Economics of Education Review 94 (2023) 102401

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Attending a school with heterogeneous peers: The effects of school detracking and its attenuation[*]

Do Won Kwak[a], Jin Young Lee[b,*]

[a] Graduate School of International Studies, Korea University, Seoul, Korea
[b] Department of Economics, Kangwon National University, 1 Gangwondaehakgil, Chuncheon-si, Gangwon-do 24341, Republic of Korea

ARTICLE INFO

JEL classification:
I20
I21
I28

Keywords:
Between-school detracking
Student performance
Peer quality
Private school
Private tutoring

ABSTRACT

This paper examines the effects of high school detracking on Korean Scholastic Aptitude Test (KSAT) scores in the subjects of Korean (reading), mathematics, and English. We found that a transition to detracking reduces test scores by 0.16–0.18 of one standard deviation. Additionally, the administrative autonomy of private schools and the use of private tutoring services at home attenuate the negative effects of detracking. Our findings are in line with those in the previous literature showing a modest effect of tracking (i.e., ability sorting) in advanced countries where the autonomous school system is well organized and commercial educational services are easily accessible and a strong positive effect of tracking in developing countries where access to well-resourced private schools and private education is generally limited.

1. Introduction

Educational tracking is the practice of sorting students into different educational tracks based on academic achievement. Although tracking has been implemented through various means, two types of tracks are currently widely used: within-school tracking, which groups students by ability at the classroom level, and between-school tracking, which assigns students to different schools according to exam scores and other admission criteria.

The introduction of tracking influences student learning because the changes in classroom peer characteristics and the behavior of teachers and parents could be affected by tracking status. For example, within-school tracking could improve educational achievement for all students by improving teaching quality but could reduce the positive spillover effect for low-achieving students by decreasing peer diversity within the classroom (Zimmer, 2003).

It is empirically challenging to identify the causal effect of tracking on student performance. One reason for this is that student assignment across schools and across classrooms within a school is generally not random. Behavioral responses from teachers, principals, and parents also make the causal link uncertain. Teachers might change instructional strategies and teaching materials to better fit heterogeneous students academically, and principals might hire more (or fewer) staff and teachers to adapt to the new system. Parents who are potentially worried about the adverse effects of detracking might employ private tutors to supplement their children's formal education in school. Certainly, the degree of discretionary responses from teachers, principals, and parents to detracking could depend upon their endowments and circumstances.

This paper examines the impact of between-school detracking on student academic performance using a data set from Korean high schools, including schools that transitioned from tracking to detracking in 2002. We use KSAT score population data from 1999 to 2007, which allows us to use within-school variations over time between tracked and detracked schools. We combine the KSAT score data with local education department data, which provides rich information on school characteristics. The combined data set allows us to examine the effects of detracking and its attenuation.

Our results from difference-in-differences (DID) estimations show that the transition to detracking significantly reduced KSAT Korean (reading), mathematics, and English scores by 0.16, 0.18, and 0.16 of

* This study was supported by 2019 Research Grant from Kangwon National University. We are grateful to the South Korean Ministry of Education for providing data. We also thank Steven Haider, Gary Solon, Haishan Yuan, Booyuel Kim, Joseph Han, Jongkwan Lee, Hee-Seung Yang and Carl Sherwood for helpful comments and suggestions. Young-wook Ku, Yongjin Jeon, and Seulki Ryu did excellent data collection work.
* Corresponding author.
E-mail addresses: dwkwak@korea.ac.kr (D.W. Kwak), jinlee@kangwon.ac.kr (J.Y. Lee).

https://doi.org/10.1016/j.econedurev.2023.102401
Received 12 January 2022; Received in revised form 24 January 2023; Accepted 23 March 2023
Available online 29 March 2023
0272-7757/

Table 1
Summary of the empirical literature on tracking/detracking.

Paper	Sample/Period	Empirical Strategy/ Data	Outcomes	Signs of effects: Overall	Signs of effects: Heterogeneous
A. Classroom-level tracking					
Betts and Shkolnik (2000a)	U.S. 7th–9th- and 10th–12th-grade students/1987–1992	OLS/Survey data	Test scores	(±)	(+) for students in the lower tracks; (–) for students in the upper tracks
Figlio and Page (2002)	U.S. 8th- and 10th-grade students/1988 and some follow-ups	OLS, 2SLS/Survey data	Test scores	(±)	Classroom tracking does not harm low-ability children.
Hoffer (1992)	U.S. 7th-grade students/1987–1989	OLS/Survey data	Test scores	(±)	(+) for advanced students; (–) for slower students
Zimmer (2003)	U.S. 8th-grade students/1981	OLS/Survey data	Test scores	(+)	(–) for low- and average-ability students; (±) for high-ability students
Duflo et al. (2011)	Kenyan 1st-grade students/2005–2007	OLS, RD/ Experimental data	Test scores	(+)	High-achieving students gained more than low-achieving students
Sund (2013)	Swedish men and women who finished compulsory school/1994–1998	DID/Administrative data	High school graduation; math grade	(±); (±)	(–) for students with a low-educated family background
B. Classroom-level detracking					
Argys et al. (1996)	U.S. 8th- and 10th-grade students/1988 and some follow-ups	OLS/Survey data	Test scores	(–)	(+) for students in the lower tracks; (–) for students in the upper tracks
Fu and Mehta (2018)	U.S. 5th-grade students/2002	Structural model estimation/ Survey data	Test scores	(–)	(+) for low-ability students; (–) for high-ability students; Parents of high-ability students increase their effort
Gamoran and Mare (1989)	U.S. high school sophomores/1980	OLS/Survey data	Test scores; high school graduation	(+); (+)	(–) for Hispanic students and those with low socioeconomic status; (+) for Black and female students
C. School-level tracking					
Clark (2010)	U.K. 9th-grade students/ 1969–1971	RD/Administrative data	Test scores; course-taking; university enrollment	(+); (+); (+)	
Galindo-Rueda and Vignoles (2007)	U.K. secondary school students aged 16/1974	OLS, 2SLS with matched sample/Survey data	Test scores; educational attainment	(±); (±)	(+) for the most able pupils; (+) for the most able pupils; These positive effects are larger for girls than for boys
Guyon et al. (2012)	Secondary school students aged 16 and 18 in Northern Ireland/ 1990–2000	DID/Administrative data	Number of students who pass exams	(+)	
Malamud and Pop-Eleches (2011)	Romanian men and women born between 1956 and 1961/1992	RD/Romanian census data	High school graduation; university completion	(+); (±)	(+) for students from poor, rural areas and with less educated parents; (±) for students from poor, rural areas and with less educated parents
Pop-Eleches and Urquiola (2013)	Romanian high school students/ 2001–2007	RD/Administrative and survey data	Test scores	(+)	Teachers sort into better-ranked schools; parents reduce effort when their children attend a better-ranked school
Jackson (2010)	10th–12th-grade students in Trinidad and Tobago/2004–2005	RD/Administrative data	Test scores	(+)	The effects are about twice as large for girls than for boys.
Kim et al. (2008)	Korean 10th–11th-grade students/2001	DID, Quantile regression/ Administrative data	Test scores	(+)	(+) for students with high ability; (±) for students with low ability
Abdulkadiroğlu et al. (2011)	U.S. 7th- and 9th-grade students/ 1997–2008 and 2001–2007	RD, 2SLS/Administrative data	Test scores; university enrollment	(+); (±)	
D. School-level detracking					
Kerr et al. (2013)	20-year-old Finnish men/1984–1986	DID/Administrative and census data	Test scores		School detracking improves test scores of students from disadvantaged backgrounds.
Meghir and Palme (2005)	Swedish 6th-grade students/1961 and 1966 (test score), 1990 (educational attainment), and 1985–1996 (earnings)	DID/Administrative and survey data	Educational attainment; earnings	(+); (±)	(+) for students with unskilled fathers and low ability; (+) for students with unskilled fathers
Pischke and Manning (2006)	U.K. students aged 11 and 16/1969–1974	OLS, 2SLS/Survey data	Test scores	(±)	
Dustmann et al. (2017)	German men and women born between 1961 and 1976/1975–2006	RD/Administrative and census data	Labor market outcomes	(±)	
Hall (2012)	Swedish men and women who finished compulsory school during 1986–1990/ 1985–2008	OLS, 2SLS/Administrative data	Educational attainment; earnings	(+); (±)	
Byun (2010)	Korean 7th-grade students/2005–2007	OLS (Hierarchical linear model estimation)/Survey data	Expenditure on shadow education	(–)	(–) for students from low-income family background
Kim and Lee (2010)	Korean 1st–12th-grade students who live in detrack regions/1997	Tobit, censored least absolute deviation/Survey data	Expenditure on shadow education	(+)	(+) for students with high ability, high family income, and highly educated parents
Kang et al. (2007)	Korean men and women/1998–2002	DID/Survey data	Earnings	(±)	(+) for students with high ability; (–) for students with low ability

(*continued on next page*)

Table 1 (*continued*)

D. School-level detracking					
Lee (2014)	Korean men and women/1998–2010	DID/Survey data	Intergenerational income elasticity	(+)	
Ahn and Goh (2021)	Korean men and women who entered high schools between 1966 and 1982/ 1998–2014	DID/Survey data	Noncognitive ability	(–)	The impacts are more pronounced for women than for men; greater decline in agreeableness for higher ability
Hahn et al. (2008)	Korean high schools/2001–2006	OLS, Tobit/Administrative data	Top university admission	(–)	

Note: In the last two columns, (+), (–), and (±) mean positive, negative, and insignificant effects, respectively.

one standard deviation (SD; 1 SD is equivalent to 30 out of 100 maximum points), respectively. Because transitions to detracking started in the 1970s in Korea, the negative effect of detracking on test scores should have been well-known when they were introduced in 2002. Thus, we examine the role of school autonomy and private tutoring in buffering the negative effect of detracking. First, to determine how schools responded to the negative effects of detracking, we exploit private school status. Unlike in the United States and other European countries, in Korea, students are randomly assigned across high schools within the detracked district, including both private and public schools. Therefore, private high schools cannot select their students by using their own admission criteria. However, private schools in Korea have autonomy and more discretion in making changes to reverse the negative effects of detracking than public schools do. Our empirical results show that public schools entirely drive the negative effect of detracking, with no statistically significant drop in test scores for private schools.

