DSSR

Discussion Paper No. 138

What do high-achieving graduates bring to nonacademic track high schools?

Yuta Kuroda

December, 2023

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

What do high-achieving graduates bring to nonacademic track high schools?

Yuta Kuroda*

December 6, 2023

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers 19K23217 and 21K13315. The administrative data for high schools used in this study were provided by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). I am grateful to DAIGAKUTSUSHIN Corporation for providing the data on the high school students' acceptance performance used in this study. I would also like to thank Kawaijuku for providing the data on the level of difficulty of the entrance examinations for the universities. I deeply appreciate the helpful comments and suggestions provided by Kentaro Nakajima. I also appreciate the participants of the Rokko Forum at Kobe University in March 2021 and the Kansai Labor Economics Workshop in May 2023. Any errors that remain are my own. The views expressed are the author's own and do not necessarily reflect those of any organizations with which the author is affiliated.

* Graduate School of Economics, Osaka Metropolitan University, 3-3-138, Sugimoto, Sumiyoshi-ku, Osaka, 558-8585 Japan (E-mail: kuroyu0725@gmail.com)

Abstract

In this study, the effects of high-achieving graduates in nonacademic track high schools with low university enrollment rates are investigated. Japanese high schools are stratified, with each high school having nearly fixed tiers of universities to which their graduates advance. Because it is so rare for students from nonacademic track schools to be accepted into top universities, students who are accepted can serve as accidental role models, positively affecting the motivation, aspiration, and knowledge of academic procedures of their lower schoolmates. I have created and used various definitions of nonacademic track schools and high-achieving graduates by using the university acceptance data of almost all high schools in Japan from 2001 to 2021. The results showed that the quasirandom appearance of high-achieving graduates improved the university acceptance outcomes of nonacademic track high schools for subsequent years. Additionally, the appearance of high-achieving graduates is not related to factors such as teacher–student ratios, regional socioeconomic characteristics, or changes in the school district system. Therefore, the observed performance improvement may result from changes in student beliefs and motivations or from the accumulation of school expertise rather than from peer effects or systemic changes.

JEL classifications: I21, I24, D83, R23

Keywords: University choice, College choice, School stratification, Role model, Subjective expectations

1. Introduction

Access to a selective university conveys significant financial and nonfinancial benefits and thus is important in reducing income inequality and improving social mobility (Chetty et al., 2020). However, because advancing to higher education is complex and psychologically taxing, factors other than academic preparation and financial constraints can become barriers (Page and Scott-Clayton, 2016). Schools with many low-income or minority students and fewer college-bound students tend to lack information about application processes and available scholarships (Roderick et al., 2011; Mulhern, 2021). Additionally, because beliefs and aspirations are socially constructed and affected by their communities, students from high schools with historically low academic achievement may have low self-confidence and tend to underestimate themselves (Janzen et al., 2017; Guyon and Huillery, 2021). Moreover, in recent years, university selection by parental socioeconomic status has begun in primary education, and students from more advantaged schools increasingly move to more selective next-level schools (Berkowitz and Hoekstra, 2011; Delaney and Devereux, 2020). Therefore, students who drop out of the academic track at an early stage may go on to less selective universities even when they have high academic potential (Hoxby and Avery, 2013; Hoxby and Turner, 2015).[1]

Input from appropriate role models can address this lack of knowledge, self-confidence, and aspirations and its resulting disparities in higher education. People

[1] Academic and nonacademic tracks are sometimes institutional and sometimes not. Due to student choice and regional characteristics, schools that are not institutionally classified may be substantially separated into academic and nonacademic tracks. In this paper, the distinction between academic and nonacademic tracks is used when the graduates' career paths differ substantially, even if they are not institutionally distinct.

improve their aspirations, efforts, and achievements through social observation, learning about the successes of their role models' with similar characteristics (Sequeira et al., 2016; Janzen et al., 2017). Additionally, in educational investments, students look to their predecessors with similar characteristics and abilities (even if they are not directly related) to determine their career paths based on their peers and organizational norms (Mulhern, 2021; Barrios-Fernández, 2022). Therefore, nonacademic track high school students with potentially high academic ability can choose a more suitable career path through exposure to appropriate role models (Hoxby and Avery, 2013; Carrell and Sacerdote, 2017).

There are many studies on the effects of role models and noteworthy predecessors in developing countries, where student aspirations, beliefs, and information are often scarce (Conley and Udry, 2010; Bernard et al., 2019). In developed countries, the effect of such role models as female entrepreneurs and scientists on women's access to higher education and STEM programs is attracting attention (Porter and Serra, 2020). However, we find it challenging to infer the impact of role models on nonacademic track students from the existing evidence. Because aspirations and beliefs are relative, distinct role model effects are expected between developed and developing countries and between academic and nonacademic tracks. Additionally, role models whose abilities are overly advanced for nonacademic track students may not be effective because people better compare themselves to people with similar characteristics (Clots-Figueras, 2012; Sequeira et al., 2016; Barrios-Fernández, 2022). Additionally, disadvantaged students may incorrectly understand the information or develop erroneous beliefs (Bettinger et al., 2012; Loyalka et al., 2013). Therefore, the role model effect in the nonacademic track, which is a relative disadvantage in developed countries, is still not understood.

Therefore, the focus of this study is on high schools with low university enrollment

rates, and the impact of unexpected student acceptance by prestigious universities (high-achieving graduates, high achievers) on such high schools is analyzed. This study is focused on the university enrollment numbers of high school rather than on individual student enrollment choices. The Japanese high schools analyzed here have been traditionally separated into those catering to selective universities (academic track) and those not catering to university admissions (nonacademic track). Since it is rare for a nonacademic track high school student to be accepted into a prestigious university, such an event can be considered quasirandom. Thus, high achievers can serve as nonexperimental and incidental role models who influence lower-level students and faculty. I found many such rare cases in the university acceptance data for almost all Japanese high schools between 2001 and 2021 and used them in this study, in which an event study framework is applied to analyze changes in the enrollment patterns of graduates in universities from high schools that had not previously sent many (or any) students to top universities following the appearance of high achievers.

As a result, I found that in those schools where high achievers appeared, the average selectivity of the universities to which students were accepted persistently increased. On the other hand, the relevant student university enrollment and application rates did not change. The critical challenge in identifying the impact of high achievers is verifying that the appearance of these high achievers is not related to any other factor. I present a variety of robustness checks to support the identification of causal effects. First, I confirmed that the appearance of high achievers was not associated with changes in teacher–student ratios or changes in the school district system. The main results also remain robust to estimation using samples that exclude small and private schools, which are prone to changes in school structure. Furthermore, the appearance of the high achievers did not

attract potentially talented students from the surrounding high schools, confirming that changes in newly enrolled students do not explain the main results.

At least four mechanisms explain the findings of this study. First, the success of close role models may affect beliefs and efforts by making students aware of a reward for their efforts (Ersoy, 2023). Second, such high achievers may inform students that selective university attendance is a realistic option. When high schools are stratified, their status perceptions form stereotypes, and their students may apply to universities they perceive as appropriate to their high school status (González-Jiménez, 2022). High achievers can overcome such stereotypes and provide a new understanding of selective university enrollment. Third, schools can accumulate further experience and information about selective university admissions, increasing the options of potential applicants. Fourth, teachers' expectations and beliefs about students' abilities can change, leading to teachers presenting students with more challenging career options (Cherng, 2017). Although these alternative mechanisms cannot be perfectly distinguished, I argue that the impact of high achievers on the knowledge, motivation, and beliefs of younger students and faculty members is a primary driver of my results.

This study contributes to the literature by identifying the role model effects on relatively low-performing schools in developed countries. Experiments that attempt to change student higher education behavior through information interventions or role model effects sometimes work (Hoxby and Turner, 2015; Mulhern, 2021) but sometimes fail to exert the desired effect (Kerr et al., 2020; Rizzica, 2020). The inability of low-achieving or disadvantaged students to correctly interpret the information they are provided is one reason why such interventions sometimes fail (Bettinger et al., 2012; Loyalka et al., 2013). Additionally, role model interventions that use overly-different

models from the focal students or present options that exceed student expectations may not be helpful to students in forming career choices (Clots-Figueras, 2012; Sequeira et al., 2016; Barrios-Fernández, 2022). Alternatively, the provision of detailed information about university admissions can often act as a negative surprise that reduces student motivation by modifying their optimism (Loyalka et al., 2013; Kerr et al., 2020). In contrast, high achievers in this study overcome the shortcomings of role model provision and information interventions presented in these previous studies. High achievers have similar characteristics to regular students, and their acceptance into a prestigious university is easier to understand than statistical information and less likely to act as a negative surprise. In this study, I present the potential for more effective interventions by identifying the effects of high achievers.

The findings of this study have important implications for educational discussion, including such topics as ability tracking and school choice. The empirical results of tracking policies have been mixed, with some finding positive impacts on average ((Duflo et al., 2011; Guyon et al., 2012) and others finding further increases in educational inequality (Pekkarinen et al., 2009; Matthewes, 2021). School choice policies have the potential to improve educational achievement through better matching, but they can also promote stratification and academic disparities (Hsieh and Urquiola, 2006; Lavy, 2021). According to my findings, such sorting may have a notable and persistent negative impact. If tracking or school choice places potential high achievers in elite schools, this could increase disparities by reducing the knowledge and confidence of students in disadvantaged schools. In contrast, my results suggest the effectiveness of policies such as the Top Ten Percent Plan, which geographically distributes outstanding students (Black et al., 2023). If the outstanding students assigned to relatively disadvantaged schools

demonstrate a high achiever effect, it could promote social mobility and reduce social inequality.

This study differs from the literature on peer effects in education. In recent years, an increasing number of studies have been focused on the peer effects of extremely high-performing students (e.g., Mouganie and Wang, 2020; Modena et al., 2022). In contrast, this study analyzes the role model effect of successful graduates rather than the peer effect of highly performing classmates. Additionally, the focus of this study is not on the effects of contact with high achievers but rather on the indirect effects, such as awareness of the high achiever's accomplishments or changes in faculty member's beliefs or knowledge. In some ways, this aligns with the literature on regional norms or entrepreneurship rooted in the community (Sorenson, 2017). Compared to peer effects, which primarily affect peers in the classroom, the norms or knowledge generated by high achievers may be more consequential for educational policy because they remain in the school.

The remainder of this paper proceeds as follows. Section 2 details the Japanese education system. Section 3 introduces the data used in the analysis. Section 4 presents the empirical strategy and identifying assumptions. Section 5 presents the main results and describes a series of robustness checks, and Section 6 concludes.

2. Institutional context

2.1. Education system in Japan

The Japanese education system includes six years in elementary school, three years in junior high school, three years in high school, and four years in university. Only elementary and junior high schools are compulsory, but approximately 99% of Japanese students advance to high school. The main career paths after high school include job

placement, vocational school, two-year junior college, and four-year university. The university enrollment rate was approximately 55% in 2020 and is on the rise.

High schools, whether public or private, have entrance exams. If students do not pass the admissions test, they mostly apply for additional openings at lower-ranked high schools. Almost all students are finally assigned to one of the high schools, while very few do not go on to high school or try again next year. Because of the high stakes of the admissions process, some students enter high schools that they do not really desire to attend.

University entrance examinations are primarily high-stakes academic tests (general examination), and admissions are further based on high school grades and interviews.[2] Typically, national universities test many subjects, while private universities often test only the specialized areas in each department. Unlike high school, if students fail the entrance exam of their preferred university, they sometimes attend a preparatory school for a year and retake the exam the following year. Such students who are neither high school nor university students are referred to by the slang term “Ronin” in Japan, and some spend years as Ronin before being accepted into highly selective universities.

2.2. Academic achievement disparity and stratification in high schools

Japanese high schools are broadly classified into two categories: schools that are designed to prepare students for higher education (general courses) and schools that primarily provide vocational training (e.g., industrial, commercial, and agricultural

[2] The examination system differs at the department level as well as at the university level, but for the sake of simplicity, I explain it here only at the university level.

courses).[3] After 2000, the percentage of students in general courses was approximately 72-73%, while the percentage in vocational courses was 18-22%. These courses are not institutionally classified into academic and nonacademic tracks. Graduates from any high school or course can apply to all types of universities. However, if the subjects required by the target university were not included in the high school curriculum, applicants must study independently or at a preparatory school.

