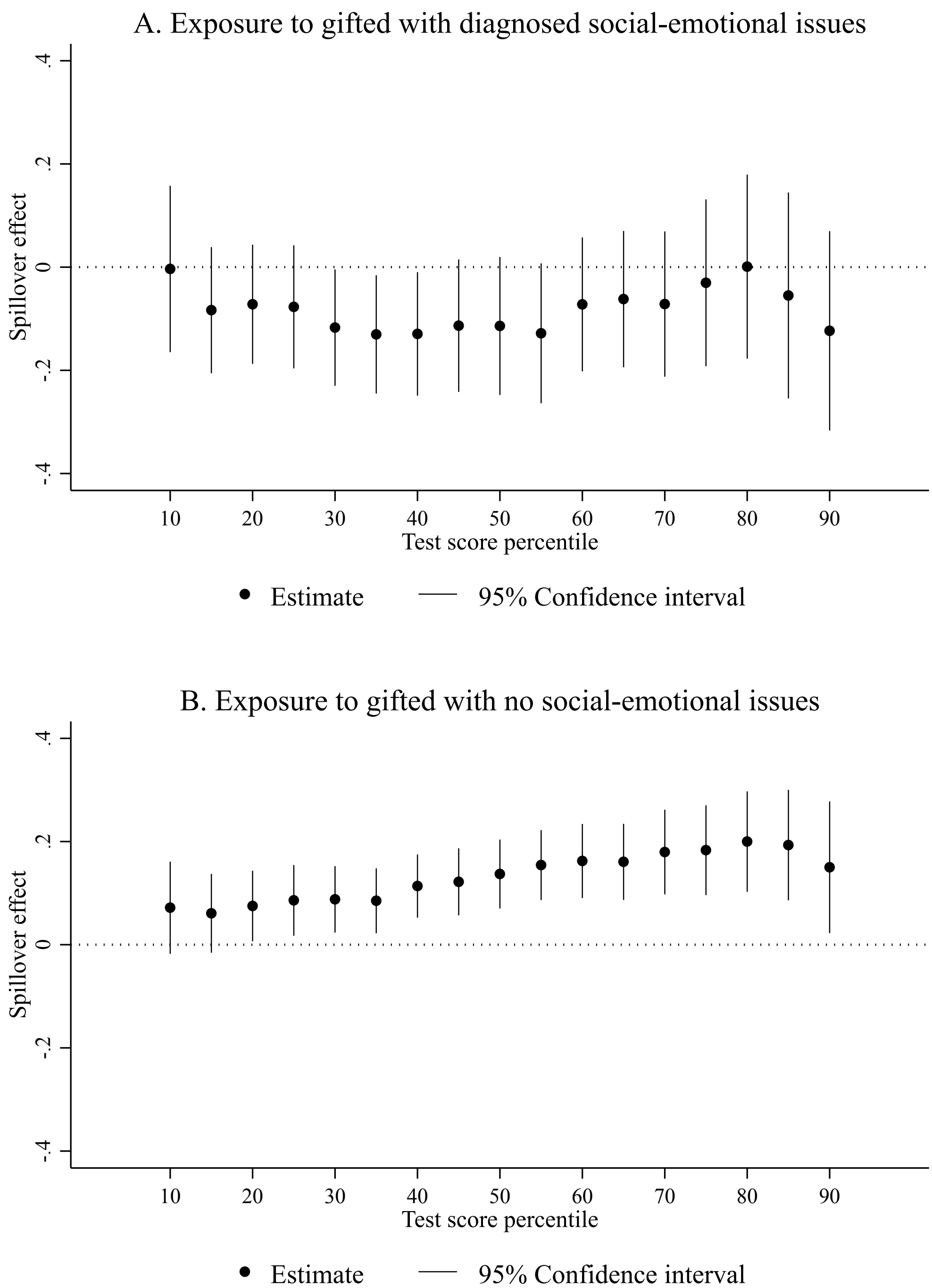
Notes: Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Figure A.6: Quantile treatment effect of exposure to gifted classmates for female students and male students



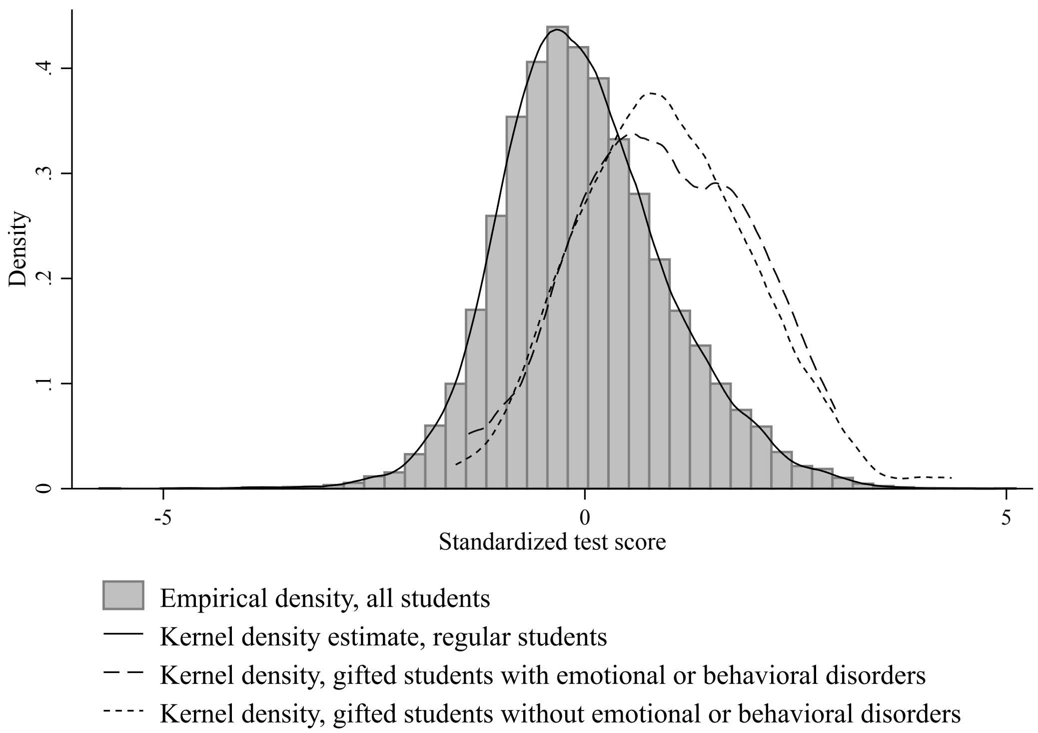
Notes: Data are from the School Psychological Service St. Gallen and the Stellwerk test service provider.

Figure A.7: Quantile treatment effect of exposure to gifted students with and without other concurring diagnosis



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

Figure A.8: Distribution of test scores for gifted children with and without emotional or behavioral disorders



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