Second, we use another source of data, the Korean Labor and Income Panel Study (KLIPS), to examine whether parents had attempted to reverse the negative shock of detracking by using private tutoring services—a popular supplement to school education. We find that students in detrack schools used private tutoring services significantly more than those in tracking schools. Notably, we find that even after controlling for household-level income, the detracking effect on the use of private tutoring services is all positive. This result implies that even poor parents invested in private tutoring services to counter the negative effects of detracking. Given that private tutoring services raise student achievement, the increase in the use of tutoring may mitigate the negative effects of detracking.

Our findings suggest several policy implications. First, school detracking did not reduce test score gaps among schools. Because the introduction of school detracking is to minimize test score inequality across schools, this finding implies that the detracking policy did not attain the intended goal but had some unintended consequences, such as the deterioration of academic achievement. Second, the effects of detracking may be much smaller in advanced countries than in developing countries. This is because the autonomous education system and commercial education market are likely to be better established in advanced countries. Finally, our results show that households at all income levels increased their private tutoring expenditures to maintain the same level of academic achievement. This result implies that school detracking may pose a financial burden on lower-income households.

The rest of the paper is organized as follows: Section 2 reviews the previous literature on tracking and discusses our study's contributions. Section 3 provides institutional details about South Korea's secondary education system and the introduction of detracking in South Korea and describes the data. Section 4 specifies our empirical model to identify detracking effects. Section 5 presents and discusses our results, and Section 6 checks their robustness. Section 7 concludes the paper.

2. Previous literature

The empirical literature on the impact of tracking has produced mixed findings. This is probably because of sample selection and endogeneity problems from multiple sources, which arise in different ways and in various contexts. Table 1 summarizes previous studies on tracking and detracking effects by track type. We tabulate the sample characteristics, empirical strategies, outcome variables, and estimated signs of the effects for each of the studies.

One strand of the literature has focused on within-school tracking, which has been widely adopted in schools in the United States and Canada. For instance, using U.S. school data, Argys et al. (1996); Fu and Mehta (2018); Gamoran and Mare (1989); Hoffer (1992), and Zimmer (2003) found that ability grouping increases educational disparity, benefitting—or at least not harming—high-achieving students and hurting low-achieving students. On the one hand, using Kenyan elementary school experiment data, Duflo et al. (2011) found that tracking significantly increases average test scores and benefits students at all percentiles of test score distributions. However, they found a slightly stronger effect for students at higher test-score percentiles. On the other hand, Abdulkadiroğlu et al. (2011), Betts and Shkolnik (2000a, 2000b), and Figlio and Page (2002) found that ability grouping has little impact on student achievement.

Another branch of the literature focuses on between-school tracking, which is common in European and East Asian countries. Kerr et al. (2013) and Meghir and Palme (2005) found that high-ability students with unskilled fathers benefited in their educational attainment from Finnish and Swedish detracking school reforms, respectively. However, using U.K. school data, Galindo-Rueda and Vignoles (2007) found that the selective school system leads to better average outcomes and more educational inequality. Using data from Northern Ireland, Guyon et al. (2012) found that an increase in the share of pupils attending elite secondary schools has a strong, positive net effect on average test scores. Using regional data from European and southern American countries, Clark (2010), Jackson (2010), and Pop-Eleches and Urquiola (2013) confirmed that attending a better-ranked school has a positive effect on academic outcomes. However, although Pischke and Manning (2006) concluded that the effect of comprehensive schooling on student achievement in the United Kingdom remains unclear, many other researchers argued that tracking or detracking has little impact on student achievement and labor market outcomes (Dustmann et al., 2017; Hall, 2012; Malamud and Pop-Eleches, 2011; Sund, 2013). Betts (2011) conducted a detailed literature review on the economics of tracking in education, and Sacerdote (2011) reviewed works that focused on the peer effects of tracking and peer effects in education.

Some empirical studies about the effects of detracking on various outcomes used Korean data. For instance, Byun (2010) and Kim and Lee (2010) examined the effect of tracking on the expenditure for private tutoring. Hahn et al. (2010) studied the effects of tracking on the probability of student advancement to elite universities. Kang et al. (2007) analyzed the effects of detracking on adulthood earnings. Lee (2014) studied the effects of detracking on intergenerational mobility, and Ahn and Goh (2021) examined soft skills and personality development effects. Kim et al. (2008) found a positive effect of detracking on test scores and argued that tracking helps high-ability students and does not harm low-ability students.

Our paper is distinct from these previous studies in three major aspects. First, our study identifies detracking effects by exploiting an abrupt policy intervention introduced only for the part of locations in the middle of the sample period. This setting provides well-defined treatment and control groups and allows the use of the DID method.

Furthermore, multiple periods for both before and after the intervention allow us to justify the DID estimates. We present event-study estimates on differences in student test scores between the treatment and control groups during the pre-transition, transition, and post-transition periods to examine whether the parallel trends assumption holds and whether the effect of detracking is temporary or persistent.

Second, we examine attenuating roles of school autonomy and private tutoring. We use a transition to detracking in Gyeonggi region in 2002, 28 years after the first transition that occurred in Seoul and Busan in 1974. Schools and parents at the time of transition to detracking in Gyeonggi region observed the previous experiences of other regions and shared worries about the negative effects of detracking on student achievement. Thus, we infer that some schools and parents could change their behaviors to minimize the negative effects of detracking if there are any. Based on a growing body of literature on the impact of school autonomy (Hahn et al., 2018; Hanushek et al., 2013) and private supplementary tutoring (Choi & Park, 2016; Dang & Rogers, 2008; Park et al., 2016) on student performance, we explicitly consider the possibility that the effect of detracking may be diluted by the behavioral responses of school principals, teachers, and parents. Specifically, we examine whether the autonomy of private school administration and supplementary private lessons could counter the negative effects of detracking. In particular, we use survey data that provides household-level information on the use of private tutors and other commercial after-school programs to examine how supplementary private lessons interacted with detracking status.

Finally, our findings —that the effects of detracking are attenuated by autonomous school systems and the use of private tutoring services— provide new insight to interpret potentially conflicting findings of previous studies on the effects of detracking. For instance, in advanced countries, where elite private schools and the use of private tutoring services are relatively prevalent, school detracking did not affect or only slightly affected student performance (Abdulkadiroğlu et al., 2011; Betts & Shkolnik, 2000a, 2000b; Figlio & Page, 2002). Thus, teachers and parents in advanced countries may have more resources and skills to revert the negative effect of detracking, and as a result, the initial negative effect of detracking on test scores may be dampened. On the other hand, in developing countries, where the autonomous school system is less well organized and the use of private tutoring services is scant, schools and parents have relatively fewer resources and tools to react against the negative effect of school detracking, and as a result, strong negative detracking effect on test score was observed (Duflo et al., 2011). Thus, our findings on the attenuating effects can help explain a modest difference in test scores between tracked and detracked students in advanced countries but a significant difference in developing countries.

3. Institutional details and data

3.1. The Korean education system and the transition from tracking to detracking

In Korea, children start primary school at the age of 6 or 7. An academic year starts in March and ends the following February. After completing six years of primary school education, students attend middle school for three years and high school for another three years. The first nine years of schooling are compulsory. Although high school education is not mandatory, most Korean middle school graduates enroll in high school.[1] First, however, they must choose their academic path from among three different kinds of schools: academic, vocational, and special purpose.[2] As of 2009, about 74% of Korean high school students were enrolled in academic high schools.[3] Our analysis focuses on academic high schools (hereafter high schools) because their main purpose is to advance students to college. Twelfth graders (i.e., high school seniors) who wish to go to college must take the KSAT national college entrance exam, which is offered only once a year. Therefore, every student who applies for college in the same year takes the same KSAT. Students can only take the KSAT once per year, but there is no restriction on the number of times in total. Thus, if students wait for one full year, they can retake the exam after graduation. It is common to take the KSAT more than once after graduating high school.

Before 1974, local education departments used entrance exam scores of ninth graders to assign them to high schools. However, in 1974, some school districts began transitioning from exam-based student assignments (i.e., tracking) to random assignments (i.e., detracking).[4] The transition started first in Seoul and Busan, the two most populous cities in Korea. Following the lead of all school districts in Seoul and Busan, about 30 districts in other regions adopted the detracking system between 1974 and 2006.[5] We only use Gyeonggi region, which went through the transition once for some school districts without any reverse transitions.during the sample period although four regions transitioned from tracking to detracking.

The transitions were on large scale in that the regime change affected about a third of a million students who entered the high school during each year of the transition period.[6] Additionally, the transition occurred quickly once the superintendent of a region decided to adopt the detracking system. The whole process of the transition, from announcement to implementation, took less than a year. Therefore, parents' preemptive responses to transitioning to a detrack system were highly implausible. Even if parents preferred the tracking system, they could only move out of detracking districts after the transition had occurred. The proportion of high school students who changed schools during the sample period was low at less than 2% (See Table A1 in Appendix).

Before the transition, under the tracking regime, ninth graders

[1] In 2009, 99.6% of 674,864 Korean middle school graduates enrolled in high school (Statistical Yearbook of Education, Korean Educational Development Institute, http://cesi.kedi.re.kr).

[2] Special-purpose high schools include foreign language, arts, and science high schools. The best ninth graders apply for these schools, and students must pass through the schools' special admission processes.

[3] Statistical Yearbook 2009 of Education, Korean Educational Development Institute, http://cesi.kedi.re.kr.