Although there are no institutional differences, significant disparities exist among the academic performances of high schools. Even in high schools with regular courses, there is a clear distinction between substantial academic track high schools and nonacademic track high schools. Traditional academic track schools send many students to prestigious universities every year. Most students in academic track schools study intensively for university admission. The school also has expertise in university admissions, and special lectures are sometimes given to prepare students for the entrance exams of prestigious universities. On the other hand, nonacademic track schools have few or no students aiming to go to university, and many students find employment or go on to vocational schools. Nonacademic track schools offer fewer opportunities to obtain information about university admissions and only provide a little preparation for entrance exams.

Thus, in Japan, attending a selective high school is crucial for advancing to a selective university. However, such selective high schools are not evenly distributed throughout the country, with many clustered in urban areas. Additionally, to enroll in such selective high schools, students must invest in education early on, such as private cram schools and tutoring. Therefore, children from nonurban areas and those with parents of low

[3] In addition to general and vocational courses, there are specialized courses such as music, art, sports, and foreign languages.

socioeconomic status face greater difficulty enrolling in selective high schools. Consequently, children with relatively disadvantaged parents sometimes attend nonacademic track high schools and consequently enroll in nonselective universities despite their potentially high academic ability.

2.3. Selectivity and classification of Japanese universities

In 2020, Japan had 86 national universities, 94 public universities, and 615 private universities. National universities tend to be more prestigious and selective than private universities. Public universities are similar to national universities but less academically selective, and they include specialized institutions such as nursing and arts. The dispersion of selectivity among private universities is vast, and while there are some outstanding universities, there are also many less selective universities, some of which even have open admissions.

There are several ways to measure the selectivity and authority of Japanese universities, and one typical indicator is the hensachi score (Araki et al., 2016; Goodman and Oka, 2018). The leading preparatory schools report hensachi scores, which are standardized scores from previous mock exams taken by the students accepted to each university. Hence, the hensachi score is not a standardized version of an entrance exam score but rather an overall indicator of the approximate difficulty and selectivity of admission to a given university or department. The Hensachi score is standardized with a mean of 50 and a standard deviation of 10, and its distribution is roughly normal. Thus, a university with a 50 hensachi score is moderately selective, while a university with a 70 hensachi score is very selective and prestigious. In the Japanese labor market, graduation from a university with a high hensachi score is highly valued and increases the probability

of finding a higher-income job. Therefore, for students choosing a university, the hensachi score is essential.

In addition to hensachi scores, other criteria for university prestige and selectivity include groups of universities such as the Ivy League in the United States and Oxbridge in the United Kingdom. University groups are sometimes classified according to their historical background, while other times, the leading preparatory schools create classifications for reputable universities. For instance, the former imperial universities are among the most highly selective and prestigious Japanese national universities. Additionally, university groups such as MARCH are prestigious private university groups defined by the leading preparatory schools and commonly used in Japan.[4] In this study, such university groups are used to define high achievers.

3. Data

There are four types of data used in this study: the university hensachi scores, the number of students accepted to universities by high school, the characteristics of high schools, and the neighborhood characteristics around the high schools. This study is aimed at identifying the impact of high achievers in nonacademic track high schools on the school's university acceptance performance of the high school in the following year and beyond. However, no institutional criteria exist for either nonacademic track high schools or high achievers. Therefore, the first step is to define nonacademic track high schools and high achievers based on hensachi scores and the acceptance performance by

[4] MARCH is an acronym for the five leading private universities in the Kanto region: Meiji University, Aoyama Gakuin University, Rikkyo University, Chuo University, and Hosei University.

high school. Second, a measure of each high school's university acceptance performance each year is developed based on how many students from each high school were accepted by universities of any selectivity level.

3.1. Hensachi score

To measure a university's selectivity, I use the hensachi scores published by Kawaijuku. Kawaijuku is one of the most well-known preparatory schools in Japan, and these data are the same as those referenced by high school students wishing to enter universities. These data include almost all universities that require academic achievement tests for admissions.[5] Hensachi scores are presented in increments of 2.5 and range from 35 to 72.5. For universities with fewer applicants than openings that have established open admissions, the hensachi score cannot be calculated (and is described as "border-free" in the data), so the hensachi score for such a university is considered 30.[6]

While hensachi scores exist for different departments and entrance exam types, the number of students accepted to universities by high school, which is discussed below, does not distinguish between departments or entrance exam types. Therefore, the average hensachi scores of the first exams (the most primary exam) of all departments is considered the university's total hensachi score. Furthermore, the average of these scores from 2012 to 2021 is a measure of the university's selectivity. In other words, hensachi

[5] The universities where hensachi scores cannot be used include nursing, arts, and sports, where the admission exam consists of an interview or tests practical skills. Appendix Table A1 shows the representativeness and capture rate of the available datasets.

[6] Excluding such universities from the analysis did not cause any change in the main results.

scores are time-invariant variables, forming one unique value for each university. This approach is plausible because the difficulty of admission, prestige, and reputation of each university barely changed during the analysis period, with an autocorrelation coefficient of approximately 0.98 for the hensachi scores.

Figure 1 shows a histogram of hensachi scores by national, public, and private universities. National universities are, on average, highly selective, and admission to most such universities requires a hensachi score of 50, which represents an average or above academic ability. Public universities also tend to have higher average hensachi scores but lower scores than national universities. In contrast, while some private universities have hensachi scores above 60, the overwhelming majority have hensachi scores below 50. The figure also shows that national universities are, on average, more selective and prestigious in Japan, with some prestigious private universities following closely behind.[7]

[7] The exception is medical schools, whether national or private, which are incredibly selective. One limitation of this study is that it cannot distinguish between medical schools and others. However, the cost of admission of such schools is very high, as private medical schools require tuition fees that are several to dozens of times higher than those of other faculties. Additionally, the number of medical schools is small, and the students who apply differ significantly from the general student population. Therefore, this study assumes that successful applicants to medical school are not affected by a role model effect and seriously impact the identification strategy.

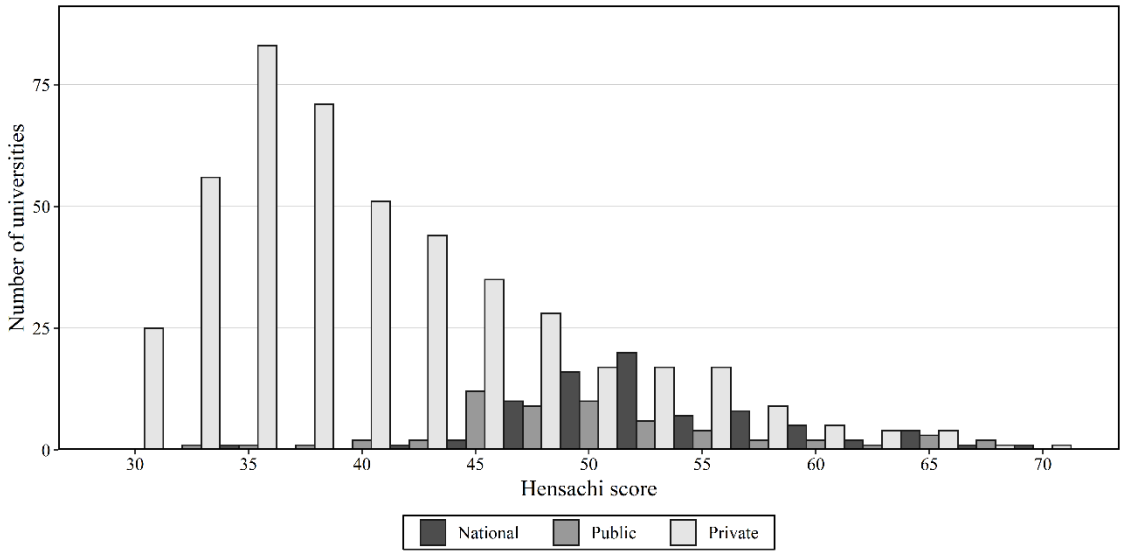


Figure 1. Histogram of hensachi scores by university type

Note: Black, dark gray, and light gray represent national, public, and private universities, respectively.

3.2. Classification of university groups

In this study, three groups of universities are defined based on selectivity, prestige, and reputation as published by preparatory schools. Figure 2 shows each group of universities, their numbers, and their mean hensachi scores. Group A consists of the 11 most selective universities, corresponding to the top 10% academically, and the graduates of these universities are highly valued in the labor market. Group B includes 27 universities, consisting of the 11 universities in Group A and the 16 next most selective universities. Group B includes prestigious national universities and groups of well-known private universities located in Japan's two largest cities, the Kanto and Kansai metropolitan areas. Group C includes 217 institutions and comprises Group B plus all national and public universities. Group C universities are not highly prestigious, but they are moderately selective, requiring at least average academic ability for admission.

University Tiers		Name of University	Number of universities	Average hensachi score
Group A	Very selective, reputable, and prestigious	**(National)** Tokyo, Kyoto, Osaka, Nagoya, Tohoku, Hokkaido, Kyushu, Hitotsubashi, and Tokyo Tech **(Private)** Keio and Waseda	11	62.889
Group B	Selective, reputable, and somewhat prestigious (or higher)	**Group A +** **(National)** Kobe, Tsukuba, YNU, and Hiroshima **(Private)** Meiji, AGU, Rikkyo, Chuo, Hosei, Gakushuin, ICU, Sophia, Kansai, KGU, Doshisha, and Ritsumeikan	27	60.338
Group C	Moderately selective and reputable (or higher)	**Group B +** All national and public universities	217	52.595

Figure 2. Classification of university groups and average hensachi scores

Note: The full name and hensachi score of each university and its rank in the major international university rankings are listed in Appendix Table A2.

3.3. University acceptance data by high school

The number of students accepted from each high school to each university every year is obtained from the "Extra issue of Sunday Mainichi, High School Achievement" published by the Mainichi Shimbun Publishing Inc. This dataset is based on an annual survey performed on four-year universities by the private company DAIGAKUTSUSHIN Corporation. These data report the total number of students accepted rather than the number of enrollments, so if one student is accepted to more than one university, that student is counted more than once. The number of high schools and universities surveyed differs slightly each year, but data are included for approximately 5,000 high schools and 500 universities. The data are highly accurate and representative, covering approximately 99% of high schools and 95% of national and public universities in Japan. Approximately 55% of private universities are included in the survey, as new and defunct private

universities and private universities with very few students are not included.[8] In this study, 21 years of data are available, covering 2001 to 2021.

Combining the number of students accepted by each university with the hensachi score, the average hensachi score of the universities accepting each high school graduate is calculated and used as the university acceptance performance of each high school. For instance, if two students from a given high school are accepted to universities with hensachi scores of 50 and 60, then the placement rate of the high school would be 55. The minimum and maximum hensachi scores of accepting universities are also used to determine whether the high achievers improved the average performance and whether the upper or lower groups increased their performances.

There are several advantages to using the average selectivity of accepted universities. In countries with high university enrollment rates, the value of a university degree in the labor market is declining, and enrollment in more selective universities is valued. Therefore, changes in student behavior that university enrollment rates cannot capture can be identified by considering the selectivity of the universities they accepted. Additionally, because nonacademic track high schools tend to avoid taking comparable academic tests between schools, they often fall short of empirical analysis due to a lack

[8] Appendix A1 compares the number of universities published by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) with the number of universities included in the High School Achievement. Appendix A3 also compares the number of accepted students published by universities with those reported in the "High School Achievement" for national and private universities around the 25th, 50th, 75th, and 99th percentiles of hensachi scores. The table shows that the higher the hensachi score is, the higher the capture rate, but even in universities with lower hensachi scores, more than 90% of the accepted students are captured.

of available data. This study provides new evidence on the nonacademic track by using university acceptance performance across all high schools in the country.

Although these data have certain advantages by capturing the outcomes of nearly all high schools across Japan over time, the potential for investigating the mechanisms underlying the effects found is limited. First, since the number of successful applicants per department or faculty is not available, it is impossible to analyze whether students choose the same department or faculty as high achievers. Additionally, since there is no information on the types of entrance exams and student applications, we cannot identify whether the changes in placement rates are due to changes in ability or changes in career choices. Therefore, other administrative high school data provide a complementary exploration of the underlying mechanisms..

3.4. Defining nonacademic track high schools and high achievers

In this study, the university placement rates by high school, hensachi scores, and university groups are used to define nonacademic track schools and high achievers. The main analysis defines high schools that have not sent students to any Group C university for at least five years as nonacademic schools. Of course, the criteria used for university group and number of years are arbitrary, thus confirming that the results are not sensitive to changes in these criteria. Then, we define high achievers as those students from nonacademic track schools that have been accepted to Group C or higher universities. Stricter criteria for high achievers, i.e., restricting enrollment to group A and B universities, would increase the possibility that high achievers are random but would result in small treatment groups. Because of these tradeoffs, the main analysis presents and compares estimates using admission into Groups A, B, and C to define high achievers.

Figure 3 illustrates how nonacademic track high schools and high achievers are defined and treatment groups and variables are created. In this example, students accepted to Group A universities are defined as high achievers. First, row (A) shows the number of students accepted to Group C universities, where one student was accepted in 2001 and then no students were accepted through 2010. Therefore, this high school has been defined as a nonacademic track school since 2007. Row (B) shows the number of years in which there were no successful applicants in Group C and is shaded to indicate when this value exceeds five years, that is, when it is considered a nonacademic track. Because there may have been high achievers before the data availability period, the data prior to 2005 were used only to define nonacademic track schools and were not used in the estimation.

		Year (2001–2021)																				
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21
Determination of nonacademic high schools	(A) Number of students accepted at Group C universities	1	0	0	0	0	0	0	0	0	0	1	2	0	3	2	0	0	0	0	0	0
	(B) Number of years with no students accepted to Group C universities	NA	0	1	2	3	4	5	6	7	8	9	0	0	1	0	0	1	2	3	4	5
Determination of high achievers	(C) Number of students accepted at Group A universities	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0
	(D) Appearance of high achiever (treatment)	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	(E) Relative time from treatment	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10
University placements of high school	(F) Number of students accepted at universities	6	4	9	3	4	4	11	7	3	8	9	10	8	11	7	9	7	10	5	6	9
	(G) Average hensachi score of the universities accepting students	39	41	41	41	41	39	40	41	42	41	48	42	43	47	44	42	41	43	41	42	44

Figure 3. Conceptual diagram of the definition of nonacademic track high schools and high achievers

Row (C) shows the number of students accepted to Group A universities, which was one in 2011 and two in 2014. Here, students accepted into Group A universities in 2011

from nonacademic track high schools are defined as high achievers. Thus, the treatment variable shown in row (D) takes 1 in 2011, and the relative time since treatment is represented by row (E). There were also successful applicants to Group A universities in 2014, but those who were accepted after the treatment are not considered. Additionally, rarely, the status of nonacademic track schools may reoccur after the appearance of high achievers. Although the second high achiever is also not considered a treatment in the main analysis, I also perform an estimation that excludes schools where the high achiever appears twice as a robustness check. Row (F) shows the number of students who were accepted by any universities, and row (G) shows the average hensachi score of the accepting universities. The key outcomes of this study are shown in row (G), and we analyze how these outcomes change with high achievers.

Only nonacademic track schools are used in the analysis. In other words, academic track schools who constantly send students to prestigious universities are not analyzed. Appendix Table A5 shows the treatment timing and the number of schools in the treatment and control groups for each of the three definitions of high achievers. Approximately 1,100 high schools, or approximately 25%, are considered nonacademic track schools. Additionally, high achievers appear without bias during the analysis period.

There are certain advantages to defining high achievers based on their university placement rates. Although previous studies have primarily defined high ability from test scores, such ability can be recognized within the same classroom but is not easily recognized in other classrooms or grades. In contrast, an achievement measure assessing whether graduates have been accepted to prestigious universities can be used to establish the reputation of a high school, which can persist over time. In Japan, when students achieve remarkable results in academics or sports, they are sometimes congratulated with

signboards or banners placed in noticeable locations around the school. These are intended to enhance the school's reputation and improve student motivation and confidence. Therefore, the appearance of high achievers likely results in a widely recognized and continuously significant impact.

3.5. Other control variables

High school data are taken from the School Basic Survey conducted by the MEXT. The number of graduates, pupils per teacher, and pupils per class are used as control variables. Additionally, the main analysis is complemented with an estimation of the job placement rates and the percentage of university applicants among graduates used as explained variables. Since the School Basic Survey is only available after 2009, the estimation controlling for these variables covers 12 years of analysis, ranging from 2009 to 2020.

In addition to high school data, schools' regional characteristics are also used in the analysis. Population density and the percentage of children (under 18 years old) are obtained from the Vital Statistics by the Ministry of Health, Labor and Welfare. The annual taxable income for each municipality is also obtained from the Ministry of Internal Affairs and Communications. Additionally, official land prices as published by the Ministry of Land, Infrastructure, Transport and Tourism are aggregated by municipality, and the average land price for the municipality in which the school is located is applied. In addition to being used as control variables, these variables are used to confirm that there is no significant change in the regional characteristics before or after the appearance of high achievers.

3.6. Dataset construction and descriptive statistics

A fundamental assumption of this study's identification strategy is that the appearance of high achievers is quasirandom. This assumption is violated when changes in the school system or student quality affect the probability of high achievers. Therefore, only those high schools that have not experienced significant structural changes during the analysis period are used. I exclude high schools that have been newly established, discontinued, consolidated, reorganized, changed ownership, or have become coeducational. Additionally, only schools surveyed at least ten times during the analysis period are used. This sample selection reduces the number of schools, including academic track schools, from 5,249 to 3,944. From this restricted sample, nonacademic track high schools are extracted and used for the main analysis.

Table 1 shows the descriptive statistics by treatment and control groups as well as their significant differences. Panels A-C present the results when high achievers are defined by acceptance into groups A, B, and C, respectively. Descriptive statistics for the treatment group show values from one year before the high achievers appeared, while the results for the control group show median values for the analysis period. Comparing each panel reveals that Group A has the slightest difference between the treatment and control groups, while Group C has the most considerable difference. This is likely due to the stricter definition of high achievers, which further randomizes their occurrence. Conversely, relaxing the definition of high achievers increases the probability that high achievers arise from factors other than coincidence, such as the efforts exerted by each high school.

	Schools with high achievers (one year before treatment)				Schools without high achievers (median for the entire period)				t value
	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	
Panel A: Group A									
Average hensachi score	39.527	2.609	36.094	50.319	38.651	2.893	31.000	51.322	1.895
Number of graduates	192.917	71.995	19.000	355.000	152.925	74.614	10.000	457.000	3.277
Number of teachers	36.639	11.960	9.000	63.000	32.889	13.324	4.000	79.000	1.845
Number of classrooms	15.808	5.200	4.000	22.000	13.333	5.275	3.000	30.000	2.393
Pupils per teacher	16.753	5.872	6.333	34.500	14.343	4.277	1.737	43.267	2.442
Pupils per class	34.742	6.402	11.400	40.917	34.907	5.552	5.500	66.500	-0.130
Ratio of applicants to universities	0.225	0.189	0.020	0.992	0.143	0.099	0.000	0.949	2.599
Population density (/1000)	4.655	5.196	0.026	19.618	1.818	3.301	0.003	20.579	3.255
Ratio of children	0.091	0.012	0.056	0.122	0.093	0.010	0.035	0.184	-1.147
Taxable income (1,000,000 JPY)	3.280	0.737	2.344	6.124	2.949	0.507	2.079	11.117	2.675
Land price (1,000,000 JPY)	0.168	0.215	0.012	1.168	0.083	0.174	0.003	3.043	2.287
Number of universities within 30 km	38.833	45.560	0.000	141.000	16.908	29.735	0.000	144.000	2.868
Number of total high schools		52				1115			
Ratio of private high schools		0.308				0.145			
Panel B: Group B									
Average hensachi score	39.319	3.236	31.072	53.214	37.725	2.969	31.000	51.322	7.129
Number of graduates	183.563	78.554	13.000	445.000	130.863	68.221	10.500	416.000	10.062
Number of teachers	36.363	13.632	5.000	82.000	29.848	12.759	4.000	74.000	6.992
Number of classrooms	15.759	5.150	3.000	27.000	11.835	5.043	3.000	27.000	9.902
Pupils per teacher	16.037	5.616	3.769	54.800	13.550	4.063	1.737	32.897	6.965
Pupils per class	36.297	5.183	11.400	68.833	33.803	5.980	5.500	65.000	5.884
Ratio of applicants to universities	0.199	0.147	0.012	1.000	0.107	0.078	0.000	0.949	10.422
Population density (/1000)	3.101	4.241	0.003	19.618	1.260	2.718	0.004	20.579	7.025
Ratio of children	0.093	0.011	0.052	0.148	0.093	0.010	0.035	0.127	1.079
Taxable income (1,000,000 JPY)	3.198	0.739	2.170	10.232	2.842	0.366	2.101	4.534	8.125
Land price (1,000,000 JPY)	0.138	0.268	0.004	3.043	0.060	0.074	0.003	0.699	5.014
Number of universities within 30 km	29.492	40.048	0.000	144.000	11.283	22.386	0.000	139.000	7.541
Number of total high schools		476				559			
Ratio of private high schools		0.227				0.129			
Panel C: Group C									
Average hensachi score	38.912	3.322	30.500	53.214	37.179	2.884	31.000	45.881	7.688
Number of graduates	169.388	76.710	13.000	445.000	118.154	66.335	10.500	416.000	10.014
Number of teachers	34.935	13.501	5.000	82.000	27.591	12.557	4.000	68.500	7.765
Number of classrooms	14.506	5.302	3.000	30.000	11.276	5.034	3.000	25.000	7.839
Pupils per teacher	15.366	5.270	3.769	54.800	13.357	3.946	2.750	26.770	6.244
Pupils per class	35.844	5.100	11.500	68.833	33.376	6.207	8.833	65.000	5.181
Ratio of applicants to universities	0.178	0.124	0.009	1.000	0.092	0.085	0.000	0.949	11.829
Population density (/1000)	2.087	3.469	0.003	19.618	1.588	3.401	0.004	20.579	1.979
Ratio of children	0.094	0.011	0.040	0.148	0.092	0.010	0.057	0.120	2.680
Taxable income (1,000,000 JPY)	3.025	0.628	2.039	11.267	2.870	0.398	2.206	4.534	4.453
Land price (1,000,000 JPY)	0.100	0.205	0.003	3.043	0.067	0.088	0.007	0.699	3.259
Number of universities within 30 km	19.604	32.700	0.000	144.000	14.420	26.575	0.000	139.000	2.475
Number of total high schools		915				257			
Ratio of private high schools		0.189				0.160			

Table 1. Comparison of the pretreatment and control groups

Overall, the treatment groups show larger school sizes in terms of the number of students and teachers and that of university applicants. However, the quality of education

is not high because of the large number of pupils per teacher and pupils per class. The population density, taxable income, land prices, and the number of surrounding universities indicate that the schools in the treatment group are located in relatively developed urban areas. Therefore, the low cost of attending university and the large number of students and university applicants make it more likely that high achievers appear by chance. I control for these characteristics and use an event study framework to identify the impact of high achievers. I also ensure that these variables are mostly consistent before and after the appearance of high achievers.

4. Identification strategy

4.1. Event study framework

To investigate how high school admissions performance changes due to the appearance of high achievers, I estimate the following event study equations.

$$Outcome_{icrt} = \alpha + \sum_{\substack{k=-12 \\ k\neq -1}}^{14} \beta_k \mathbf{1}[t - High_i = k] + X_{it} + C_{ct} + S_i + Y_t + E_{rt} + \varepsilon_{ict} \quad (1)$$

where $Outcome_{icrt}$ is the outcome of interest in year t for high school i in municipality c covered by the board of education r, and $High_i$ is the year high achievers appeared in high school i. X_{it} and C_{ct} are vectors of time-varying school and regional characteristics that may affect the results, respectively. S_i is a school-fixed effect controlling for time-invariant characteristics at the school level. Y_t is a year-fixed effect controlling for year-specific factors. Ert is the intersection term of the Board of Education and year dummy.[9] Although E_{rt} flexibly controls for unobservable annual

[9] The Japanese Board of Education has authority over prefectures or ordinance-designated cities and can determine policy for the area under its jurisdiction.

institutional changes and trends in the area, it is only used in the main analysis because it absorbs most of the variation arising from the lack of schools in each area. All standard errors are clustered at the school level to address the possibility of the temporal correlation of error terms within schools.

The relative time indicators, $\mathbf{1}[t - High_i = k]$ are equal to 1 when the observed value is $k = -12, \dots, 14$ years from the year in which high achievers appeared in a given high school and zero for all schools with no high achievers. The omitted year, $k = -1$, corresponds to the year before the appearance of high achievers. Thus, $k = 0$ corresponds to the year that high achievers appeared, and $k = 1$ corresponds to after the subsequent year. The coefficient β_k captures the trend of the outcome of interest before and after the appearance of high achievers.