[4] High school detracking implies the random assignment of middle school graduates to high schools within a school district, but this has not been strictly implemented over time for some regions. These assignment processes are still random in the sense that student assignments are randomized using a lottery and are partially random in the sense that students can submit an application that lists preferences for schools. Students can express their preferences among schools within a district, but they cannot control their assignment because it is done by lottery. A substantial proportion of students can be assigned to less preferred schools if their list of preferences only contains competitive schools in terms of enrollment. However, whether the assignment is random or partially random is not an issue for our identification. Our identification strategy only requires that a transition to detracking definitely increases diversity in test scores (i.e., a proxy for cognitive ability) within the classroom and school. Korean law guarantees this such that in the detracking regime, students cannot be systematically sorted by exam scores.

[5] We call the broadest local district level a region. As of 2009, South Korea has 16 regions. The transitional regions have both detracking and tracking districts because only some of the districts within a region went through a transition.

[6] According to the Korean Statistical Information Service, about 361,000 students entered public high schools in 2010.

(senior students in middle school) took a high school entrance exam and applied for schools in accordance with their exam scores. High school ranking was usually determined by the school's overall student achievement ranking, such as its average KSAT score, which is well known to the public. The top-ranked high school (according to exam performance), with a capacity for 500 first-year students, could admit the top 500 students according to the entrance exam scores. The number-two school, with a capacity for 350 students, could take the 350 next-best students. However, after the transition to detracking, ninth graders no longer took an entrance exam and were randomly assigned by lottery among schools in the district where they lived. Therefore, the transition resulted in an academically heterogeneous student population across classrooms within a school and across schools within a district.[7]

In our study, a portion of school districts in Gyeonggi region transitioned to detracking in 2002. The secretary of the Korea Department of Education decided whether to track or detrack high schools in a district in the early 2000s. Practically, however, it is the superintendent of a region who decides whether to transition to detracking. The Secretary of Education approves a request for tracking or detracking from the superintendent as long as the superintendent has gathered the opinions of local residents through public hearings and polls (Hwang et al., 2010). The superintendent of Gyeonggi region in the early 2000s was elected through an indirect vote by about 28,000 members of school governance committees in the region.[8] Therefore, whether a school district transitions to detracking depended largely on the opinions of school governance committee members of the schools in the district.

Because the transitions to the detracking system were not randomly decided, we need a valid identification strategy for the causal effect of detracking on test scores. We provide two sets of statistical evidence to support the validity of our identification strategy. First, we visually illustrate outcome differences across track and detrack districts over time. We also provide statistical tests of mean difference results to show no significant outcome difference between the track and detrack districts before the transition to detracking. Second, we provide falsification test results, which indicate that no significant difference arose at the time of transition, using variables that detracking should not affect.

3.2. Self-Selection of a transition to detracking

When the detracking system was first introduced in Korea, the primary objective was to narrow the achievement gap among high schools. For instance, Kim et al. (2008) showed that detracking schools performed worse than tracking schools on test scores, but this did not mean the detracking policy had failed.

However, there is still a possibility that some parents may strongly dislike the negative effect of detracking and act on the transition to detracking. For instance, informed parents might know about the negative effects of detracking on test scores and may try to move out of a detracking district and into a tracking district. Specifically, if affluent parents can afford to move to a tracking district and are better informed, they are more likely to transfer their children from a detracking school to a tracking school after the transition. This could cause selection problems such that only the students of less informed and poorer parents are more likely to remain in the detracking school districts.

Furthermore, the self-selection of teachers into detracking schools can be a problem. However, the self-selection of teachers across schools is unlikely to arise in Korea because local governments heavily regulate public high schools with regard to curriculum and teacher assignment.[9] For instance, public high school teachers cannot choose the district or school in which they teach but are randomly assigned. And, teachers rotate among different schools within the same district every five years. The assignment regulation implies that schools that teachers work in should not be systematically correlated with teacher characteristics.

Table 2
Comparison between treatment and control districts.

A. Transition to detracking in Gyeonggi region

	Transition year	Districts	Proportion of students at the first year of transition (2006)
Always detracked during sample period	1979, 1980	(1) Suwon*, (2) Seongnam (sub-districts: Soojeong, Jungwon)*	30%
Transition from track to detrack during sample period	2002	(1) Goyang (2) Bucheon (3) Anyang, (4) Gwacheon, (5) Uiwang-Goonpo, (6) Seongnam (sub-district: Bundang)*	40%
Always tracked during sample period	NA	(1) Gwangmeong, (2) Ansan, (3) Uijeongbu, (4) Youngin, (5) Peongtack, (6) Gwangju, (7) Hanam	30%

B. Track versus detrack districts in Gyeonggi region

	Track	Transition to detrack	Detrack
Number of districts	7	5	1
Number of schools	100	83	45

Note: We indicate Seongnam and Suwon districts which are excluded in the estimations with *. The number of schools reported is based on the value at the end of the sample year.

3.3. Data

Our sample combines the KSAT score population data for all KSAT examinees from 1999 to 2007 with the annual *Regional Statistical Yearbook of Education*. Along with examinees' KSAT scores, the KSAT data includes examinees' demographic and high school information: their gender, high school enrollment status, current or graduated school location (city, suburb, or rural school districts), school region, and school ID number. Scores in mathematics, Korean, and English are our dependent variables, which are standardized to have a mean of zero and a SD of 1 for each cohort. Using the enrollment statuses and school IDs of examinees, we match the KSAT data with information on the school-level characteristics obtained from the *Regional Statistical Yearbook of Education*, which the local Education Department publishes annually.[10] School-level characteristics include establishment type (public versus private), coeducation versus single-sex education, number of students (total and by gender), number of classes, average class size,

[7] Schools typically assign students randomly across classrooms, and school-level heterogeneity is maintained to classroom-level heterogeneity.

[8] The number of members of school governance committees is from a local newspaper article about the Gyeonggi superintendent election in 2009 (http://www.anseongnews.com/front/news/view.do?articleId=16999). The school governance committee of a school consists of five to 15 members who are representative of the teachers and parents of the school and leaders of the community.

[9] Articles No. 32 and No. 40 of the Enforcement Act of the Elementary and Secondary School Education Law define teacher assignment. They specify systemic requirements to assign teachers equitably. See Kim and Han (2002) for more details.

[10] School-level information is available at http://dl.nanet.go.kr/. This document is distinguished from the Statistical Yearbook of Education, which has local-level average data published by the Ministry of Education.

Table 3
Summary statistics.

	Total		Transition		Track	
	Mean	(Std. dev.)	Mean	(Std. dev.)	Mean	(Std. dev.)
KSAT scores						
Mathematics	54.41	(26.47)	52.44	(26.59)	55.01	(26.41)
Reading	55.30	(25.99)	53.89	(26.38)	55.73	(25.86)
English	55.60	(25.52)	54.59	(25.99)	55.91	(25.36)
Student characteristics						
Male (1 if male)	0.55	(0.50)	0.56	(0.50)	0.55	(0.50)
School characteristics (school average)						
Private (1 if private school)	0.43	(0.49)	0.56	(0.49)	0.39	(0.49)
Number of students	535.0	(133.6)	556.5	(140.8)	528.6	(130.7)
Per teacher	8.00	(1.19)	8.40	(1.18)	7.88	(1.17)
Per administrative staff	106.5	(33.0)	104.4	(31.2)	107.0	(33.53)
Class size	49.1	(4.17)	49.72	(4.15)	48.90	(4.15)
Number of high schools	183		100		83	

Note: We use average values across 1999 and 2001 during the pre-intervention periods.
The number of schools reported is based on the value at the end of the sample year.

student–teacher ratio, number of teachers (total and by gender), number of staff, and number of graduates.

Table 2 shows the number of schools and the number of students across the tracked and detracked districts. We use a sample from Gyeonggi region, which had 14 school districts. Of the 14 school districts, two districts, Suwon and Seongnam, were detracked before the beginning of the sample period and they were always detracked districts during the sample period. Five districts, Goyang, Bucheon, Anyang, Gwacheon, and Uiwang-Goonpo, transitioned from track to detrack in the middle of the sample period and the rest seven districts stayed tracked during the sample period. We exclude Suwon and Seongnam from our sample and use the remaining 12 districts (five transition districts and seven non-transition districts) in our main analysis. We drop Seongnam district because Bundang, a subdistrict of Seongnam adopted a different education policy from the other subdistricts of Seongnam and drop Suwon district because it had been always detracked during the sample period.

We provide the basic summary of the statistics for the dependent variables and the main covariates used in the estimations in Table 3. For the schools included in our sample, average test scores in all three subjects are slightly higher in tracked districts than in detracked districts. We also find that school-level characteristics are a bit more favorable in tracked schools in the sense that the student-teacher ratio is lower and class size is smaller for tracked schools than for detracked schools.

4. The model

4.1. Model description

The causal effect we want to capture is the difference in KSAT scores between having three years of high-school education in the detracking regime and having three years of high-school education in the tracking regime while everything else remains the same.

We employs the DID framework for our main analyses. Because our sample covers the period from 1999 to 2007 and a policy intervention happened in the middle of sample period, we have an ideal setting for the DID framework with well-defined pre- and post-intervention periods and treatment and control groups such that the transition schools belong to the treatment group (i.e., detracked schools) and the remaining nontransition schools belong to the control group (i.e., tracked schools). We call senior students who took the KSAT in year t cohort t. For example, freshmen who entered detracked high schools in 2002 are cohort 2004 because they took the KSAT in 2004. Because our sample covers cohort 1999 to cohort 2007 (i.e., senior students who took the KSAT from 1999 to 2007), we have three pre-transition cohorts (1999–2001 cohorts), four transition cohorts (2002–2005 cohorts), and two post-transition cohorts (2006–2007 cohorts). For the transition cohorts, students from both the tracking and detracking regime shared the school campus. In a detracking district, the 2004 and 2005 cohorts are detracked students who shared the campus with older tracked cohort (s) while the 2002 and 2003 cohorts are tracked students who shared the campus with younger detracked cohort(s). We treat the 2002–2005 cohorts as a mixed treatment group and include the indicator of the group in the estimations.

We later establish the credibility of our identification strategy by visually and statistically examining whether the parallel trends assumption on the outcomes (i.e., three subjects of KSAT scores) is satisfied.