Following previous studies (Bailey and Goodman-Bacon, 2015; Acton, 2022), I also present the grouped DID equation estimates to succinctly summarize the results.

$$\begin{aligned} Outcome_{icrt} = \alpha &+ \beta_1 \mathbf{1}[-5 \le t - High_i \le 1] + \beta_2 \mathbf{1}[t - High_i = 0] \\ &+ \beta_3 \mathbf{1}[1 \le t - High_i \le 2] + \beta_4 \mathbf{1}[3 \le t - High_i \le 5] \\ &+ \beta_5 \mathbf{1}[t - High_i \ge 6] + X_{it} + C_{ct} + S_i + Y_t + E_{rt} + \varepsilon_{ict} \end{aligned} \quad (2)$$

where the coefficients β_1, β_2, β_3, β_4, and β_5 are used to capture changes in the outcome of interest from one to five years before, from the occurrence year, from one to two years after, from three to five years after, and from six or more years after the appearance of high achievers, respectively. The baseline is more than six years before the appearance of high achievers, and β_1 captures whether there was a different trend between the treatment and control groups before the appearance of high achievers. β_2 is the impact of the year that high achievers appeared, and it includes the outcomes of high achievers and the peer effects arising from their effect. The students one to two years after the treatment enrolled before the appearance of high achievers, and for them, the

treatment can be considered an exogenous shock. Thus, β_3 captures the impact of high achievers on the confidence or information available to the lower grades and does not include the impact of their school choice. Because β_4 and β_5 capture the outcome changes for students enrolled after the appearance of high achievers, they also potentially include enrollment characteristic changes due to high achievers. The other control variables and the clustering of standard errors are the same as those in equation (1).

Equations (1) and (2) use a two-way fixed effects (TWFE) approach that compares the outcomes for schools in which high achievers have just appeared with the outcomes for schools in which high achievers have yet to appear, have already appeared, or will never appear. However, recent studies (e.g., Sun and Abraham, 2021) have suggested that when the treatment timing and effect are correlated, TWFE models might provide biased estimates. Therefore, I confirm the robustness of the results using an alternative approach that compares the treated schools with only a untreated control group of schools (Sun and Abraham, 2021).

4.2. Identifying assumptions

The identification strategy used in this paper relies on whether schools without high achievers can be effectively compared to those with high achievers. Confirming this condition requires that the parallel trend assumption is first satisfied, which can be examined by checking the pretreatment changes in equations (1) and (2). Second, there must be no contemporaneous changes regarding high achievers that could affect the outcome of the treatment group. For example, such changes in a school district system and curriculum can improve student capabilities, leading to the appearance of high achievers. While it is impossible to rule out all of these changes, several observable

factors are addressed and examined in this study.

Here, I investigate the validity of the identification strategy by examining whether the occurrence of high achievers in high schools is associated with changes in the characteristics of the relevant schools or regions. Significant correlations would indicated that the assumption that the appearance of high achievers is quasirandom is not satisfied. School characteristics include the number of enrollments, teachers, and classes, and regional characteristics include population density, taxable income, and land prices. I estimated equation (1), controlling only for school and year fixed effects, using each effect as a dependent variable.

The results are shown in Figure 4, with the left side displaying the results for school characteristics and the right displaying the results for regional characteristics. The top, middle, and bottom panels show the estimation results using Groups A, B, and C to define high achievers. Panels (a) and (b) indicate that when acceptance into Group A defines a high achiever, neither school nor area characteristics are significantly related to the appearance of high achievers. Panels (c) and (e) suggest that the number of enrollments and number of teachers tends to increase in schools with high achievers. Additionally, in panels (d) and (f), the schools with high achievers show increases in population density and land prices from those of prior periods. Neither shows any systematic shift immediately before or after the appearance of high achievers, and the trends are consistent throughout the period. One interpretation is that high achievers are more likely in areas with a large or growing population. Therefore, the estimation gas not been confounded by some systematic change associated with the appearance of high achievers. In the main specification, I control for these trends in the estimation by using school and regional characteristics as explanatory variables.

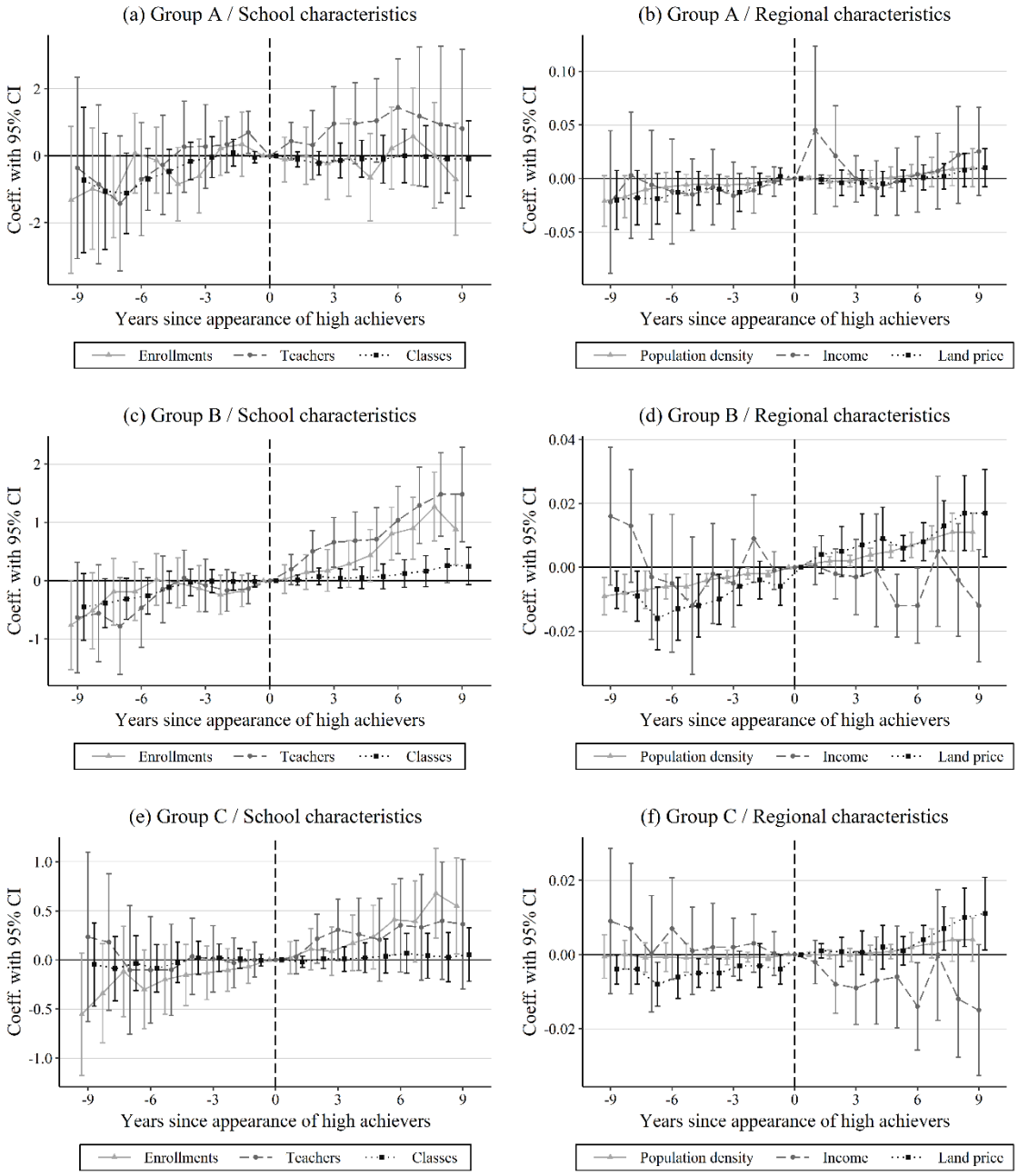


Figure 4. Changes in various factors surrounding the appearance of high achievers

Note: This figure shows the changes in several variables over the nine years before and after the appearance of high achievers. Each figure shows estimates of the coefficients β_k in equation (1), controlling only for year and school fixed effects. All standard errors are clustered at the school level.

5. Results

5.1. Main results

Figure 5 displays the event study estimates detailing how high achievers change the selectivity of the universities accepting students from that high school. Panels (a), (c), and (e) on the left side of the figure show baseline estimates controlling only for school- and year-fixed effects and estimates including additional control for school and regional characteristics. Panels (b), (d), and (f) on the right side of the figure provide estimates controlling for the cross terms of year and board of education fixed effects and estimates that apply the Sun and Abraham (2021) specification. The results are all very similar, confirming that the main results are not sensitive to changes in variables or specifications.

First, the coefficients prior to the occurrence of high achievers were hardly significant, while the acceptance performance increased significantly afterwards. The only slight upward trend appears in the estimates with controls in panel (a), but there is no significant difference between the treatment and control groups from one and those from six years prior. Therefore, a parallel trend, with no gradual improvement in performance due to school efforts or structural changes, can be confirmed.

In all results, the year in which high achievers appear significantly increases the selectivity of the accepting universities, which is unsurprising since such selectivity includes the results of the high achievers themselves. Notably, the coefficients are always positive after the appearance of high achievers. Panels (a) and (b), which show the results of defining high achievers by Group A, demonstrate that the acceptance performance improved after the appearance of high achievers, although the statistical significance is low, and this impact has continued for almost ten years. Panels (c) and (d) show that the

appearance of high achievers increases the selectivity of accepting universities by approximately one point per decade. Panels (e) and (f) indicate that high achievers improve acceptance performance and that their impact gradually strengthens this effect. However, the fact that the estimated results of achievers defined by acceptance into Group C, may include the effects of structural changes because of the significant difference in unobserved characteristics between the treatment and control groups.

Table 2 shows the grouped DID estimates for the main results above using equation (2). Columns (1)-(3), (4)-(6), and (7)-(9) provide the results of the estimation with high achievers defined by groups A, B, and C, respectively. The coefficients from 1-5 years prior to treatment are either not statistically significant or not very significant, again confirming that there is no difference in prior trends between the treatment and control groups. The results show that 1-2 years after high achievers appear, the hensachi scores of accepted universities increase by approximately 0.8 to 1.8 points. This effect results from students who enrolled without anticipating the appearance of high achievers. Thus, it is suggested that high achievers exert a role model effect, which improves the confidence and motivation of younger students or enhances their knowledge regarding university entrance exams. It also indicates that the positive and significant effect persists more than three years after the appearance of the high achiever and intensifies after another six years or longer. This might be due to the development of school reputation, which attracts more qualified enrollments, or it could be due to the gradual accumulation of expertise in university entrance examinations.

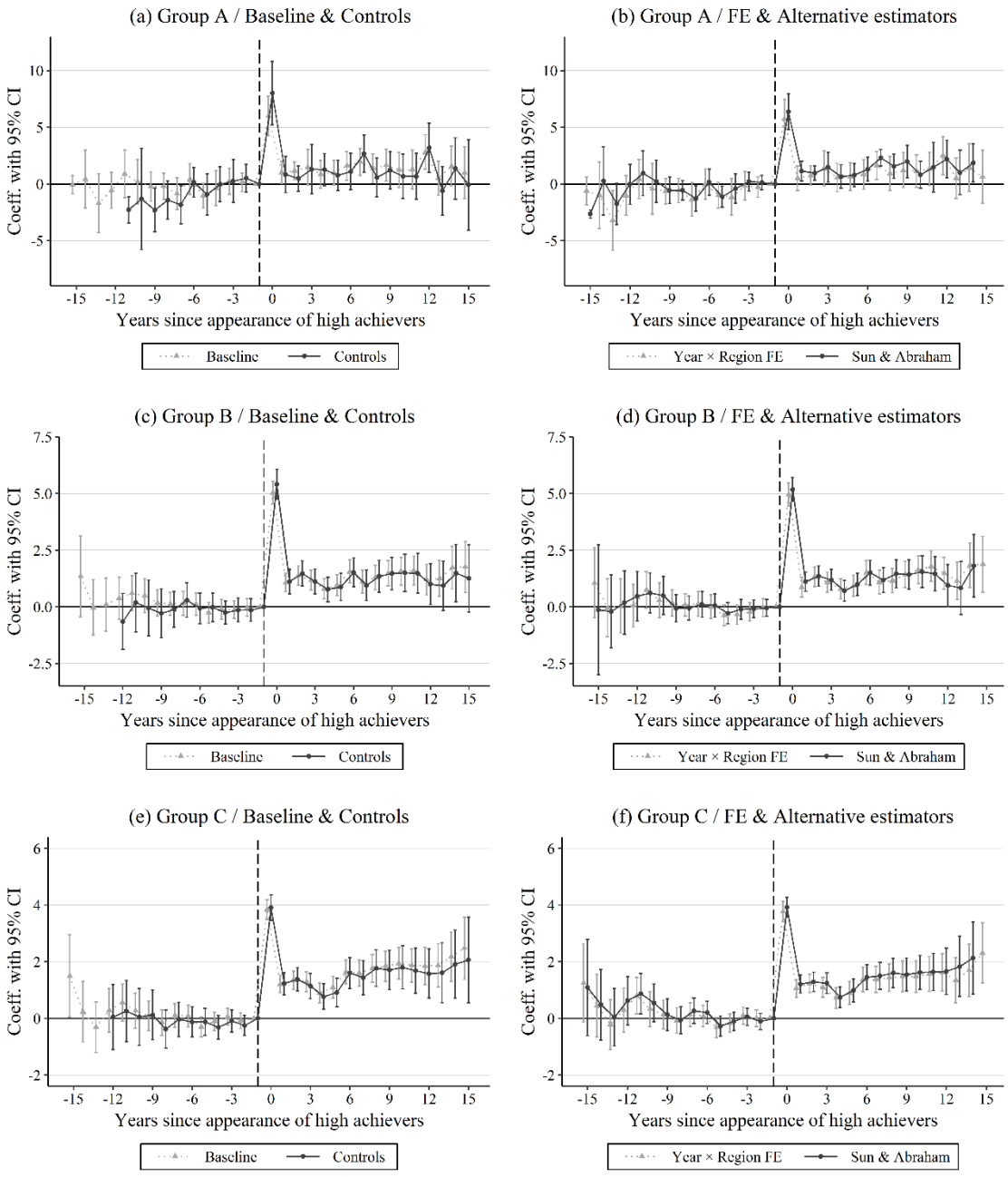


Figure 5. Event study estimates of the impact of high achievers

Note: Each figure shows estimates of the coefficient β_k in equation (1). All standard errors are clustered at the school level. Panels (a) and (b), (c) and (d), and (e) and (f) show the estimation results with high achievers defined by acceptance into groups A, B, and C, respectively.

	Group A			Group B			Group C		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Time since treatment									
1–5 years prior	-0.049	1.12	1.21	-0.319*	-0.102	-0.401	-0.223	-0.088	-0.310
	(0.399)	(0.698)	(0.789)	(0.160)	(0.254)	(0.238)	(0.129)	(0.201)	(0.185)
Year of appearance	6.12***	9.15***	8.42***	4.81***	5.41***	4.39***	3.66***	3.95***	3.21***
	(1.06)	(1.56)	(1.66)	(0.287)	(0.432)	(0.385)	(0.212)	(0.304)	(0.280)
1–2 years after	1.20*	1.78*	1.74	0.991***	1.30***	0.778*	1.09***	1.34***	0.858***
	(0.511)	(0.847)	(0.961)	(0.232)	(0.337)	(0.318)	(0.192)	(0.275)	(0.254)
3–5 years after	1.24*	2.27*	1.82	0.745**	0.929**	0.437	0.895***	0.964**	0.426
	(0.611)	(0.916)	(1.01)	(0.238)	(0.355)	(0.330)	(0.211)	(0.298)	(0.274)
6+ years after	1.67**	2.41**	2.26*	1.17***	1.32**	0.774*	1.52***	1.63***	0.860**
	(0.635)	(0.922)	(0.994)	(0.273)	(0.403)	(0.378)	(0.251)	(0.350)	(0.322)
Controls									
Number of graduates		0.077	0.154		0.058	-0.066		0.118	0.153
		(0.262)	(0.249)		(0.261)	(0.261)		(0.232)	(0.226)
Pupils per teacher		0.029	0.055		-0.029	-0.004		-0.012	0.009
		(0.061)	(0.054)		(0.060)	(0.057)		(0.060)	(0.053)
Pupils per classroom		0.039	0.023		0.072*	0.048		0.056	0.039
		(0.030)	(0.029)		(0.031)	(0.031)		(0.031)	(0.030)
Population density		0.124	-0.266		0.299	-0.150		0.305	-0.161
		(0.265)	(0.307)		(0.270)	(0.297)		(0.251)	(0.292)
Ratio of children		-7.82	0.584		-12.4	-0.422		-18.4	-5.13
		(16.1)	(16.2)		(17.6)	(16.8)		(16.3)	(16.5)
Taxable income		0.279	0.509		0.062	0.167		0.187	0.158
		(0.383)	(0.411)		(0.404)	(0.393)		(0.411)	(0.410)
Land price		0.179	-0.017		0.027	0.277		0.047	0.380
		(0.645)	(0.623)		(0.393)	(0.411)		(0.384)	(0.373)
School FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year×Region FE	NO	NO	YES	NO	NO	YES	NO	NO	YES
Observations	16,280	10,433	10,433	14,354	8,854	8,854	16,649	10,225	10,225
Adjusted R-squared	0.3736	0.3991	0.4841	0.4261	0.4359	0.5274	0.4177	0.4262	0.5118

Table 2. The impact of high achievers on university selectivity toward their school

Note: Robust standard errors are clustered at the school level and appear in parentheses. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively.

5.2. Alternative outcomes

The main results show that the appearance of high achievers improves university acceptance performance over the long term. In this section, through the use of explained variables other than university selectivity, some alternative explanations are rejected and

the mechanism for deriving the results is examined. Table 3, Panels a-c show the estimation results from defining high achievers by acceptance into groups A, B, and C, respectively. Column (1) restates the baseline estimates displayed Table 2 for reference.

Identifying whether high achievers affect only the top tier of students or all students in a school would be interesting. Additionally, one potential concern is that the appearance of high achievers may have prompted the creation of connections (e.g., recommendation quotas) with particular universities.[10] To address these concerns, (2) median, (3) maximum, and (4) minimum values of hensachi scores, rather than the mean value, are used for estimation. The results indicate an increase in not only the maximum value but also the minimum and median values. Thus, the appearance of high achievers increased the average level of university selectivity, confirming that the scenario where a few students increase the average does not explain the results. Column (5) also shows the estimated hensachi score results from using only universities other than those accepting the high achiever, which are almost identical to the main results. Therefore, these results are derived from students who were accepted to different universities than those of the high achievers, and thus the high school-university connection explanation is not persuasive.

Additionally, previous studies have indicated that higher education programs exert different impacts on high school student applications, acceptance, and enrollment (e.g., Phillips and Reber, 2022). To understand the impact of high achievers on student behavior,

[10] Appendix Figure A1 shows the average number of successful applicants by relative time since the appearance of high achievers for the same universities as high achievers and other universities, respectively. This shows that many students have been accepted to the same university as the high achiever after the high achiever's appearance.

columns (6), (7), and (8) present estimation results from using graduate job placement rates, university application rates, and acceptance rates for all universities as the explained variables, respectively. The results suggest that high achievers have little or no significant effect on employment, higher education aspirations, or university acceptance rates. The results in Panel B, column (6), imply that there may have been a slight increase in the number of students who went to university rather than finding a job, but this is of little statistical significance or impact. The findings indicate that high achievers do not increase the number of applicants or to the acceptance rates of universities but only significantly impact the selectivity of the universities to which they are accepted.

	Baseline	Hensachi score				Student behavior		
		Med.	Max.	Min.	Excluding the same university	Job placement rate	Ratio of university applicants	University acceptance rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Group A								
1–5 years prior	1.12	1.67*	-0.574	1.14	1.16	0.040**	-0.023*	-0.004
	(0.698)	(0.734)	(1.42)	(0.846)	(0.699)	(0.014)	(0.010)	(0.006)
Year of appearance	9.15***	8.31***	18.7***	4.43*	5.32***	0.023	-0.029	0.005
	(1.56)	(1.84)	(1.75)	(1.95)	(1.53)	(0.023)	(0.022)	(0.022)
1–2 years after	1.78*	1.60*	1.75	1.83*	1.80*	0.019	-0.011	0.010
	(0.847)	(0.813)	(1.81)	(0.900)	(0.815)	(0.019)	(0.014)	(0.011)
3–5 years after	2.27*	2.45*	2.40	1.86	1.92*	0.024	-0.017	0.007
	(0.916)	(0.955)	(1.91)	(1.02)	(0.919)	(0.020)	(0.014)	(0.012)
6+ years after	2.41**	2.62**	1.67	1.94	2.36*	0.019	-0.017	0.012
	(0.922)	(0.982)	(1.95)	(1.05)	(0.933)	(0.023)	(0.017)	(0.016)
Observations	10,433	10,433	10,433	10,433	10,431	11,768	11,768	11,768
Panel B: Group B								
1–5 years prior	-0.102	-0.105	-0.935*	0.519*	-0.182	-0.006	-0.003	-0.009*
	(0.254)	(0.289)	(0.437)	(0.222)	(0.245)	(0.006)	(0.004)	(0.005)
Year of appearance	5.41***	4.31***	13.3***	1.66***	1.59***	-0.011	-0.004	0.012
	(0.432)	(0.511)	(0.515)	(0.431)	(0.373)	(0.007)	(0.006)	(0.007)
1–2 years after	1.30***	1.13**	2.46***	0.780**	0.920**	-0.018*	0.0005	-0.003
	(0.337)	(0.377)	(0.589)	(0.288)	(0.323)	(0.008)	(0.005)	(0.005)
3–5 years after	0.929**	0.804*	1.71**	0.596	0.536	-0.019*	-0.002	-0.005
	(0.355)	(0.403)	(0.603)	(0.311)	(0.338)	(0.008)	(0.006)	(0.006)
6+ years after	1.32**	1.28**	2.06**	0.961**	0.994*	-0.021*	-0.004	-0.002
	(0.403)	(0.456)	(0.663)	(0.362)	(0.388)	(0.010)	(0.006)	(0.007)
Observations	8,854	8,854	8,854	8,854	8,839	10,085	10,085	10,085
Panel C: Group C								
1–5 years prior	-0.088	-0.101	-0.940**	0.460**	-0.196	-0.002	-0.002	-0.006*
	(0.201)	(0.232)	(0.343)	(0.170)	(0.195)	(0.004)	(0.003)	(0.003)
Year of appearance	3.95***	3.08***	9.11***	1.60***	0.551*	-0.004	0.001	0.011*
	(0.304)	(0.352)	(0.448)	(0.303)	(0.268)	(0.006)	(0.004)	(0.004)
1–2 years after	1.34***	1.10***	2.27***	0.874***	0.847**	-0.005	-0.0001	-0.001
	(0.275)	(0.315)	(0.458)	(0.243)	(0.263)	(0.006)	(0.004)	(0.004)
3–5 years after	0.964**	0.756*	1.42**	0.812**	0.513	-0.006	-0.002	-0.003
	(0.298)	(0.345)	(0.489)	(0.272)	(0.286)	(0.007)	(0.004)	(0.005)
6+ years after	1.63***	1.48***	2.01***	1.31***	1.17***	-0.003	-0.005	0.001
	(0.350)	(0.403)	(0.553)	(0.332)	(0.342)	(0.008)	(0.005)	(0.005)
Observations	10,225	10,225	10,225	10,225	10,160	11,360	11,360	11,360
School characteristics	YES	YES	YES	YES	YES	YES	YES	YES
City characteristics	YES	YES	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Year×Region FE	NO	NO	NO	NO	NO	NO	NO	NO

Table 3. Estimation results using different explained variables

Note: Robust standard errors are clustered at the school level and appear in parentheses. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively.

5.3. Robustness checks

Table 4 provides the results of a series of robustness checks using sample changes. Column (2) provides the results of using all high schools, including reorganized and newly established schools, to ensure that the results are not derived by arbitrary sample selection. Conversely, Columns (3)-(5) report the results of a stricter sample selection. Column (3) presents results excluding schools with large fluctuations in enrollment, thus more strictly eliminating the possibility that the school structure had changed. Column (4) shows the results after excluding small schools with fewer than 100 students, where outcomes can fluctuate. Column (5) presents the results excluding private schools and integrated middle and high schools, which are more likely to involve variance in the input factors. These results were generally consistent with the main results, confirming that the results are not derived from the sample characteristics.