4.2. Estimation model

In our baseline regression, we apply the simplest version of the DID method and obtain the overall effect of detracking. Under the parallel trends assumption, the baseline model is as follows:

$$y_{ijkt} = \gamma_1 \cdot DiD_{k,t} + \gamma_2 \cdot ShareCamp_{k,t} + Z^1_{jkt}\pi + v_j + f_t + \varepsilon_{ijkt}, \quad (1)$$

where y_{ijkt} is the outcome of interest, which is the standardized KSAT score of a senior student i who attends school j in school district k in calendar year t; $DiD_{k,t}$ has a value of 1 if the school is located in detracking district k, and the year t is the period from 2006 or later and has a value of 0 otherwise; $ShareCamp_{k,t}$ has a value of 1 if the school is located in a detracking district k with year t in the period from 2002 to 2005, and it has a value of 0 otherwise; and Z^1_{jkt} includes a few time-varying school characteristic variables, such as class size, student-

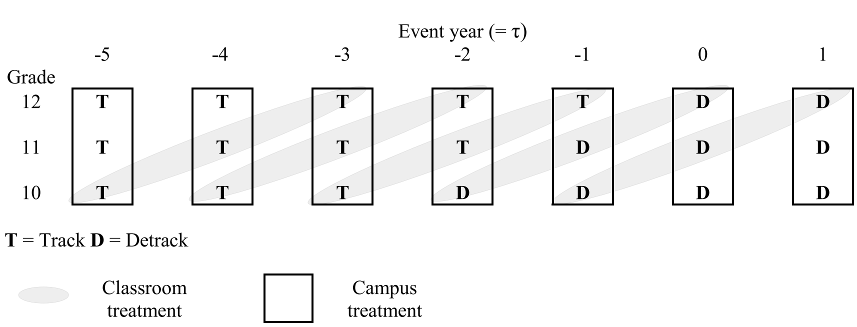

Fig. 1. Campus and classroom treatment.

Table 4
The detrack effect on standardized test scores.

	(1)	(2)	(3)	(4)	(5)
A. Korean					
Detrack	-0.135**	-0.150**	-0.150**	-0.157**	-0.117***
	(0.064)	(0.058)	(0.064)	(0.064)	(0.042)
Mixed campus	-0.039	-0.026	-0.024	-0.020	-0.035
	(0.039)	(0.043)	(0.042)	(0.041)	(0.038)
R-squared	0.075	0.181	0.187	0.187	0.195
Number of obs	636,355	636,355	636,355	636,355	636,355
B. Mathematics					
Detrack	-0.142**	-0.174**	-0.176**	-0.179**	-0.130***
	(0.065)	(0.062)	(0.062)	(0.064)	(0.040)
Mixed campus	-0.050	-0.032	-0.030	-0.027	-0.044
	(0.040)	(0.044)	(0.043)	(0.043)	(0.041)
R-squared	0.069	0.202	0.202	0.202	0.210
Number of obs	658,025	658,025	658,025	658,025	658,025
C. English					
Detrack	-0.146**	-0.154**	-0.154**	-0.159**	-0.117***
	(0.068)	(0.064)	(0.063)	(0.064)	(0.044)
Mixed campus	-0.049	-0.038	-0.035	-0.031	-0.049
	(0.044)	(0.047)	(0.046)	(0.046)	(0.042)
R-squared	0.090	0.218	0.222	0.222	0.231
Number of obs	633,756	633,756	633,756	633,756	633,756
Student characteristics	Yes	No	Yes	Yes	Yes
School characteristics	No	No	No	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes
District time trends	No	No	No	No	Yes
School FEs	No	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors clustered by school are reported in the parentheses. Additionally, ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. School FEs absorb any variation for the detrack treatment variable and cohort FEs absorb before and after indicator variables. Detrack is the variable of the DID interaction term (After * Treat). Student characteristics includes gender (1 if girl). School characteristics includes the number of students, student-teacher ratio, student-administrative staff ratio, and class size.

teacher ratio, student-administration staff ratio, and total number of students. We use test scores for three subjects—Korean, mathematics, and English—as outcome variables. γ_1 captures the estimated detracking effects. In Eq. (1), we account for the time-invariant average differences between tracked and detracked schools using school fixed effects (v_j) and account for aggregate shocks to Gyeonggi region through the year (i.e., cohort) fixed effects (f_t).

To visualize the detracking effect, we extend the DID framework using the event-study framework. When an event, the introduction of detracking, occurs at the beginning of $\tau = 0$, our DID estimating equation could be extended using the following event study framework:

$$y_{ijkt} = \sum_{\tau=-5}^{-1} \beta_\tau Event_{k,t}^{\tau} + \sum_{\tau=0}^{3} \gamma_\tau Event_{k,t}^{\tau} + \gamma_2 \cdot ShareCamp_{k,t} + Z_{jkt}^{1}\pi + \theta_\tau + v_j + f_t + \varepsilon_{ijkt}, \quad (2)$$

where the subscript τ denotes the event year, which is the number of years that have passed since the introduction of detracking, and $Event_{k,t}^{\tau}$ is an indicator variable that equals 1 for detracked districts at year t when τ periods elapse after the introduction and equals 0 otherwise.

We denote the cohorts of $\tau = -2$, $\tau = -1$, $\tau = 0$, and $\tau = 1$ as mixed campus cohorts or transition cohorts. As Fig. 1 shows, the cohorts of $\tau = -2$ and $\tau = -1$ (i.e., the 2002 and 2003 cohorts) attended tracked classrooms for all three years of high school but shared their campus with detracked cohorts. They are transition cohorts. Our pure control cohorts are pre-transition cohorts of $\tau \leq -3$ who attended tracked classrooms and shared their campus only with tracked cohorts. Similarly, the cohorts of $\tau = 0$ and $\tau = 1$ (i.e., the 2004 and 2005 cohorts) attended detracked classrooms for 3 years but shared their campus with tracked cohorts. As a result, the pure treatment cohorts are post-transition cohorts of $\tau \geq 2$ who attended detracked classrooms and shared their campus only with detracked cohorts. In equation (2), the year fixed effects, θ_τ, are included to capture the deviation of the outcome at the event year τ from the baseline year outcome. γ_τ (for $\tau = 2$ and $\tau = 3$) captures the detracking effects at τ periods after the introduction of detracking relative to the control cohorts.

5. Results

5.1. The overall effect of detracking

In Table 4, we show the DID estimates of the detrack effect with various model specifications for three subjects – Korean, Mathematics, and English in panels A, B, and C, respectively. In panel A for the Korean subject, a transition to detracking reduced test scores by 0.12 SD to 0.16 SD, and all coefficients are significant at the 5% level. In all three subjects, coefficients for the detracking effect are significant at the 5% level, and the negative effect estimates are stable around 0.12 SD to 0.16 SD for Korean, around 0.13 SD to 0.18 SD, and around 0.12 SD to 0.16 SD for English. In all three subjects, the negative detracking effect increased in magnitude by adding school-level characteristics. When district-specific time trends are included, for all three subjects, the negative

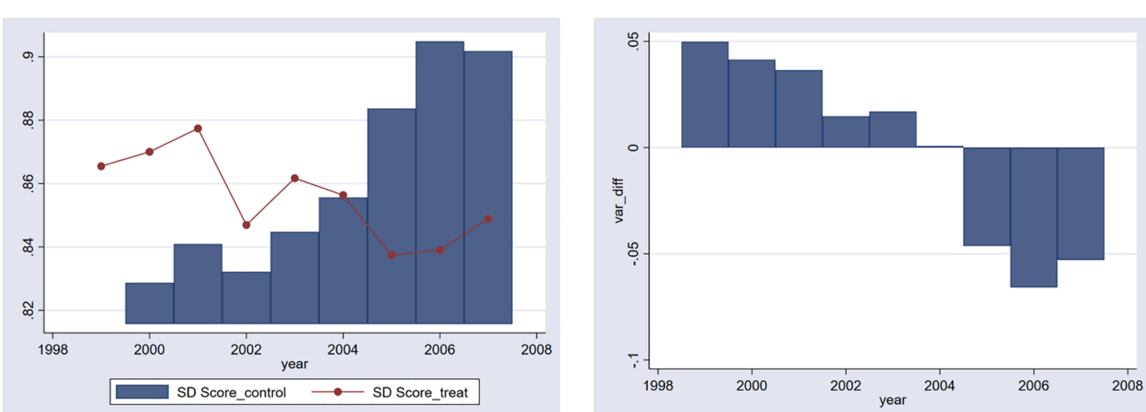

Fig. 2. Dispersion of test scores across tracked and detracked districts. Note: In the left panel, each bar represents standardized standard deviation of average KSAT scores in all track districts and the red line represents standardized standard deviation of average KSAT scores in all detrack districts. In the right panel, we use a bar graph to show the difference of standard deviation between detrack and track districts. Negative values imply standard deviation in detrack districts is smaller than that in track districts. The periods can be classified into three groups: (1) pre-transition (1997–2001), mixed campus (2002–2005), and post-transition (2006–2008).

Table 5
Detrack effect by school-level average score.

	(1) Korean	(2) Math	(3) English	(4) Korean	(5) Math	(6) English
Detrack	0.012	0.022	0.024	0.012	0.023	0.024
	(0.086)	(0.099)	(0.113)	(0.086)	(0.098)	(0.113)
Detrack*G2	0.062	0.103	0.122	0.062	0.103	0.122
	(0.087)	(0.099)	(0.109)	(0.087)	(0.099)	(0.109)
Detrack*G3	-0.146*	-0.187**	-0.153	-0.146*	-0.187**	-0.153
	(0.078)	(0.089)	(0.099)	(0.078)	(0.089)	(0.099)
Detrack*G4	-0.512***	-0.606***	-0.614***	-0.512***	-0.606***	-0.614***
	(0.087)	(0.107)	(0.115)	(0.087)	(0.107)	(0.115)
Mixed (combined)	-0.102	-0.123	-0.100			
	(0.064)	(0.075)	(0.077)			
Mixed(track cohorts)				-0.095	-0.117	-0.094
				(0.072)	(0.085)	(0.087)
Mixed(detrack cohorts)				-0.109	-0.130*	-0.106
				(0.061)	(0.072)	(0.074)
Girl	0.158***	-0.004	0.154***	0.158***	-0.004	0.154***
	(0.010)	(0.012)	(0.012)	(0.010)	(0.012)	(0.012)
Class size	0.011**	0.013***	0.012**	0.011**	0.013***	0.012**
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Student-teacher ratio	0.052***	0.036**	0.057***	0.052***	0.036**	0.057***
	(0.017)	(0.018)	(0.020)	(0.017)	(0.018)	(0.020)
Student-admin staff ratio	0.001*	0.002*	0.002**	0.001*	0.002*	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	183	183	183	183	183	183
Number of observation	649,238	667,205	647,190	649,238	667,205	647,190
R-squared	0.257	0.297	0.305	0.257	0.297	0.305

Note: Robust standard errors clustered by school are reported in the parentheses. Additionally, ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. School FEs absorb any variation for the detrack treatment variable and cohort FEs absorb before and after indicator variables. Detrack is the variable of the DID interaction term (After * Treat). School characteristics includes student-teacher ratio, student-administrative staff ratio, and class size.

detracking effect decreased in magnitude by about 0.04 SD or 0.05 SD with district-specific time trends as column (5) shows.

5.2. The reduction of student sorting by introducing detracking

Korean local education departments introduced detracking to mitigate test performance dispersion from student sorting by test score. They wanted to attain more balanced education services across schools within the region. Thus, if the detracking policy was successful as intended, there should be less difference in the average student's test score across schools within the detracked districts. In Fig. 2, the SD of average KSAT score shows neither an increasing nor a decreasing trend in detracked districts, but the SD of average KSAT score had been increasing in tracked districts during the sample period. Specifically, while the SD of KSAT scores had been increasing in tracking districts, the SD of KSAT scores was lower during the post-transition period than during the pre-transition period.

Because detracked students were randomly assigned to schools, all schools in a detracked district had the same expectation on student test score distribution in the post-transition period. In the pre-transition period, however, schools were tracked and stratified by their student test score. If the random assignment of detracking was well implemented, detracked students who attended high-ranking schools on the basis of the pre-transition average KSAT score of their older tracked

Table 6
School-level response to detracking transition by school type.

	(1) Class size	(2) Student-teacher ratio	(3) Student-admin staff ratio	(4) Number of Students	(5) Class size	(6) Student-teacher ratio	(7) Student-admin staff ratio	(8) Number of Students
Detrack	1.104***	0.168**	0.219	8.407	1.422***	0.251**	-0.669	23.888
	(0.489)	(0.085)	(0.058)	(13.671)	(0.546)	(0.099)	(4.114)	(15.700)
Detrack *Private					-1.343	-0.350**	3.753	-65.481***
					(1.124)	(0.142)	(6.643)	(24.492)
Mixed campus	-0.431	-0.027	-3.514	-5.641	-0.452	-0.033	-3.456	-6.667
	(0.331)	(0.083)	(2.545)	(6.508)	(0.335)	(0.084)	(2.538)	(6.504)
School FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	183	183	183	183	183	183	183	183
Number of observations	636,355	658,025	633,756	636,355	636,355	658,025	633,756	636,355
R-squared	0.078	0.071	0.092	0.187	0.078	0.071	0.092	0.187

Note: Robust standard errors clustered by school are reported in the parentheses. Additionally, ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. School FEs absorb any variation for the detrack treatment variable and cohort FEs absorb before and after indicator variables. Detrack is the variable of the DID interaction term (After * Treat), and Private is an indicator variable of whether a school is a private school. In all columns student and school characteristics variables are included as control variables. Student characteristics includes gender (1 if girl) and school characteristics includes the number of students, student-teacher ratio, student-administrative staff ratio, and class size.

Table 7
The detrack effect on standardized test scores by school type (Public vs private).

	(1) Korean	(2) Math	(3) English	(4) Korean	(5) Math	(6) English
Detrack	-0.189***	-0.214***	-0.195***	-0.240***	-0.267***	-0.248***
	(0.066)	(0.072)	(0.070)	(0.089)	(0.096)	(0.096)
Detrack * Private	0.138	0.142	0.155	0.150	0.169	0.184
	(0.098)	(0.103)	(0.099)	(0.130)	(0.135)	(0.132)
Mixed campus	-0.017	-0.024	-0.029			
	(0.042)	(0.043)	(0.046)			
Girl	0.152***	-0.007	0.124***	0.121***	-0.012	0.101***
	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
Class size	0.007	0.007	0.006	0.017**	0.012	0.015**
	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)
Student-teacher ratio	-0.002	-0.023	-0.001	0.017	-0.007	0.018
	(0.023)	(0.024)	(0.023)	(0.029)	(0.030)	(0.030)
Student-admin staff ratio	0.001	-0.0008	0.001	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Number of students	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
School FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Excluding mixed campus cohorts	No	No	No	Yes	Yes	Yes
Number of clusters	183	183	183	183	183	183
Number of observations	636,355	658,025	633,756	446,732	465,308	444,532
R-squared	0.188	0.203	0.223	0.160	0.185	0.189

Note: Robust standard errors clustered by school are reported in the parentheses. Additionally, ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. School FEs absorb any variation for the detrack treatment variable and cohort FEs absorb before and after indicator variables. Detrack is the variable of the DID interaction term (After * Treat). Private is an indicator variable of whether a school is a private school and is absorbed by School FEs.

students would have lower test scores than their tracked counterparts. Similarly, detracked students who attended low-ranking schools where older tracked cohorts were low achievers would have higher test scores than their tracked counterparts. To examine this, we extend estimating equation (2) to the following equation:

$$y_{ijkt} = \gamma_1 \cdot DiD_{k,t} + \sum_{l=2}^{4} \gamma_j \cdot DiD_{k,t} G_j(l) + \delta_2 \cdot ShareCamp_{k,t} + Z_{jkt}^1 \pi + v_j + f_t + \varepsilon_{ijkt}, \quad (3)$$

where $G_j(l = 2)$ has a value of 1 if student i attends school j that belongs to the 25–50 percentile in terms of pre-transition school ranking by average KSAT score and 0 otherwise; $G_j(l = 3)$ has a value of 1 if student i attends school j that belongs to the 50–75 percentile in terms of pre-transition school rank by average KSAT score and 0 otherwise; $G_j(l = 4)$ has a value of 1 if student i attends school j that belongs to the 75–100 percentile in terms of pre-transition school ranking by average KSAT score and 0 otherwise.

The results in the estimation of Eq. (3) are reported in Table 5. We find that the negative effect of detracking was the largest in magnitude for detracked cohorts who attended the top quarter of schools and the smallest for detracked cohorts who attended the bottom quarter of schools. This finding of a stronger negative effect with pre-transition school ranking implies that the random assignment of detracked students led to expected increases in student academic heterogeneity.

Table 8
Detrack policy and participation in commercial after-school programs.

	Dependent variable: 1 = Participation in any ASP		
	(1)	(2)	(3)
Detrack	0.149***	0.154***	0.139***
	(0.020)	(0.039)	(0.049)
75th–100th percentile of household income		0.281***	0.220**
		(0.058)	(0.085)
50th–75th percentile of household income		0.166***	0.093
		(0.047)	(0.070)
25th–50th percentile of household income		0.090**	0.133*
		(0.043)	(0.073)
Detrack*75th–100th percentile of household income			0.088
			(0.106)
Detrack*50th–75th percentile of household income			0.101
			(0.085)
Detrack*25th–50th percentile of household income			-0.069
			(0.070)
Cohort FEs	Yes	Yes	Yes
District FEs	Yes	Yes	Yes
Number of clusters	47	47	47
Number of observations	2731	932	932
R-squared	0.065	0.122	0.126

Data source: Ninth and eleventh core and supplemental surveys of the Korean Labor and Income Panel Study.
Note: Detrack is the variable for the DID interaction term (After * Treat). Robust standard errors clustered by district are reported in the parentheses. ASP refers to after-school programs (i.e., private tutoring service) for KSAT preparation run by commercial organizations. Additionally, ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively.

5.3. The attenuation of the detracking effects from private schools and family input

In this section, we examine how schools and parents respond to a detracking transition and influence the outcome. Aligned with our finding that the overall effect of detracking is negative and substantial in magnitude, we infer that teachers, principals, and parents may exert some effort in mitigating or even reversing the effects of detracking.

First, to examine the degree of attenuation from the response of teachers and principals at the school level to a transition to detracking, we exploit the effect variation between public and private schools. Although both private and public high schools admit students who have been assigned randomly, private schools may be in a better position to respond to a negative shock than public schools because of their greater administrative autonomy. In Table 6, we estimate the following Eq. (4), which extends Eq. (1) allowing effect heterogeneity by type of school, public or private schools. We also replace the outcome of KSAT score

Table 9
Falsification test using cohorts who shared a campus with the original treatment cohort as treated cohorts.

	(1) Korean	(2) Math	(3) English	(4) Korean	(5) Math	(6) English
False detrack_0203	0.020	0.032	0.039			
	(0.060)	(0.065)	(0.072)			
False detrack_02				0.029	0.034	0.043
				(0.060)	(0.069)	(0.072)
Girl	0.208***	0.047***	0.209***	0.199***	0.031*	-0.213***
	(0.012)	(0.015)	(0.015)	(0.013)	(0.017)	(0.017)
Class size	0.001	0.008*	0.002	-0.001	0.005	-0.000
	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)
Student-teacher ratio	0.029*	0.011	0.028	0.027*	0.007	0.021
	(0.016)	(0.018)	(0.019)	(0.015)	(0.018)	(0.018)
Student-admin staff ratio	0.001	0.001	0.002*	0.001	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of Students	0.001***	0.001**	0.001**	0.001**	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
School FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
# of clusters	164	164	164	153	153	153
# of obs	411,874	411,743	411,6392	358,041	357,960	357,898
R-squared	0.380	0.344	0.404	0.390	0.346	0.409

Note: Detrack is the variable of the DID interaction term (After * Treat). Robust standard errors clustered by school are reported in the parentheses. # refers to number. ***, **, * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively.

with school-level resources such as class size, student-teacher ratio, student-administration staff ratio, and number of students.

$$y_{ijkt} = \gamma_1 \cdot DiD_{k,t} + \gamma_j \cdot DiD_{k,t} Private_j + \delta_2 \cdot ShareCamp_{k,t} + Z^1_{jkt}\pi + v_j + f_t + \varepsilon_{ijkt}, \quad (4)$$

The estimates on "Detrack," which captures the effect for public schools, in columns (5) and (6) in Table 6 imply that a transition to detracking increases class size and student-teacher ratio by 1.42 and 0.25, respectively. However, the estimates on "Detrack * Private," which captures differences of the effect of detracking for private schools, in columns (5) and (6) imply that the student-teacher ratio is significantly smaller for detracked private schools compared to detracked public schools. This means that detracked private schools made choices that led to a lower student-teacher ratio, which are potentially favorable characteristics for students' academic achievement. In Table 7, to examine how the effect of detracking differs between public and private schools, we estimate a three-way interaction model using a private-school dummy as a moderator. In columns (1), (2), and (3) of Table 7, we report results for Korean, mathematics, and English, respectively, with the mixed campus indicator included as a control variable. For public schools—the reference group—a transition to detracking decreased KSAT scores by 0.19 SD, 0.21 SD, and 0.20 SD for Korean, mathematics, and English, respectively, and the effects were all statistically significant at the 1% level. Moderation by private schools of the detracking effect was positive for all three subjects and large in magnitude. After calculating the moderation effect of private schools, we find that the net negative effects of detracking for private schools were 0.05 SD, 0.07 SD, and 0.04 SD for Korean, mathematics, and English, respectively, although the negative effects were not statistically significant. This result suggests that the effect of detracking was attenuated in private schools.

Next, we look at the attenuation of the effect due to private supplementary education, which could involve the potential response of parents who wish to complement their children's school education with private tutoring services. Because our KSAT data do not provide any information on parents of students, we examine whether a transition to a detrack regime leads to more participation in commercial ASPs, including private supplementary tutoring, using another data set, KLIPS, which provides indicators of supplementary education for respondents in eleventh grade.[11] Additionally, the data set provides information on household wealth, which may affect the decision to participate in ASPs and to what extent and should be controlled for in regressions.

The KLIPS data are from the ninth and eleventh waves of the core and supplemental surveys. The surveys provide detailed information on individuals' education history, including the full name of the high school from which a respondent graduated, high school graduation year, high school location, and ASPs in which the respondent participated. Using the location and graduation year, we make a detrack dummy with a value of 1 when a respondent received a treatment of detrack classes for three years and a value of 0 otherwise.

To determine how the detracking policy affects the use of ASPs and whether the detrack effect varies by household wealth, we divide a KLIPS sample into four groups based on household income percentile at the high school graduation year. The first group contained the quarter of high school graduates in the 75th to 100th percentile of the household income; the second group contained graduates in the 50th to 75th percentile; the third contained graduates in the 25th to 50th percentile; and the fourth contained graduates below the 25th percentile. We use three household income group percentile indicator variables to generate interaction terms, and they are used as proxies for household wealth at the time individuals were attending high school. Table 8 shows the results from regressions of the ASP participation indicator variable on the detrack indicator variable with and without control of household income indicators and their interaction terms. The sample used for Table 9 contains individuals who graduated from high school between 1998 and 2006. The results in column (1) in Table 8 imply that a student who attended a school in a detrack district was more likely to participate in an ASP by about 15 percentage points. These results barely changed when we add covariate variables for household income indicator

[11] KLIPS is a longitudinal study of a representative sample of Korean households and individuals living in an urban area. The study began in 1998 with a sample of 5,000 households and members of households aged 15 or older. Information on individuals and households was collected annually, including data on household characteristics, economic activities, labor movement, income expenditure, education, job training, and social activities.

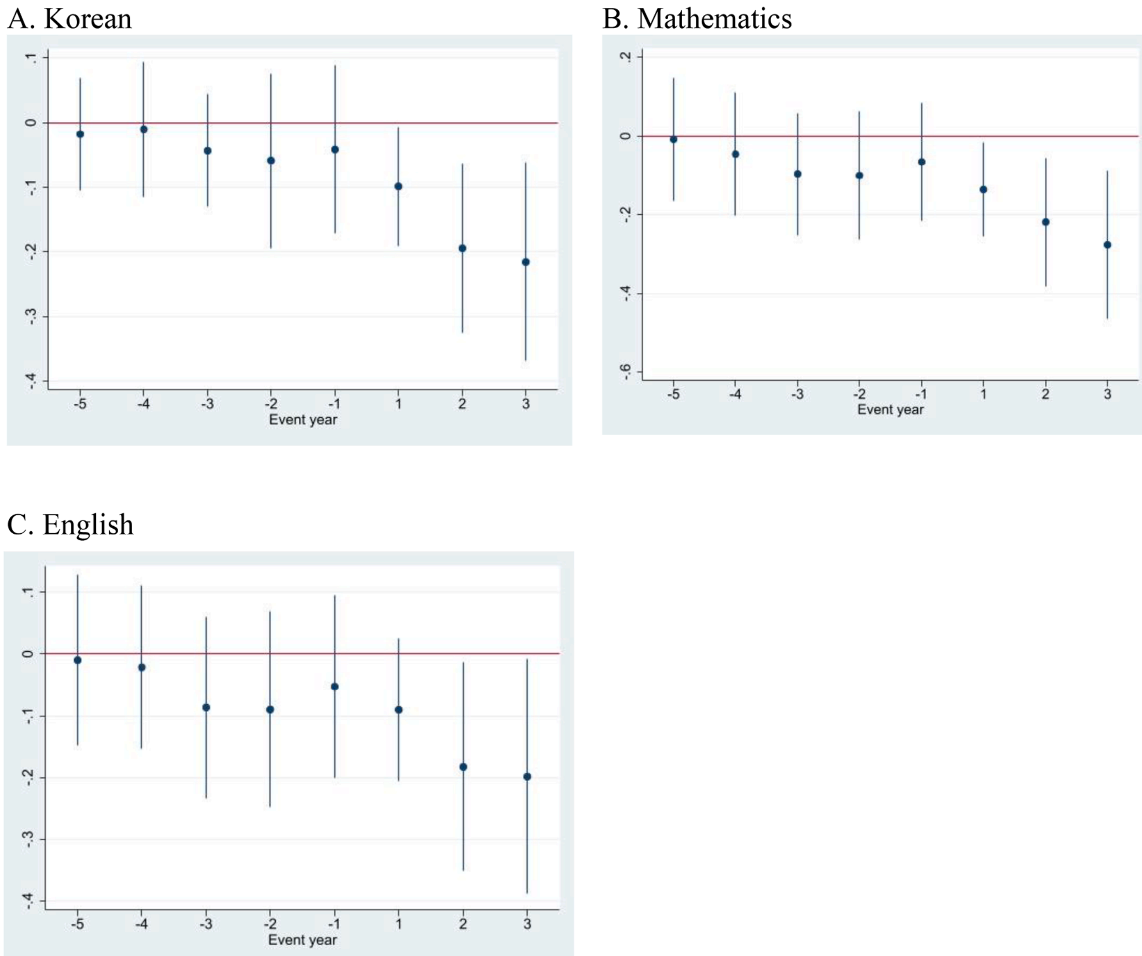

Fig. 3. Event Study Plots. Note: Each point represents the difference in average outcomes between the treatment and control groups for each cohort compared with the baseline cohort, the 2004 cohort (on the red line). The vertical line for each point represents a 95% confidence interval. The event years, which span –5, –4, –3, –2, –1, 1, 2, and 3, indicate 1999, 2000, 2001, 2002, 2003, 2005, 2006, and 2007, respectively. For the cohorts from 2002 to 2005, both track and detrack students shared the same campus. Periods can be classified into three groups: (1) pre-transition (1997–2001), mixed campus or transition (2002–2005), and post-transition (2006–2008).

variables as in column (2). These results imply that parents whose children attended detracking schools increased the use of private tutoring services and other commercial ASPs, holding household income constant. The results in column (2) also indicate that the more household income that individuals had available, the more likely they were to participate in ASPs. However, the detrack effect on private tutoring does not vary significantly by household income. The results in column (3) of Table 8 suggest that the quarter of individuals with the top household income who attended high school in a detracked district was 8.8 percentage points more likely to participate in ASPs than the quarter of individuals with the bottom household income who attended high school in a detracked district. However, the difference was not statistically significant at the 10% level. Similarly, students from households in the 50–75 percentiles of the household income were 10.1 percentage points more likely to participate in ASPs than students from households at the bottom quartile (0–25 percentiles) in household income, but the coefficient estimate was not statistically significant. Given that parents' investment in private supplementary education programs positively affected children's academic performance, we infer that the negative effect of detracking may become modest as the use of private tutoring services increases. In sum, although we find the expected signs and substantial magnitude of the coefficient for the heterogeneous effect of detracking on ASPs by household income level, the coefficient estimates were not statistically significant in all estimations because of large standard errors.

6. Robustness checks

In this section, we examine the robustness of our DID estimation results using a few sensitivity analyses.

6.1. Event-study framework

We show the event study framework, which tests the parallel trends of outcomes between treatment and control groups and visually examines the outcome difference between treatment and control groups in the post-transition periods. Specifically, we report the results of the overall effect of detracking on three KSAT scores in Fig. 3. In this figure, event year –5, –4, –3, –2, –1, 1, 2, and 3 indicate 1999, 2000, 2001, 2002, 2003, 2005, 2006, and 2007, respectively, and we use 2004 as the baseline year. For all three subjects, the negative effect of detracking turns negatively significant at event years 2 and 3, which are pure detracked cohorts.

6.2. Falsification test A: transition cohorts and retakers

We perform two falsification tests using KSAT score of false-treated cohorts. First, instead of using pure detracked cohorts in Section 4, we

Table 10
Falsification test using retakers as treated cohorts.

	(1) Korean	(2) Math	(3) English
False detrack	0.049	0.044	0.073
	(0.047)	(0.061)	(0.057)
Girl	0.166***	-0.069***	0.115***
	(0.016)	(0.022)	(0.022)
Class size	-0.000	-0.000	0.002
	(0.004)	(0.006)	(0.006)
Student-teacher ratio	-0.012	-0.030	-0.017
	(0.014)	(0.019)	(0.017)
Student-admin staff ratio	0.001*	0.002**	0.002*
	(0.001)	(0.001)	(0.001)
Number of Students	0.001***	0.001**	0.001**
	(0.000)	(0.000)	(0.000)
School FEs	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes
Number of clusters	162	162	162
Number of observations	114,878	114,753	114,618
R-squared	0.315	0.312	0.329

Note: Detrack is the variable of the DID interaction term (After * Treat). Robust standard errors clustered by school are reported in the parentheses. ***, **, * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively.

use transition cohorts of 2002 and 2003 who were assigned by test score in the track regime but who shared a campus with detracked cohorts. In Table 9, we report the results from estimating equation (5) while changing treated cohorts. The first three columns in the table show results using both 2002 and 2003 cohorts as false-treated cohorts, and the last three columns show results using only the 2002 cohort as a false-treated cohort. For these false-treated cohorts, we found no significant difference for all three subjects. The magnitude of all coefficient estimates on the (false) DID interaction variable was very small for all three subjects compared with the estimated standard errors.

Next, we conduct another falsification test with KSAT retakers. In our data set, we identifies whether each student was a first-time KSAT taker or a retaker. For instance, the 2004 cohort who took the KSAT for the first time were assigned to high schools under a detrack regime, but retakers who took the exam in 2004 were assigned to high schools under a track regime. Thus, retakers belonging to the 2004 cohort were those assigned by test score in a track regime in both the treatment and control districts. We use these retakers from the treated districts as a false-treated detrack group and estimated Eq. (1) again. Table 10 shows the results. For all three subjects, we find no detracking effect. All coefficient estimates were positive and statistically insignificant.

6.3. Falsification test B: school and district characteristic trends for track and detrack districts

To show there were no differences in trends of characteristics not directly related to the transition between tracked and detracked districts before and after the event year, we examine a few school and district characteristic trends. Fig. 4A presents trends of the amount of property tax collected, which is a proxy variable for wealth of residents, across tracked and detracked districts. According to the figure, both of the

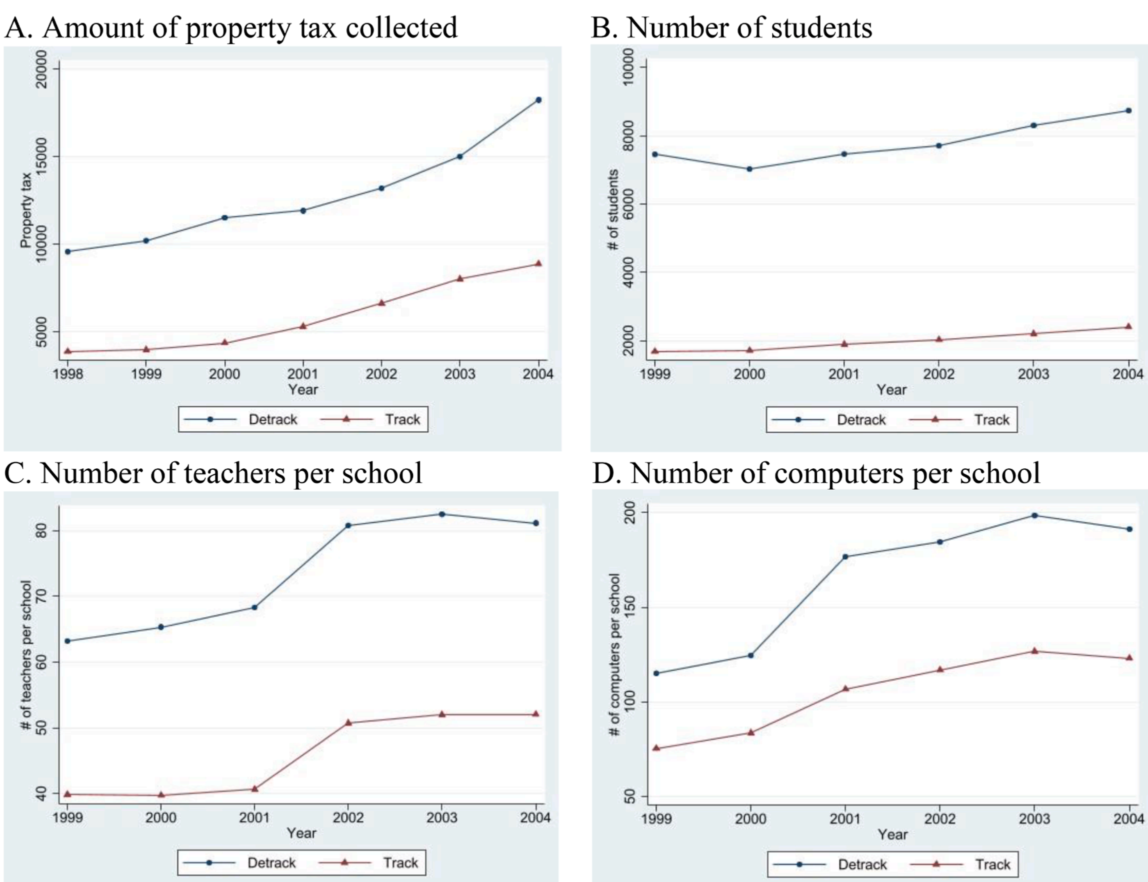

Fig. 4. Characteristics of Gyeonggi Region across Track and Detrack Districts. Note: The amount of property tax collected and the number of students are averaged over districts using population weights. The unit of the amount of property tax collected is one million Korean Won.

trends in the amount of property tax collected for tracked and detracked

Table 11
The impact of detracking on the proportion of students taking KSAT.

	Dependent variable: Proportion of students taking KSAT		
	(1)	(2)	(3)
Policy variable			
Detrack	0.016	0.002	0.009
	(0.011)	(0.007)	(0.008)
Control variable			
Girl	0.001	-0.001	-0.001
	(0.004)	(0.001)	(0.001)
Private	0.004	0.002	0.002
	(0.004)	(0.003)	(0.003)
District-level characteristics			
Log(wealth)			-0.005
			(0.006)
School-level characteristics			
Total number of students		0.000	0.000
		(0.000)	(0.000)
Student-teacher ratio		-0.049	-0.050
		(0.024)	(0.026)
Year FEs	Yes	Yes	Yes
District FEs	Yes	Yes	Yes
R-squared	0.155	0.285	0.264
Number of observations	215,492	214,722	193,508

Note: Robust standard errors clustered by school are reported in parentheses. School characteristic variables include establishment type (public versus private) and school gender type (coeducation versus single-sex), number of students (total and by gender), class size, student-teacher ratio, number of teachers (total and by gender). We use the amount of property tax as a proxy for wealth. School characteristic variables include establishment type (public versus private) and school gender type (coeducation versus single-sex), number of students (total and by gender), student-teacher ratio, number of teachers (total and by gender). We use the amount of property tax as a proxy for wealth.

districts increased from 1998 to 2003, and then the two trends diverged slightly in 2004.[12] This figure implies that there was not much change in the trend of residents' wealth for both track and detrack districts before 2004. Fig. 4B shows trends of the number of 10th graders across tracked and detracked districts. According to the figure, the number of students grew at an almost constant rate from 2000 to 2004. Figs. 4C and 4D present the trends of the number of teachers per school and the trends of the number of computers per school for tracked and detracked districts, respectively. The last two figures imply that there is no significant difference in trends of the level of school resources between tracked and detracked districts. Combined with the evidence in the appendix table that very few students transferred schools during the sample period, Fig. 4B–D indicate that the transition to detracking rarely changed the allocation of school resources per student by the government both before and after the transition.

6.4. Selective KSAT participation

Another concern about the balance between tracked and detracked districts may arise from the KSAT participation rate. If the proportion of students taking the KSAT changed over time and the change showed a different pattern between tracked and detracked districts, then it could bias our estimates. For instance, if more under-achieving students took the KSAT after the school's transition to detracking, then the negative effects of detracking would be overestimated. To examine this possibility, we run a linear regression of the proportion of students taking the KSAT on the indicator of detrack status with other controls. Table 11 shows the results: there is no evidence that the proportion of students taking the KSAT differed between tracked and detracked districts.

[12] The unit of the amount of property tax collected is one million Korean Won.

Table A1
Annual turnover rates for high school students in Gyeonggi.

Year	1995	1996	1997	1998
Grade	10th, 11th and 12th	10th, 11th and 12th	10th, 11th and 12th	10th, 11th and 12th
# of moving-in students (ratio,%)				
within or outside the region	2923	2504	2246	2887
	(1.98)	(1.49)	(1.17)	(1.39)
Total # of students	147,334	168,029	191,596	208,307

Year	1999			2000			2001			2002		
Grade	10th	11th	12th	10th	11th	12th	10th	11th	12th	10th	11th	12th
# of moving-in students (ratio,%)												
within the region	764	384	54	802	356	35	872	321	41	844	307	42
	(1.04)	(0.55)	(0.08)	(1.13)	(0.49)	(0.05)	(1.15)	(0.46)	(0.06)	(1.09)	(0.42)	(0.06)
outside the region	1181	694	113	1309	687	82	1304	675	59	1429	577	59
	(1.61)	(0.99)	(0.16)	(1.84)	(0.95)	(0.11)	(1.72)	(0.97)	(0.08)	(1.85)	(0.78)	(0.09)
Total # of students	73,547	70,183	70,183	71,097	72,434	72,434	75,658	69,593	71,231	77,253	73,889	67,940

Year	2003			2004			2005			2006		
Grade	10th	11th	12th	10th	11th	12th	10th	11th	12th	10th	11th	12th
# of moving-in students (ratio,%)												
within the region	1767	355	27	854	366	37	4118	3438	1612	949	377	28
	(2.16)	(0.47)	(0.04)	(0.98)	(0.46)	(0.05)	(4.44)	(4.02)	(2.03)	(0.97)	(0.42)	(0.03)
outside the region	1564	632	83	1415	653	96	2471	2145	1275	1511	670	61
	(1.91)	(0.84)	(0.11)	(1.62)	(0.81)	(0.13)	(2.66)	(2.51)	(1.61)	(1.55)	(0.74)	(0.07)
Total # of students	81,909	75,559	72,915	87,219	80,367	74,574	92,824	85,490	79,359	97,357	90,741	84,624

Data: *Statistical Yearbook of Education* published by each local Education Department.

6.5. Selective attrition

In the table in the appendix, we report annual school transfer rates between districts for 10th, 11th, and 12th graders to examine whether the movement of students across districts within Gyeonggi region at the time of transition affected turnover rates. Incidences of noncompliance in school assignment among students who transferred from a detracked school to a tracked school or vice versa were quite small. Table A1 indicates that across all transitional districts and all periods, the percentage of movement among students was less than 2% for all grades in high school and less than 0.1% for 12th graders.[13] Thus, it is quite unlikely that attrition occurred selectively across detracked and tracked districts for Korean high school students during the sample period.

7. Conclusion and discussion

This paper investigates the effects of high school detracking on Korean SAT scores. First, we examine the overall effect of detracking and whether the responses of schools and parents attenuated the effects. Our results showed that the transition to detracking significantly reduced test scores by 0.16 SD, 0.18 SD, and 0.16 SD for Korean, mathematics, and English subjects, respectively. For attenuation effects, we find significant negative effect of detracking for public schools. Specifically, the transition to detracking decreased KSAT scores by 0.24 SD, 0.27 SD, and 0.25 SD for Korean, mathematics, and English, respectively. However, the negative effect of detracking came very close to zero in magnitude in all three subjects for private schools. This implies that the administrative autonomy of private schools may reverse the negative effect of detracking. Finally, we find evidence that student participation in ASPs such as private tutoring services may also mitigate the negative effect of detracking. Students at all household income levels in a detracked district increased the use of commercial ASPs regardless of whether they were from low-income or affluent families. Therefore, the potential attenuated effect from commercial ASPs was present for all households regardless of household income level.

Our analysis has a few limitations. First, we leave the mechanism of the attenuation unexplained. Although we find that the negative effect of detracking is mitigated for private schools, we do not know how they obtain better results. We find that detracked private schools have a smaller student-per-teacher ratio, but we could not conduct in-depth investigation of how private schools with autonomy alter their behaviors.

Second, limited available data restrict our analysis on the responses of parents. We find that households at all income levels in detracked districts were more likely to use private tutoring, and the heterogeneous effect by household wealth was not significantly observed. While the coefficient estimates for heterogeneous effect on wealthier families in detracked districts tended to be positive and substantial in magnitude, the estimates were not statistically significant, probably because of the small size of the data. Future studies equipped with rich information on the behaviors of families and schools can investigate more thoroughly the topic of how schools and parents respond to a transition to detracking and achieve better results.

Third, our results should be carefully interpreted and compared with previous studies dealing with younger children. Our study uses high school student data, whereas many other previous studies used elementary school student data or other younger age cohort data. This could be an important issue because, at least in Korea, the variance of test scores becomes greater as students age or advance in their grades. The difference in peers' test scores between tracking (i.e., sorting of students by test score) and detracking (i.e., random assignment) is likely to be greater for middle and high school students than for elementary school students. The large negative effect in our study may not be reproduced in other studies with younger students if there are only minor differences in the test scores for elementary students.

Finally, we only examine the effect of detracking on KSAT test scores. There are numerous other important outcomes that administration officers, teachers, and parents care about. Therefore, our results are only partly suggestive of the adverse effects of the detracking transition policy.

[13] This may be because the transition to detracking mostly occurred in major cities first. Thus, it would be unusual for parents who were concerned about their children's achievement to move from a very big city with many well-known and prestigious schools to a small city nearby. Further, school transfers within the same district were prohibited.

Declaration of Competing Interest

The authors of the manuscript, Dr. Kwak and Dr. Lee, warrant that the article is the authors' original work, hasn't received prior publication and isn't under consideration for publication elsewhere. Additionally, the authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

Data availability

Data will be made available on request.

Appendix

References

Abdulkadiroğlu, A., Angrist, J., & Pathak, P. (2011). The elite illusion: Achievement effects at Boston and New York exam schools. *Econometrica: Journal of the Econometric Society, 82*(1), 137–196.

Ahn, T., & Goh, Y. (2021). The long-term influences of ability mixing on soft skills. *Journal of Economic Behavior and Organization, 191*(C), 367–386.

Argys, L. M., Rees, D. I., & Brewer, D. J. (1996). Detracking America's schools: Equity at zero cost? *Journal of Policy Analysis and Management, 15*(4), 623–645.

Betts, J. R. (2011). The economics of tracking in education. *Handbook of the Economics of Education, 3*, 341–381.

Betts, J. R., & Shkolnik, J. L. (2000a). The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review, 19*(1), 1–15.

Betts, J. R., & Shkolnik, J. L. (2000b). Key difficulties in identifying the effects of ability grouping on student achievement. *Economics of Education Review, 19*(1), 21–26.

Byun, S. Y. (2010). Does policy matter in shadow education spending? Revisiting the effects of the high school equalization policy in South Korea. *Asia Pacific Education Review, 11*, 83–96.

Choi, Y., & Park, H. (2016). Shadow education and educational inequality in South Korea: Examining effect heterogeneityof shadow education on middle school seniors' achievement test scores. *Research in Social Stratification and Mobility, 44* (June), 22–32.

Clark, D. (2010). Selective schools and academic achievement. *B.E. Journal of Economic Analysis & Policy, 10*(1) (Advances), Article 9.

Dang, H., & Rogers, F. (2008). The growing phenomenon of private tutoring: Does It Deepen human capital, widen inequalities, or waste resources? *World Bank Research Observer, 23*(2), 161–200.

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.

Dustmann, C., Puhani, P., & Schönberg, U. (2017). The long-term effects of early track choice. *The Economic Journal, 127*(August), 1348–1380.

Figlio, D. N., & Page, M. E. (2002). School choice and the distributional effects of ability tracking: Does separation increase inequality? *Journal of Urban Economics, 51*(3), 497–514.

Fu, Chao, & Mehta, Nirav (2018). Ability tracking, school and parental effort, and student achievement: A structural model and estimation. *Journal of Labor Economics, 36*(4), 923–979.

Galiendo-Rueda, F., & Vignoles, A. (2007). The heterogeneous effect of selection in U. K. secondary schools. In L. Woessmann, & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 103–128). MIT Press.

Gamoran, A., & Mare, R. D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology, 94*(5), 1146–1183.

Guyon, N., Maurin, E., & McNally, S. (2012). The effect of tracking students by ability into different schools: A natural experiment. *Journal of Human Resources, 47*(3), 684–721.

Hahn, S., Sung, T., & Baek, J. (2008). Mixing versus sorting: Entering top universities. *Economics Letters, 100*, 43–46.
Hahn, Y., Wang, L. C, & Yang, H. (2018). Does greater school autonomy make a difference? Evidence from a randomized natural experiment in South Korea. *Journal of Public Economics, 161*, 15–30.
Hall, C. (2012). The effects of reducing tracking in upper secondary school: Evidence from a large-scale pilot scheme. *The Journal of Human Resource, 47*(1), 237–269.
Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics* (Volume 104,(September), 212–232, 104.
Hoffer, T. B. (1992). Middle school ability grouping and student achievement in science and mathematics. *Educational Evaluation and Policy Analysis, 14*(3), 205–227.
Hwang, J., Park, J., Koh, J., Kim, S., & Hah, B. (2010). *An analysis on determining districts that implement high school detracking*. Korean Educational Development Institute.
Jackson, C. K. (2010). Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago. *The Economic Journal, 120* (December), 1399–1429.
Kang, C., Park, C., & Lee, M. J. (2007). Effects of ability mixing in high school on adulthood earnings: Quasi-experimental evidence from South Korea. *Journal of Population Economics, 20*(2), 269–297.
Kerr, S. P., Pekkarinen, T., & Uusitalo, R. (2013). School tracking and development of cognitive skills. *Journal of Labor Economics, 31*(3), 577–602.
Kim, E and Han, Y., 2002. “Background report: Korea, oecd background report: Attracting, developing and retaining effective teachers”, Korean Educational Development Institute.
Kim, S., & Lee, J. H. (2010). Private tutoring and demand for education in South Korea. *Economic Development and Cultural Change, 58*(2), 259–296.
Kim, T., Lee, J. H., & Lee, Y. (2008). Mixing versus sorting in schooling: Evidence from the equalization policy in South Korea. *Economics of Education Review, 27*, 697–711.
Lee, Y. (2014). Exams, districts, and intergenerational mobility: Evidence from South Korea. *Labour Economics, 29*, 62–71.
Malamud, O., & Pop-Eleches, C. (2011). School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics, 95*(11–12), 1538–1549.
Meghir, C., & Palme, M. (2005). Educational reform, ability, and family background. *American Economic Review, 95*(1), 414–424.
Park, H., Buchmann, C., Choi, J., & Merry, J. (2016). Learning beyond the school walls: Trends and implications. *Annual Review of Sociology, 42*(1), 231–252.
Pischke, J., and Manning, A., 2006, “Comprehensive versus selective schooling in England in Wales: What do we know?” NBER Working Paper Series No. 12176.
Pop-Eleches, A., & Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review, 103*(4), 1289–1324.
Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education, 3*, 249–277.
Sund, K. (2013). Detracking Swedish compulsory schools: Any losers, any winners? *Empirical Economics, 44*(2), 899–920.
Zimmer, R. (2003). A new twist in the educational tracking debate. *Economics of Education Review, 22*(3), 307–315.