The main analysis defines nonacademic track schools as high schools that have not sent graduates to universities included in Group C for over five years. However, as seen in Appendix Figure A1, some schools had students who met the definition of high achievers more than six years before the appearance of high achievers in that school. Such schools may be inherently more effective than schools in the control group, which could bias the results. Therefore, we performed estimations using samples that excluded high schools with students who had previously been accepted to the same university group as the high achievers. The results shown in Column (6) confirm the high achiever effect even when excluding such schools. Column (7) also presents the results of excluding schools in which two or more successful applicants simultaneously met the definition of high achievers. These estimates make it more likely that the sample comprises exclusively nonacademic track schools and confirm that, even so, the results are consistent with the

main results.

	Baseline	All schools	Stable enrollment schools	Excluding small schools	Excluding upper secondary & private schools	No history of high achievers	Only one high achiever at a time
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Group A							
1–5 years prior	1.12	0.118	1.61	1.12	1.12	1.21	1.21
	(0.698)	(0.745)	(1.02)	(0.697)	(0.699)	(0.754)	(0.842)
Year of appearance	9.15***	8.66***	11.0***	9.15***	9.15***	9.85***	9.39***
	(1.56)	(1.40)	(2.13)	(1.56)	(1.56)	(1.76)	(1.87)
1–2 years after	1.78*	1.38	2.70*	1.80*	1.80*	1.93*	1.58
	(0.847)	(0.840)	(1.14)	(0.842)	(0.843)	(0.953)	(1.01)
3–5 years after	2.27*	1.87*	2.99*	2.27*	2.28*	2.81**	2.20*
	(0.916)	(0.852)	(1.27)	(0.913)	(0.917)	(1.01)	(1.09)
6+ years after	2.41**	2.30*	3.31**	2.42**	2.43**	2.84**	2.17*
	(0.922)	(0.923)	(1.26)	(0.918)	(0.922)	(0.998)	(1.07)
Observations	10,433	12,538	5,942	10,197	10,254	10,381	10,355
Panel B: Group B							
1–5 years prior	-0.102	-0.140	-0.326	-0.160	-0.123	0.299	-0.312
	(0.254)	(0.234)	(0.300)	(0.249)	(0.256)	(0.310)	(0.319)
Year of appearance	5.41***	5.54***	4.90***	5.36***	5.45***	6.51***	4.57***
	(0.432)	(0.394)	(0.511)	(0.431)	(0.438)	(0.537)	(0.528)
1–2 years after	1.30***	1.40***	1.12**	1.25***	1.21***	1.75***	1.29**
	(0.337)	(0.317)	(0.405)	(0.335)	(0.337)	(0.395)	(0.433)
3–5 years after	0.929**	1.10***	0.926*	0.867*	0.883*	1.31**	0.647
	(0.355)	(0.334)	(0.424)	(0.355)	(0.358)	(0.417)	(0.443)
6+ years after	1.32**	1.52***	1.22*	1.23**	1.22**	1.89***	1.14*
	(0.403)	(0.377)	(0.495)	(0.404)	(0.407)	(0.461)	(0.486)
Observations	8,854	10,700	5,104	8,644	8,679	7,623	7,059
Panel C: Group C							
1–5 years prior	-0.088	-0.151	-0.321	-0.148	-0.102	0.666*	-0.040
	(0.201)	(0.185)	(0.252)	(0.201)	(0.202)	(0.279)	(0.251)
Year of appearance	3.95***	3.96***	3.25***	3.84***	3.95***	5.24***	3.52***
	(0.304)	(0.280)	(0.379)	(0.306)	(0.307)	(0.416)	(0.385)
1–2 years after	1.34***	1.31***	1.10**	1.30***	1.28***	2.14***	1.44***
	(0.275)	(0.255)	(0.347)	(0.277)	(0.276)	(0.378)	(0.352)
3–5 years after	0.964**	1.01***	0.759*	0.890**	0.923**	1.66***	0.677
	(0.298)	(0.278)	(0.373)	(0.301)	(0.300)	(0.400)	(0.368)
6+ years after	1.63***	1.65***	1.32**	1.53***	1.56***	2.23***	1.43***
	(0.350)	(0.327)	(0.430)	(0.354)	(0.352)	(0.449)	(0.428)
Observations	10,225	12,032	5,986	9,984	10,023	6,404	6,588
School characteristics	YES	YES	YES	YES	YES	YES	YES
City characteristics	YES	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Year×Region FE	NO	NO	NO	NO	NO	NO	NO

Table 4. Robustness checks using altered sample selections

Note: Robust standard errors are clustered at the school level and appear in parentheses. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively.

Appendix Table A6 also presents the estimation results using alternative definitions of nonacademic track schools and high achievers. Columns (1)-(3) present estimation results where high achievers are defined using certain hensachi score thresholds rather than the groups of universities shown in Figure 2. Columns (4)-(6) show the estimation results using samples in which the threshold for defining nonacademic track schools is extended from five to ten years. Columns (7) and (8) report estimates that define nonacademic track schools by Group B rather than Group C. All estimated results are consistent with the main results, confirming that the main results are not sensitive to changes in the definition of nonacademic track schools or high achievers.

5.4. Heterogeneity of impact

Because peer and role model effects can be heterogeneous (e.g., Clots-Figueras, 2012; Luppino, 2015), a subsample analysis is used here to examine the heterogeneous effects of high achievers by school and regional characteristics. Table 5 presents the estimation results from using the two divided samples according to the number of students, the taxable income, the percentage of university graduates in the neighborhood, and whether the school offers a regular course. Columns (2) and (3) indicate that smaller schools are more affected by high achievers, which is interesting given that high achievers are more likely to appear in larger schools. This could be because smaller schools are more likely to recognize high achievers, and student confidence and application information are scarce in such schools. Columns (4)-(7) show that the impact of high achievers is more

remarkable in areas with lower taxable income and education, which is consistent with the explanation that the more disadvantaged an area is, the greater the room for improvement. Columns (8) and (9) show that the appearance of high achievers positively impacts both general and vocational high schools but has a relatively weak impact on vocational high schools. This indicates that the effect of high achievers is weak for students who do not wish to attend university, reinforcing that high achievers are the cause of the main results.

	Baseline	Number of students		Taxable income		Ratio of university or higher graduates		With or without a general course	
		Large	Small	High	Low	High	Low	With	Without
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Group A									
1–5 years prior	1.12	0.847	1.55*	0.814	1.53	1.51	-0.051	0.667	1.25
	(0.698)	(0.961)	(0.643)	(0.564)	(1.34)	(0.805)	(1.29)	(0.618)	(1.07)
Year of appearance	9.15***	7.46***	12.6***	6.24***	12.9***	7.94***	10.9***	5.02***	13.8***
	(1.56)	(1.86)	(2.10)	(1.58)	(2.34)	(1.84)	(2.39)	(1.04)	(2.27)
1–2 years after	1.78*	2.02	0.852	1.88**	1.51	2.29*	0.404	1.34	1.81
	(0.847)	(1.10)	(0.893)	(0.643)	(1.78)	(1.01)	(1.52)	(0.790)	(1.36)
3–5 years after	2.27*	1.83	3.43*	1.66	3.21	2.25	1.82	1.95*	1.94
	(0.916)	(1.14)	(1.38)	(0.869)	(1.74)	(1.15)	(1.27)	(0.955)	(1.44)
6+ years after	2.41**	2.44*	1.50	1.63	3.84*	2.45*	1.77	1.70	2.24
	(0.922)	(1.14)	(1.37)	(0.897)	(1.72)	(1.16)	(1.33)	(0.933)	(1.50)
Number of treatments	52	38	14	34	18	33	19	30	22
Number of controls	1115	540	575	546	569	547	568	519	596
Observations	10,433	5,899	4,534	5,482	4,951	5,513	4,920	4,046	6,387
Panel B: Group B									
1–5 years prior	-0.102	-0.430	0.421	-0.062	-0.374	-0.218	0.034	0.478	-0.473
	(0.254)	(0.271)	(0.510)	(0.248)	(0.600)	(0.264)	(0.530)	(0.331)	(0.352)
Year of appearance	5.41***	4.08***	8.55***	4.28***	7.53***	4.33***	7.36***	4.76***	5.88***
	(0.432)	(0.413)	(0.960)	(0.411)	(0.975)	(0.438)	(0.891)	(0.555)	(0.611)
1–2 years after	1.30***	0.651	2.63***	1.02**	1.73*	1.06**	1.66*	1.42**	1.26**
	(0.337)	(0.354)	(0.736)	(0.351)	(0.759)	(0.364)	(0.684)	(0.450)	(0.470)
3–5 years after	0.929**	0.543	1.54*	0.594	1.37	0.750	1.08	1.32**	0.690
	(0.355)	(0.384)	(0.743)	(0.387)	(0.762)	(0.389)	(0.702)	(0.487)	(0.491)
6+ years after	1.32**	0.758	2.28**	0.738	2.09**	0.960*	1.79*	1.92**	0.954
	(0.403)	(0.441)	(0.832)	(0.459)	(0.802)	(0.457)	(0.740)	(0.583)	(0.545)
Number of treatments	476	322	154	310	166	304	172	250	226
Number of controls	559	190	369	204	355	210	349	242	317
Observations	8,854	5,009	3,845	4,721	4,133	4,775	4,079	3,325	5,529
Panel C: Group C									
1–5 years prior	-0.088	-0.453*	0.321	0.010	-0.219	-0.034	-0.145	0.516	-0.455
	(0.201)	(0.226)	(0.359)	(0.230)	(0.340)	(0.226)	(0.341)	(0.329)	(0.253)
Year of appearance	3.95***	2.76***	5.51***	3.27***	4.60***	3.43***	4.51***	3.94***	4.01***
	(0.304)	(0.328)	(0.545)	(0.368)	(0.491)	(0.364)	(0.497)	(0.469)	(0.400)
1–2 years after	1.34***	0.710*	2.09***	1.09**	1.56***	1.08**	1.60***	1.66***	1.18**
	(0.275)	(0.314)	(0.481)	(0.344)	(0.438)	(0.340)	(0.436)	(0.431)	(0.358)
3–5 years after	0.964**	0.423	1.56**	0.629	1.22**	0.775*	1.09*	1.48**	0.660
	(0.298)	(0.349)	(0.513)	(0.378)	(0.471)	(0.366)	(0.474)	(0.470)	(0.388)
6+ years after	1.63***	0.938*	2.39***	0.965*	2.24***	1.06*	2.19***	2.28***	1.25**
	(0.350)	(0.415)	(0.585)	(0.445)	(0.545)	(0.422)	(0.559)	(0.574)	(0.447)
Number of treatments	915	511	404	486	429	490	425	478	437
Number of controls	257	72	185	96	161	93	164	110	147
Observations	10,225	5,754	4,471	5,326	4,899	5,450	4,775	4,102	6,123
School characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES
City characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year×Region FE	NO	NO	NO	NO	NO	NO	NO	NO	NO

Table 5. Subsample analysis for heterogeneous effects

Note: Robust standard errors are clustered at the school level and appear in parentheses. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively.

5.5. Spillover effect

The main results show that schools with high achievers have been continually improving their admission performance for over a decade. This is a result of high achievers improving the confidence and application information of the students, thus improving their performance and career choices. However, the reputation of high achiever can attract outstanding students from surrounding schools. In this scenario, the appearance of high achievers can lower the placement rates of the surrounding high schools.

To address this concern, we estimate the average outcome of high schools within 10 kilometers of a given high school as the explained variable and report the results in Table 6. Columns (1)-(3) present the estimation results using the surrounding high schools' mean, maximum, and minimum hensachi scores as explained variables. The results showed that the appearance of high achievers had no significant impact on the university acceptance performance of the surrounding schools. Column (4) provides the estimation results using the number of students enrolled in the surrounding schools as the explained variable. Although there appears to be a slight increase in the number of students following the appearance of high achievers only when Group A is used in the definition, it is not statistically significant. Columns (5), (6), and (7) present estimation results from using the rate of job placement, university application, and university acceptance in the surrounding schools as explained variables. These results suggest that the occurrence of high achievers do not attract the brightest students from the surrounding schools. Rather, high school students seem slightly more likely to choose higher education over employment in those areas where high achievers have appeared.

	Hensachi score			Student behavior			
	Avg.	Max.	Min.	Number of students enrolled	Job placement rate	Ratio of university applicants	University acceptance rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Group A							
1–5 years prior	0.177	-0.275	-0.013	21.4	-0.0007	0.004	0.042*
	(0.261)	(0.609)	(0.239)	(87.7)	(0.005)	(0.004)	(0.018)
Year of appearance	0.760	-0.434	0.075	112.7	0.007	-0.006	0.053
	(0.632)	(0.859)	(0.436)	(114.7)	(0.009)	(0.006)	(0.030)
1–2 years after	-0.012	-0.751	-0.061	172.7	0.010	-0.012	0.040
	(0.872)	(0.881)	(0.278)	(123.5)	(0.009)	(0.008)	(0.025)
3–5 years after	0.182	0.255	0.167	219.4*	0.012	-0.014	0.070*
	(0.313)	(0.775)	(0.326)	(111.7)	(0.009)	(0.007)	(0.030)
6+ years after	-0.466	0.564	0.035	161.7	0.016	-0.012	0.113**
	(0.604)	(0.974)	(0.327)	(122.8)	(0.011)	(0.009)	(0.035)
Observations	10,696	10,696	10,696	10,768	10,756	10,756	10,756
Panel B: Group B							
1–5 years prior	-0.063	0.145	0.079	2.15	-0.004*	0.003	0.007
	(0.078)	(0.138)	(0.072)	(25.6)	(0.002)	(0.003)	(0.009)
Year of appearance	0.064	0.238	0.179	4.02	-0.005	0.002	0.020
	(0.166)	(0.175)	(0.101)	(33.9)	(0.003)	(0.004)	(0.012)
1–2 years after	-0.125	0.109	0.172	41.0	-0.005	0.002	0.019
	(0.139)	(0.177)	(0.090)	(33.6)	(0.003)	(0.004)	(0.012)
3–5 years after	-0.125	0.348	0.136	18.6	-0.008*	0.002	0.032*
	(0.149)	(0.191)	(0.096)	(33.9)	(0.004)	(0.004)	(0.014)
6+ years after	-0.229	0.354	0.145	9.86	-0.010*	-0.0002	0.040*
	(0.180)	(0.238)	(0.115)	(39.6)	(0.004)	(0.005)	(0.017)
Observations	9,220	9,220	9,220	9,271	9,259	9,259	9,259
Panel C: Group C							
1–5 years prior	-0.061	0.047	0.007	-1.36	-0.004	0.004	-0.003
	(0.104)	(0.139)	(0.063)	(17.1)	(0.002)	(0.003)	(0.008)
Year of appearance	0.070	0.027	0.020	-6.48	-0.005	0.006	0.005
	(0.182)	(0.186)	(0.083)	(22.5)	(0.003)	(0.003)	(0.011)
1–2 years after	-0.126	0.149	0.033	10.3	-0.004	0.006	0.004
	(0.162)	(0.176)	(0.080)	(23.3)	(0.003)	(0.004)	(0.011)
3–5 years after	-0.166	0.291	0.002	-17.9	-0.004	0.006	0.014
	(0.190)	(0.203)	(0.092)	(26.7)	(0.004)	(0.004)	(0.013)
6+ years after	-0.319	0.432	-0.002	-16.0	-0.004	0.004	0.013
	(0.249)	(0.246)	(0.115)	(33.6)	(0.004)	(0.005)	(0.016)
Observations	10,343	10,343	10,343	10,398	10,387	10,387	10,387
School characteristics	YES	YES	YES	YES	YES	YES	YES
City characteristics	YES	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Year×Region FE	NO	NO	NO	NO	NO	NO	NO

Table 6. Spillover effect on neighboring high schools (within 10 km)

Note: Robust standard errors are clustered at the school level and appear in parentheses. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively.

6. Conclusion

In this study, an event study framework is used to confirm that prestigious university acceptances (high achievers) appearing in nonacademic track high schools improve the university acceptance performance of their schools over the long term. The appearance of high achievers is unrelated to changes in school input factors, such as the teacher per student ratio, suggesting that high achievers may change school-specific factors such as student and faculty motivation, self confidence, and knowledge of the university admissions process. Although it is impossible to separately identify the potential mechanisms of the effect, it is worth pointing out that high achievers may lead to a sustained improvement in acceptance performance. These results underscore the importance of familiar role models for underrepresented students.

The findings of this study present several important policy implications. First, providing role models for students may be a cost-effective policy because it does not require additional expensive resources. The impact of graduate university selectivity on income and occupation is complex, and the empirical evidence in Japan is scarce, making it difficult to conduct a cost–benefit analysis. Nevertheless, the result of high achievers increasing university selectivity by approximately 0.1-0.2 standard deviations over a decade is substantial. Policies that distribute the brightest students, such as the Top N% Plan, can be expected to exert a long-term improvement effect through high achievers. Conversely, policies such as school choice and tracking may increase social disparities over the long term through reduced confidence or knowledge resulting from fewer high achiever.

The inability to identify potential underlying mechanisms is a limitation of this study. We cannot determine whether high achievers have improved the academic performance

or application destination. defining this pathway through student achievement test scores and application information is a critical future task. Other survey data should also be used to examine whether high achievers change school-specific knowledge of entrance examinations or other factors.

Furthermore, considering the heterogeneity of high achiever effects is also a fruitful future task. Due to the data availability issues in this study, little to nothing is known about the characteristics of high achievers. Recent empirical studies suggest the importance of more familiar and homogeneous role models (Kofoed and McGovney, 2019; Porter and Serra, 2020). Therefore, an analysis using high achiever race, gender, and academic major data would help address issues such as the disparity in STEM enrollment rates between genders.

Finally, understanding the factors that lead to the appearance of high achievers is also essential. Although this study assumes that the appearance of high achievers is quasirandom, it could be that they are more or less likely to appear in a particular institution or environment. For example, school choice and financial aid policies may increase sorting, making it more difficult for high achievers to appear. Conversely, a school district system based on residency can produce a big-fish-little-pond effect as well as a high achiever effect, resulting in increased academic achievement for all. Assessing the academic dispersion within and between schools (within regions) is important to examining these effects to develop a more efficient educational policy.

Declaration of competing interests

The author declares no known competing financial interests or personal relationships that can have appeared to influence the work reported in this paper.

Data availability statement

The datasets provided by DAIGAKUTSUSHIN Corporation and Kawaijuku are commercial, and the raw data cannot be shared without the providers' permission. The raw datasets can be purchased by applying to the respective companies. Although the data from the School Basic Survey are also not permitted to be shared, they can be applied to the MEXT. All other datasets are publicly available. The dataset used can be provided upon request for reproducing the results.

References

Acton, R. K. (2022). Is a name change a game change? The impact of college-to-university conversions. *Economics of Education Review*, *88*, 102240.

Araki, S., Kawaguchi, D., & Onozuka, Y. (2016). University prestige, performance evaluation, and promotion: Estimating the employer learning model using personnel datasets. *Labour Economics*, *41*, 135–148.

Bailey, M. J., & Goodman-Bacon, A. (2015). The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans. *American Economic Review*, *105*(3), 1067–1104.

Barrios-Fernández, A. (2022). Neighbors' Effects on University Enrollment. *American Economic Journal: Applied Economics*, *14*(3), 30–60.

Berkowitz, D., & Hoekstra, M. (2011). Does high school quality matter? Evidence from admissions data. *Economics of Education Review*, *30*(2), 280–288.

Bernard, T., Dercon, S., Orkin, K., & Taffesse, A. S. (2019). Parental Aspirations for Children's Education: Is There a "Girl Effect"? Experimental Evidence from Rural Ethiopia. *AEA Papers and Proceedings*, *109*, 127–132.

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment. *Quarterly Journal of Economics*, *127*(3), 1205–1242.

Black, S. E., Denning, J. T., & Rothstein, J. (2023). Winners and Losers? The Effect of Gaining and Losing Access to Selective Colleges on Education and Labor Market Outcomes. *American Economic Journal: Applied Economics*, *15*(1), 26–67.

Carrell, S., & Sacerdote, B. (2017). Why Do College-Going Interventions Work? *American Economic Journal: Applied Economics*, *9*(3), 124–151.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2020). Income Segregation and Intergenerational Mobility Across Colleges in the United States. *Quarterly Journal of Economics*, *135*(3), 1567–1633.

Clots-Figueras, I. (2012). Are Female Leaders Good for Education? Evidence from India. *American Economic Journal: Applied Economics*, *4*(1), 212–244.

Conley, T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review*, *100*(1), 35–69.

Delaney, J. M., & Devereux, P. J. (2020). Choosing differently? College application behavior and the persistence of educational advantage. *Economics of Education Review*, *77*, 101998.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American*

Economic Review, *101*(5), 1739–1774.

Ersoy, F. (2023). Effects of perceived productivity on study effort: Evidence from a field experiment. *Journal of Economic Behavior and Organization*, *207*, 376–391.

González-Jiménez, V. (2022). Social status and motivated beliefs. *Journal of Public Economics*, *211*, 104662.

Goodman, R., & Oka, C. (2018). The invention, gaming, and persistence of the hensachi ('standardised rank score') in Japanese education. *Oxford Review of Education*, *44*(5), 581–598.

Guyon, N., & Huillery, E. (2021). Biased Aspirations and Social Inequality at School: Evidence from French Teenagers. *Economic Journal*, *131*(634), 745–796.

Guyon, N., Maurin, E., & McNally, S. (2012). The Effect of Tracking Students by Ability into Different Schools. *Journal of Human Resources*, *47*(3), 684–721.

Hoxby, C., & Avery, C. (2013). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students. *Brookings Papers on Economic Activity*, *2013*(1), 1–65.

Hoxby, C. M., & Turner, S. (2015). What high achieving low income students know about college. *American Economic Review*, *105*(5), 514–517.

Hsieh, C. T., & Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics*, *90*(8–9), 1477–1503.

Janzen, S. A., Magnan, N. P., Sharma, S., & Thompson, W. M. (2017). Aspirations failure and formation in rural Nepal. *Journal of Economic Behavior and Organization*, *139*, 1–25.

Kerr, S. P., Pekkarinen, T., Sarvimäki, M., & Uusitalo, R. (2020). Post-secondary education and information on labor market prospects: A randomized field experiment. *Labour Economics*, *66*, 101888.

Kofoed, M. S., & McGovney, E. (2019). The effect of same-gender or same-race role models on occupation choice: Evidence from randomly assigned mentors at west point. *Journal of Human Resources*, *54*(2), 430–467.

Lavy, V. (2021). The Long-Term Consequences of Free School Choice. *Journal of the European Economic Association*, *19*(3), 1734–1781.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., Zhang, L., Shi, Y., Chu, J., & Rozelle, S. (2013). Can information and counseling help students from poor rural areas go to high school? Evidence from China. *Journal of Comparative Economics*, *41*(4), 1012–1025.

Luppino, M. (2015). Peer turnover and student achievement: Implications for classroom

assignment policy. *Economics of Education Review*, *46*, 98–111.

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

Modena, F., Rettore, E., & Tanzi, G. M. (2022). Asymmetries in the gender effect of high-performing peers: Evidence from tertiary education. *Labour Economics*, *78*, 102225.

Mouganie, P., & Wang, Y. (2020). High-Performing Peers and Female STEM Choices in School. *Journal of Labor Economics*, *38*(3), 805–841.

Mulhern, C. (2021). Changing College Choices with Personalized Admissions Information at Scale: Evidence on Naviance. *Journal of Labor Economics*, *39*(1), 219–262.

Page, L. C., & Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, *51*, 4–22.

Pekkarinen, T., Uusitalo, R., & Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, *93*(7–8), 965–973.

Phillips, M., & Reber, S. (2022). Does Virtual Advising Increase College Enrollment? Evidence from a Random-Assignment College Access Field Experiment. *American Economic Journal: Economic Policy*, *14*(3), 198–234.

Porter, C., & Serra, D. (2020). Gender Differences in the Choice of Major: The Importance of Female Role Models. *American Economic Journal: Applied Economics*, *12*(3), 226–254.

Rizzica, L. (2020). Raising Aspirations and Higher Education: Evidence from the United Kingdom's Widening Participation Policy. *Journal of Labor Economics*, *38*(1), 183–214.

Roderick, M., Coca, V., & Nagaoka, J. (2011). Potholes on the Road to College. *Sociology of Education*, *84*(3), 178–211.

Sebastian Cherng, H. Y. (2017). If they think I can: Teacher bias and youth of color expectations and achievement. *Social Science Research*, *66*, 170–186.

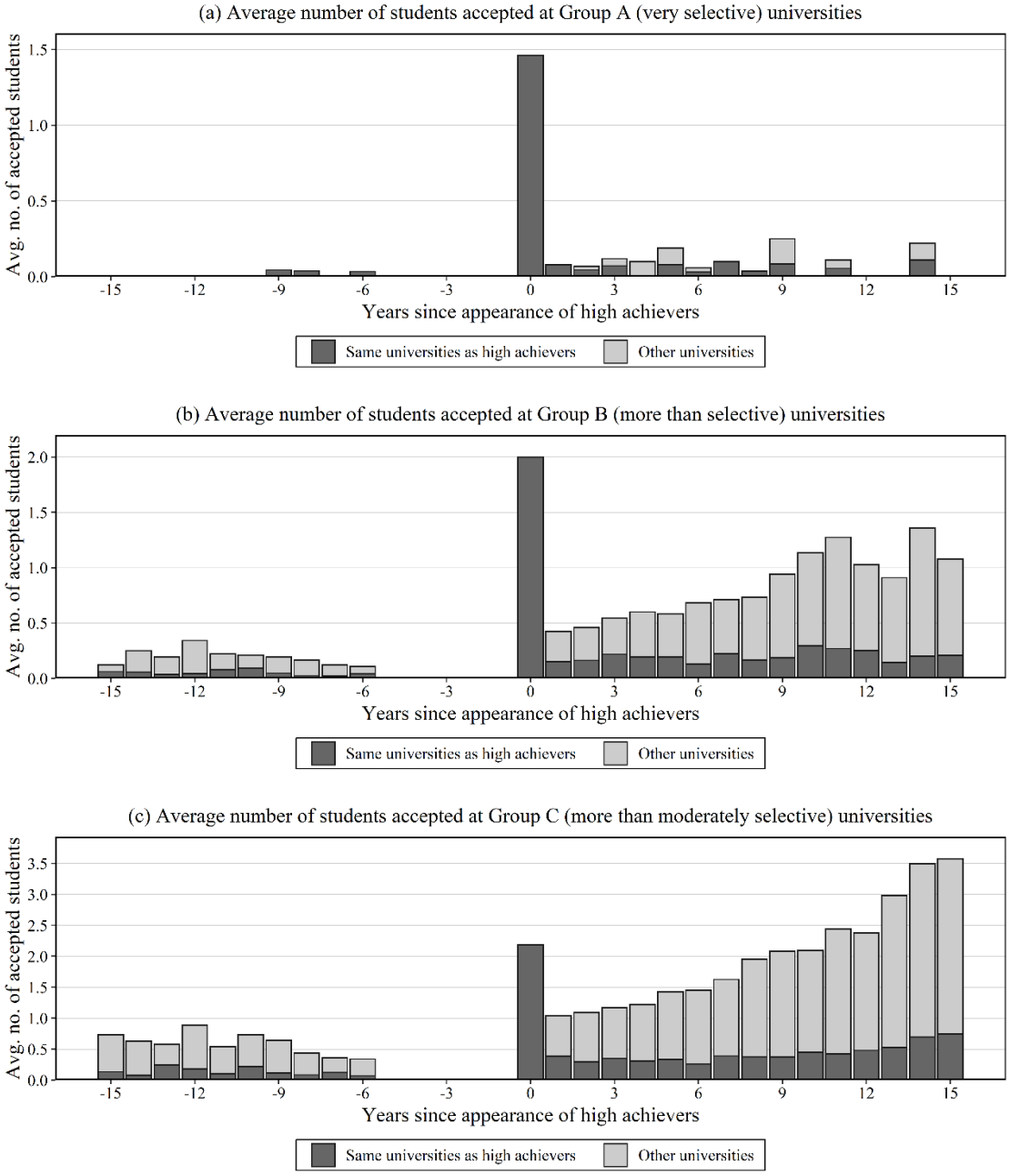
Sequeira, S., Spinnewijn, J., & Xu, G. (2016). Rewarding schooling success and perceived returns to education: Evidence from India. *Journal of Economic Behavior and Organization*, *131*, 373–392.

Sorensony, O. (2017). Regional ecologies of entrepreneurship. *Journal of Economic Geography*, *17*(5), 959–974.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Appendix

Figure A1. Average number of successful applicants by relative time period since the appearance of high achievers



Note: Panels (a), (b), and (c) show the results of defining high achievers by universities in Groups A, B, and C, respectively. In each panel, the vertical axis represents the average

number of successful applicants, and the horizontal axis represents the relative time since the appearance of high achievers. In each high school, the number of students accepted to the same university as the high achiever is accepted to is represented by dark gray bars, and the number of students accepted to different universities than the high achiever is represented by light gray bars.

Table A1. Representativeness and availability of university data

	Number of universities				Availability of hensachi score	
	MEXT Survey		University acceptance data		Unique universities in university acceptance data	Universities with available hensachi scores
	2010	2020	2010	2020		
Total	778	795	484 (62.21%)	507 (63.77%)	619	568 (91.76%)
National	86	86	82 (95.35%)	82 (95.35%)	96	92 (95.83%)
Public	95	94	79 (83.16%)	92 (97.87%)	82	58 (70.73%)
Private	597	615	317 (53.10%)	333 (54.15%)	441	432 (97.96%)

Note: The number of universities of each type is based on the School Basic Survey of the MEXT. University acceptance data are based on a survey by DAIGAKUTSUSHIN Corporation, and hensachi scores are based on reports from Kawaijuku.

Table A2. Hensachi scores and international rankings of the universities in the prestigious group

University name	Abbreviation	Type	Group		Hensachi score	Hensachi ranking	QS WUR 2022	THE JUR 2023	THE WUR 2023
			A	B					
The University of Tokyo	Tokyo	National	1	1	68.9	2	23	2	39
Hitotsubashi University	Hitotsubashi	National	1	1	67.1	5	531-540	16	-
Keio University	Keio	Private	1	1	66.6	6	201	12	801-1000
Kyoto University	Kyoto	National	1	1	64.9	11	33	5	68
Waseda University	Waseda	Private	1	1	64.8	13	203	14	1001-1200
Tokyo Institute of Technology	Tokyo Tech	National	1	1	64.4	15	56	4	301-350
Osaka University	Osaka	National	1	1	61.3	24	75	3	201-250
Nagoya University	Nagoya	National	1	1	59.5	32	118	8	301-350
Tohoku University	Tohoku	National	1	1	58.6	38	82	1	201-250
Hokkaido University	Hokkaido	National	1	1	58.3	40	145	7	501-600
Kyushu University	Kyushu	National	1	1	57.5	48	137	6	501-600
International Christian University	ICU	Private	0	1	65.8	8	801-1000	10	-
Sophia University	Sophia	Private	0	1	62.9	21	801-1000	22	1500+
Rikkyo University	Rikkyo	Private	0	1	60.2	27	1001-1200	40	1001-1200
Meiji University	Meiji	Private	0	1	60.1	28	1001-1200	45	1500+
University of Tsukuba	Tsukuba	National	0	1	59.8	30	285	9	501-600
Aoyama Gakuin University	AGU	Private	0	1	59.5	33	1201+	52	1500+
Doshisha University	Doshisha	Private	0	1	59.4	35	1001-1200	43	1500+
Kobe University	Kobe	National	0	1	59.0	36	386	13	801-1000
Chuo University	Chuo	Private	0	1	58.1	42	-	54	1500+
Yokohama National University	YNU	National	0	1	57.9	43	801-1000	28	1201-1500
Hosei University	Hosei	Private	0	1	57.5	47	-	59	1201-1500
Gakushuin University	Gakushuin	Private	0	1	56.9	51	-	49	-
Kwansei Gakuin University	KGU	Private	0	1	55.9	58	-	44	1500+
Ritsumeikan University	Ritsumeikan	Private	0	1	55.5	61	751-800	31	1201-1500
Kansai University	Kansai	Private	0	1	54.8	67	-	66	1500+
Hiroshima University	Hiroshima	National	0	1	54.1	75	343	11	801-1000

Note: International university rankings are based on QS World University Rankings 2022, Times Higher Education World University Rankings 2023, and Times Higher Education Japan University Rankings 2023.

Table A3. Number of secondary schools per government statistics and our dataset

	School Basic Survey by MEXT				University acceptance data	
	All schools		Only schools with college-bound students			
	2010	2020	2010	2020	2010	2020
High school	4,950	4,582	4,844	4,535	4,803	4,449
National	17	17	17	17	17	17
Public	3,630	3,264	3,531	3,220	3,496	3,138
Private	1,303	1,301	1,296	1,298	1,290	1,294
Upper secondary school	24	51	23	51	23	51
National	2	4	2	4	2	4
Public	13	31	12	31	12	31
Private	9	16	9	16	9	16
Total	5,225	4,820	4,867	4,586	4,826	4,500

Note: The number of schools of each type is based on the School Basic Survey of the MEXT.

Table A4. Example of the number of successful applicants in both university-released data and my data

	Hensachi score	Number of successful applicants			
		Published by University		University acceptance data	
		2010	2020	2010	2020
National					
University of Tokyo	68.88	3,109	3,098	3,081 (99.11%)	3,040 (98.13%)
Tohoku University	58.55	2,602	2,489	2,552 (98.08%)	2,435 (97.83%)
Okayama University	53.51	1,909	2,399	1,907 (99.90%)	2,336 (97.37%)
Niigata University	50.79	2,428	2,467	2,203 (90.73%)	2,297 (93.11%)
Private					
Keio University	66.63	10,253	8,590	1,0176 (99.25%)	8,531 (99.31%)
Kanagawa University	45.90	8,947	6,167	8,877 (99.22%)	6,133 (99.45%)
Meikai University	38.05	753	1,085	751 (99.73%)	1,073 (98.89%)
Seitoku University	34.93	1,964	896	1,956 (99.59%)	883 (98.55%)

Note: The number of successful applicants published by each university is obtained from each university's website.

Table A5. Timing of treatments (appearance of high achievers)

Definition of high achievers	Number of treatments in each year																Total number of treated schools	Number of never-treated schools
	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21		
Group A	6	4	4	2	3	4	1	3	3	5	3	2	2	3	6	1	52	1,115
Group B	39	40	35	37	37	26	23	35	30	30	35	35	16	22	20	16	476	559
Group C	62	61	65	75	81	58	55	68	69	54	58	59	30	38	44	38	915	257

Table A6. Robustness checks using alternate definitions for schools and students

	University groups based on hensachi scores			Extension of years nonacademic track definition			Group B nonacademic track	
	Above	Above	Above	GroupA	Group B	Group C	GroupA	Group B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Time since treatment								
1–5 years prior	-0.030	-0.283	-0.186	1.60	0.639	0.515*	0.226	0.091
	(0.542)	(0.251)	(0.194)	(0.939)	(0.343)	(0.239)	(0.438)	(0.171)
Year of appearance	7.47***	5.93***	4.79***	13.0***	7.36***	5.49***	5.48***	4.20***
	(1.07)	(0.461)	(0.313)	(2.48)	(0.659)	(0.417)	(0.846)	(0.278)
1–2 years after	0.379	1.35***	1.30***	2.07	1.88***	2.09***	0.838	1.08***
	(0.734)	(0.346)	(0.290)	(1.23)	(0.439)	(0.348)	(0.588)	(0.232)
3–5 years after	0.588	0.992**	1.16***	2.75	1.29**	1.67***	1.18*	0.910***
	(0.741)	(0.378)	(0.329)	(1.56)	(0.486)	(0.391)	(0.590)	(0.255)
6+ years after	1.08	1.49***	1.67***	3.89**	1.73**	2.27***	1.06	1.17***
	(0.834)	(0.451)	(0.393)	(1.45)	(0.598)	(0.481)	(0.605)	(0.297)
School characteristics	YES	YES	YES	YES	YES	YES	YES	YES
City characteristics	YES	YES	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Year×Region FE	NO	NO	NO	NO	NO	NO	NO	NO
Number of treated schools	90	344	752	13	152	257	144	958
Number of control schools	909	521	214	725	477	341	1303	586
Observations	8,879	7,497	8,526	6,177	5,112	4,819	13,359	13,744
Adjusted R-squared	0.3739	0.4045	0.4112	0.3948	0.4326	0.4397	0.4214	0.4634

Note: Robust standard errors are clustered at the school level and appear in parentheses.

*, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively.