When Success Becomes Salient: Local Role Models and Schooling Decisions[⋆]

Wei Huang[a], Zhejian Wang[b,*], Ruoming Zhang[b]

[a] *National School of Development, Peking University, Beijing, China*
[b] *Department of Economics, University of Delaware, Newark, DE, USA*

Abstract

We ask whether locally salient academic achievement affects the schooling decisions of nearby cohorts. Using a hand-collected database of provincial Gaokao top scorers in China between 1990 and 2005 and county-by-cohort outcomes from the 2010 Population Census, we treat the first appearance of a top scorer in a county as a localized exposure event and estimate event-study and staggered difference-in-differences specifications. Cohorts one to two years younger than the top scorer, who were still in senior high school when the event was publicized, show meaningful gains in educational attainment: high school completion rises by 4.5 percentage points, four-year college completion rises by 2.9 percentage points, and years of schooling increase by 0.27 years. Effects are larger in counties with stronger pre-existing educational capacity, indicating that the gains depend on local capacity to convert the response into completed attainment. The findings are consistent with a role-model channel through which a salient, locally identified academic achievement shifts the schooling decisions of nearby cohorts.

Keywords: Role models, Aspirations, Educational attainment, China, Gaokao, Difference-in-differences

[⋆] We are grateful to Jeremy Tobacman, Adrienne Lucas, Barton Jay Willage, and Sabrin Beg for valuable comments and discussions. We also thank seminar and conference participants for helpful feedback. All remaining errors are our own.

[*] Corresponding author. Email: zhejianw@udel.edu.

1. Introduction

Educational decisions depend not only on costs and returns, but also on what students believe is attainable for someone from their local environment (Genicot and Ray, 2017; Dalton et al., 2016). A growing literature in the economics of education studies whether visible high achievers can shift this perception and thereby alter the schooling decisions of nearby peers (Riley, 2024), with implications for the design of low-cost recognition-based and informational interventions (Hoxby et al., 2013). Such beliefs may be especially important in high-stakes education systems, where small changes in effort and persistence can affect access to selective schooling and higher education. We examine this question through a naturally occurring source of public exposure to local academic success: the annual identification of provincial top scorers in China's national college entrance examination, the *Gaokao*. Each year, the highest-scoring students within each exam track in a province are publicly identified out of several hundred thousand candidates;[1] the top scorer's high school and the county in which the school is located are reported in the local press, recognized in school and government ceremonies, and revisited in subsequent years through return speeches and other school-organized events. We treat the first appearance of a provincial top scorer in a county as a localized exposure event, and ask whether it alters the schooling decisions of cohorts who were still in school when the announcement was made.

We combine a hand-collected database of 277 first appearances of provincial top scorers between 1990 and 2005 with individual-level microdata from the 2010 Population Census of China (National Bureau of Statistics of China, 2010, 2012), and estimate event-study and staggered difference-in-differences

[1] In practice, provinces designate separate top scorers in the science and humanities tracks. Our analysis focuses on the science-track top scorer, who attracts the most public attention; see Section 3.1.

specifications using the framework of Callaway and Sant'Anna (2021). The unit of analysis is a county-by-birth-cohort cell. Identification compares cohorts in counties first exposed to a top scorer to cohorts in counties not yet exposed at the same cohort date, conditional on county and cohort fixed effects and on county-by-cohort controls for ethnic and rural composition. The design therefore asks whether, within a county, cohorts just young enough to observe the local success signal before completing secondary school experience larger gains than older cohorts, relative to same-birth-year cohorts in counties whose first top-scorer event occurs later. The main identifying assumption is that, absent the event, educational trajectories in treated and not-yet-treated counties would have evolved similarly conditional on the controls.

The estimates concentrate, in both magnitude and timing, on cohorts a few years younger than the top scorer, that is, students finishing senior high school when the event was publicized. Pooling across these cohorts, exposure raises high school completion by 4.5 percentage points, lifts four-year college completion by 2.9 percentage points, and increases years of schooling by 0.27 years. Relative to pre-treatment means, these estimates represent sizable increases, especially for four-year college completion. Pre-trends are flat across all four educational outcomes we examine, and a within-city specification that compares treated counties to untreated counties within the same prefecture-level city yields broadly similar estimates.

Three patterns make a pure fiscal or supply-side explanation less likely and are consistent with a role-model or aspirations channel. The response is concentrated in cohorts that were still in senior high school at the time of the announcement, consistent with a behavioral adjustment among students who could still alter effort and persistence at the margin. County-level public education expenditures show no detectable change around the event, weakening the case that the response is driven by a contemporaneous fiscal or supply-side adjustment. Humanities-track top scorers, who receive substantially less media attention and represent a smaller and more selectively composed peer group, generate effects close to zero, providing suggestive evidence that the salience of the achievement and the size of the exposed peer group matter for the response. Salience-driven behavioral responses to information have been documented in related contexts such as television exposure and observable peer behavior (La Ferrara et al., 2012; Bursztyn and Jensen, 2015), and the broader peer-effects literature offers complementary evidence on social influence in schooling (Sacerdote, 2001; Duflo et al., 2011), building on the longer-standing identification challenges raised in Manski (1993).

Our paper makes two main contributions to the economics literature. First, we contribute to the literature on aspirations, role models, and educational investment by studying a naturally occurring public signal rather than a researcher-delivered intervention. Existing evidence in this literature largely relies on randomized exposure or administrative matching: female political reservation in Indian villages (Beaman et al., 2012), randomized assignment to a film depicting a Ugandan chess prodigy (Riley, 2024), and exposure during adolescence to a female general practitioner among Norwegian women (Riise et al., 2022); visibility-based peer effects have also been documented at later career stages, including persistent productivity declines among co-authors of star scientists who exit their fields (Azoulay et al., 2010). We instead examine a signal generated within the education system itself, the public emergence of a locally identified Gaokao top scorer, around which we find no evidence of contemporaneous changes in county-level public education expenditure.

Second, we contribute to research on human-capital formation in high-stakes examination systems. In rank-based systems, individual success is publicly interpretable and locally salient, but the broader effects of such success on nearby students are not well documented. Our identification builds on a hand-collected panel of 277 top-scorer events that are rare and geographically dispersed: a typical county had under a one-percent annual probability of producing a provincial top scorer over our sample period. The county-by-cohort exposure design links these events to census-based outcomes and can be adapted to study other public success signals in similar institutional contexts, complementing evidence on how students adjust schooling decisions when their perceived returns are updated through informational interventions (Wiswall and Zafar, 2015).

Beyond these two contributions, we document an additional pattern of heterogeneity in the response: the educational gains are concentrated in female cohorts and appear broadly insensitive to the gender of the role model. Girls in counties with male top scorers and girls in counties with female top scorers both show statistically significant gains in high-school completion and four-year college completion, while boys show effects close to zero in either case. The pattern complements existing role-model evidence that has emphasized same-gender matching as a defining feature (Carrell et al., 2010; Riise et al., 2022), and suggests that baseline scope for updating aspirations may matter for who responds. We interpret this gender heterogeneity cautiously, given the relatively small number of female-scorer counties and the consequent imprecision of the female-scorer estimates.

Our findings show that a discrete public realization of academic success can shift completed schooling in nearby cohorts even when no formal policy accompanies the event, with the response concentrated in cohorts with the most scope to update aspirations under prevailing gender gaps in tertiary attainment. The implied gains in years of schooling are non-trivial relative to standard cross-country returns to schooling (Montenegro and Patrinos, 2014), and they complement broader evidence that local human-capital accumulation generates substantial labor-market externalities (Moretti, 2004; Acemoglu and Angrist, 2000) as well as non-labor-market social returns such as reductions in crime (Lochner and Moretti, 2004). Because higher education admission in our setting is rank-based, however, these reduced-form gains in treated counties should be interpreted as effects on local educational attainment rather than as direct evidence of aggregate welfare gains: increases in college attainment in treated counties may partly reflect displacement from other counties competing for the same admission slots. The mechanism we identify, namely visible and locally attributed academic achievement that shifts what nearby students take to be attainable, operates in many other rank-based examination systems, and our county-by-cohort exposure design can be adapted to study them.

The remainder of the paper proceeds as follows. Section 2 describes the Gaokao system and the institutional features that make a top-scorer realization a publicly visible local event. Section 3 introduces the data and the construction of county-by-cohort exposure. Section 4 sets out the empirical framework. Section 5 reports the main estimates and the within-city specification. Sections 6 and 7 present robustness checks and heterogeneity analyses. Section 8 discusses interpretation, and Section 9 concludes.

2. Institutional Background

2.1. The Gaokao system and the production of top scorers

China's National College Entrance Examination is a centrally administered, high-stakes examination that governs admission to higher education. Although the Gaokao is a national institution, admissions competition is primarily province-specific: universities allocate quotas by province, and admission thresholds are determined within provincial applicant pools (Davey et al., 2007). Within-province ranking is therefore the primary margin of competition.

Within this rank-based system, provinces typically identify the highest-scoring students within major exam tracks. Our analysis focuses on science-track provincial top scorers, who were among the most publicly visible academic achievers during the period we study. The total number of the Gaokao examinees nationwide ranged from approximately 2.8 million in 1990 to 3.8 million in 2000 (Ministry of Education of the People's Republic of China, 2005), implying that provincial top scorers were selected from large examinee pools and represented extreme realizations in the upper tail of the achievement distribution. These outcomes are generally perceived as credible because they are produced by a standardized, high-stakes examination system and interpreted through province-level rank-based allocation. At the county-year level, first appearances are rare, occurring in fewer than one percent of county-year cells in our analysis sample.

2.2. Local visibility and geographic anchoring

Top-scorer outcomes are geographically anchored. Top scorers are almost always publicly identified with a specific senior high school and with the county in which that school is located. Contemporaneous local media coverage emphasizes this geographic attribution.[2] Appendix 1 provides a more detailed accounting of the publicity channels, including newspaper coverage, school-level recognition, local government ceremonies, and the regulatory response that this publicity has generated.

[2] In 1995, China had about 14,000 regular senior high schools nationwide (Ministry of Education of the People's Republic of China, 2005); over a similar period, China had about 2,850 county-level administrative units in total, including urban districts, county-level cities, and counties (National Bureau of Statistics of China, 2015). A simple ratio implies roughly five regular senior high schools per county-level unit on average.

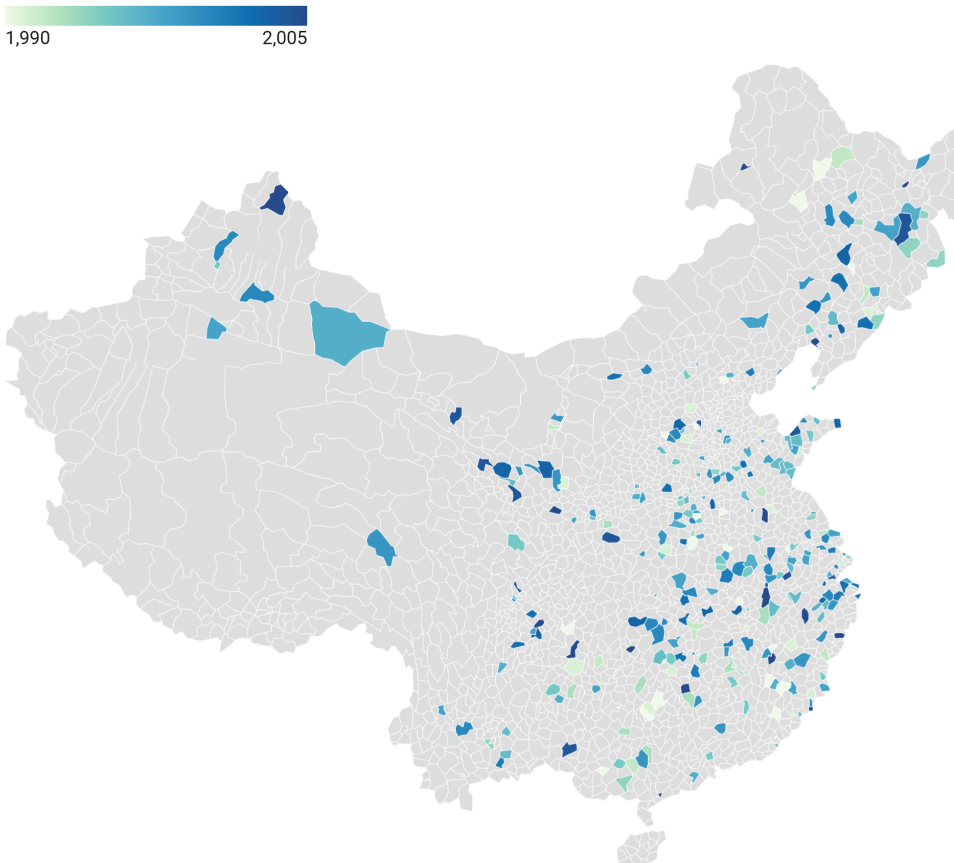


Figure 1: Geographic Timing of the First Appearance of Provincial Top Scorers in China

Notes: This figure maps the first year in which a county-level administrative unit produced a provincial top scorer during the sample period from 1990 to 2005. Counties are shaded according to the year of first appearance, with lighter colors indicating earlier years and darker colors indicating later years. Counties that did not produce a provincial top scorer during the sample period are shown in gray. County-level units include counties, county-level cities, and urban districts, harmonized to a common set of administrative boundaries. For direct-controlled municipalities, core urban districts are aggregated to the city-proper unit for visualization due to limitations in historical boundary definitions. In total, 277 out of 2,543 county-level units experienced a first provincial top-scorer appearance during the sample period. County attribution is based on the location of the scorer's senior high school as reported by official examination records and contemporaneous local media coverage. Counties designated as national-level poverty counties in official government lists are excluded from the analysis sample and are therefore not shown.

Figure 1 maps the first year in which a county-level administrative unit produced a provincial top scorer between 1990 and 2005. Three features stand out. First, first appearances occur throughout the sample period rather than clustering in a small number of years. Second, they arise in both urban and non-urban counties across all major regions of China. Third, the majority of county-level units never produce a top scorer during the sample period: of 2,543 county-level units in the analysis sample, 277 experienced a first top-scorer event during 1990–2005. Treated counties are spatially dispersed rather than forming large contiguous clusters, which limits concerns that the estimates are driven mechanically by spillovers across neighboring counties. We return to broader regional shocks in the empirical analysis through within-city comparisons and additional robustness checks.

Although provincial top scorers are defined at the province level, exposure to their achievement is highly localized. With only a handful of regular senior high schools per county-level jurisdiction, students in the same county are more likely than students elsewhere to be exposed to the same local school networks, media coverage, and public narratives surrounding the top scorer. Local governments, schools, and media outlets emphasize the geographic attribution, framing the achievement as a local accomplishment and as a replicable benchmark for students in the same county. News coverage typically highlights the county in which the scorer's high school is located, the scorer's schooling history, and the local study environment. Such coverage is most intensive in the school locality and may be repeated through school ceremonies, return speeches, and local media reports, transforming a single examination outcome into a persistent local presence. The resulting exposure is spatially concentrated and cohort-specific even though the underlying examination system is national in scope.

These features make the first local appearance of a provincial top scorer a discrete, geographically anchored public signal. The empirical strategy below uses the staggered timing of these first appearances to compare counties before and after the event with counties not yet exposed at the same cohort date, while assessing the identifying assumptions through event-study pre-trends and robustness checks.

3. Data

3.1. Provincial Gaokao Top-Scorer Records, 1990–2005

Our first dataset is a manually constructed province-year compilation of *Gaokao top scorers* covering 1990–2005. For each province and year, we collect the top scorer information whenever available, with a focus on the high school associated with the top-scoring students and the corresponding year. Because our empirical analysis is conducted at the county level, we use the reported school name to locate the top scorer geographically and assign the event to a county administrative unit. We standardize school names across sources and map each school to a county identifier based on the school's location in the relevant year. This procedure yields a county-by-year record of whether a county produced a provincial top scorer, which serves as the foundation for constructing the treatment-timing variable in our staggered-adoption design.

Top-scorer records are hand-collected from archival sources, including contemporaneous newspaper reports, news coverage, and official or government-affiliated announcements. Coverage is incomplete for earlier years, reflecting limited availability and digitization of historical local newspapers and official releases in the early 1990s. Consequently, some province-year observations are missing, with missingness plausibly more common in the early part of the sample. We treat these missing observations as a data limitation rather than true zeros and discuss the potential implications for measurement error in treatment timing. In the empirical analysis, we conduct robustness checks that restrict attention to periods and provinces with more complete reporting and show that the main results are not driven by years with the highest missing rates.

3.2. 2010 Population Census Microdata

Our second data source is an anonymized individual-level microdata sample from China's Sixth National Population Census conducted in 2010, originally produced by the National Bureau of Statistics of China (National Bureau of Statistics of China, 2010). The 2010 Census is a nationally representative enumeration that records rich individual- and household-level information, including demographics, hukou status and place of residence, educational attainment and years of schooling, and other socioeconomic characteristics (National Bureau of Statistics of China, 2012). The microdata contain geographic identifiers that allow us to locate individuals in specific administrative units and to aggregate outcomes to the county-by-birth-cohort level, which matches the unit of analysis in our empirical design. In practice, we construct county $\times$ birth-cohort pseudo-panels by collapsing individual observations within each county and birth-year cell and computing cohort-level outcome measures and co-variate rates. We also retain cell sizes for use as analytic weights in the county-by-cohort analysis.

Linking the census-based outcomes to the top-scorer dataset, we assign counties to treatment timing based on whether and when a county produced a provincial top scorer. We then follow cohorts over time using birth year as the cohort dimension and measure educational outcomes observed at the 2010 census date. Because outcomes are observed at a common census cross-section, our analysis focuses on differences across cohorts within the same county and compares counties that experience top-scorer exposure at different times. We apply standard sample restrictions to ensure comparability across cohorts and to reduce confounding from selective migration, and we use the large census sample size to obtain precise estimates when aggregating to county-by-cohort cells.

For completeness, Table A1 reports additional descriptive statistics that summarize baseline county characteristics by the timing of the first appearance of a provincial top scorer. These tables document the staggered and spatially dispersed nature of top-scorer realizations and the heterogeneity in pre-period county conditions, and are intended to provide background context for the empirical analysis rather than to assess identifying assumptions.

3.3. Sample Construction

Our analysis combines a province-level Gaokao top-scorer compilation covering 1990–2005 with educational attainment outcomes measured in the 2010 population census microdata. The choice of 1990 as the start year reflects four considerations. First, the local media infrastructure that made provincial top scorers publicly salient to nearby cohorts, including local newspapers, school honor rolls, and provincial radio and television coverage, only became broadly and systematically active around the early 1990s; while sporadic publicity existed earlier, the kind of consistent, geographically targeted visibility that underlies our exposure mechanism is a feature of the post-1990 period (see Appendix 1). Second, systematic archival reporting of provincial top scorers becomes consistently available from 1990 onward, while pre-1990 records are too sparse to support reliable treatment timing. Third, the post-1990 window predates the major Gaokao reforms initiated in 2014–2017, which restructured subject testing and admissions tracks, and therefore covers a comparatively stable institutional environment. Fourth, ending the compilation in 2005 ensures that the youngest exposed cohort, born around 1987, is at least 23 years old at the 2010 census, so that four-year college completion can be observed for all exposed cohorts.

We restrict the main analysis cohorts to birth years 1965–1987. This cohort window guarantees a sufficiently long pre-event period even for counties whose first top-scorer appearance occurs in 1990, providing a meaningful baseline for assessing pre-trends, while ensuring that the youngest cohorts are old enough by 2010 to have plausibly completed four-year college.

We define a top-scorer county using the location of the top scorer's high school. Some counties experience multiple top-scorer events during 1990–2005. In the main specification, we retain only the first observed top-scorer year for each county so that treatment timing is uniquely defined by the county's initial exposure to a top-scorer signal. As a robustness check, we re-estimate the main specifications after dropping these multi-event counties entirely to verify that repeated top-scorer occurrences are not driving the results.

Our baseline identification uses counties that are not yet treated as the comparison group in each event time. In additional robustness checks, we expand the control group to include never-treated counties to assess sensitivity to alternative comparison sets.

After applying these restrictions, the final estimation sample contains 277 provincial top scorers, providing substantial spatial and temporal variation in event timing across counties. Appendix Table A2 reports detailed information for each top scorer used in the analysis.

4. Empirical Strategy

We estimate the causal effect of county-level exposure to a provincial Gaokao top scorer on the educational attainment of later birth cohorts using a staggered difference-in-differences design. The unit of analysis is a county-by-birth-cohort pseudo-panel constructed from the 2010 Population Census. Let c index counties and t index birth cohorts. Let Y_{ct} denote a county-by-cohort outcome, such as mean years of schooling or the share completing high school or attaining higher education. We include a vector of county-by-cohort controls X_{ct}, weight each county-by-cohort cell by its size N_{ct}, and cluster standard errors at the county level. A stylized belief-updating framework in Appendix 4 rationalizes our staggered DID design by modeling a top-scorer realization as a locally salient signal that shifts perceived returns to educational effort and can generate differential responses along extensive (staying in school) and intensive (upgrading from three-year to four-year college) margins.

4.1. Exposure timing and treatment definition

For each county c, let T_c denote the calendar year in which the county first produces a provincial Gaokao top scorer, as identified from the top-scorer database and the school-to-county mapping described in Section 3.1. Since the Gaokao is taken at the end of senior high school and examinees are typically around age 18, we translate the event year into a cohort-level exposure timing as

$$G_c \equiv T_c - 18.$$

This mapping interprets G_c as the approximate birth cohort of the top scorer in county c and aligns treatment timing with the schooling lifecycle. The cohorts most plausibly affected by the top-scorer signal are those strictly younger than the top-scorer cohort, who may update aspirations and schooling decisions while still enrolled. Accordingly, we define the post-exposure indicator as

$$D_{ct} \equiv \mathbb{1}[t \geq G_c + 1],$$

so that cohorts born in $G_c + 1$ and later are classified as post-exposure for county c, while cohorts born in G_c and earlier form the pre-exposure period.

4.2. CSDID estimation: baseline pre–post and event-study specifications

Our main estimates are based on the group-time average treatment effects framework of Callaway and Sant'Anna (2021), which we abbreviate CSDID throughout the paper. The estimator is designed for staggered adoption and permits treatment effects to vary across exposure cohorts and event times. The identifying variation comes from comparing cohorts within a county before and after exposure, relative to cohorts in counties that have not yet been exposed at the same cohort date (and, when applicable, counties that are never exposed during the sample window). Identification relies on a conditional parallel trends assumption: absent exposure, outcomes for counties first exposed at a given time would have evolved in parallel with outcomes in the comparison group, conditional on X_{ct}.

Baseline CSDID estimator. We first summarize the overall impact of exposure using a pre–post specification with county and cohort fixed effects:

$$Y_{ct} = \alpha_c + \lambda_t + \beta\, D_{ct} + X_{ct}'\gamma + \varepsilon_{ct}, \tag{1}$$

where α_c are county fixed effects absorbing time-invariant differences across counties and λ_t are cohort fixed effects capturing common cohort-level shifts in educational attainment. The coefficient β summarizes the average post-exposure change in Y_{ct} for exposed counties relative to the contemporaneous evolution of the comparison group. In the staggered setting, we interpret (1) as a convenient representation of an aggregated effect constructed from the underlying group-time ATTs in Callaway and Sant'Anna (2021), rather than as a literal two-way fixed effects estimator.

Event-study specification. To characterize dynamics and assess identifying assumptions, we estimate an event-study version that allows effects to vary with event time:

$$Y_{ct} = \alpha_c + \lambda_t + \sum_{k \in \mathcal{K}} \beta_k \cdot \mathbb{1}[t - (G_c + 1) = k] + X_{ct}'\gamma + \varepsilon_{ct}, \tag{2}$$

where k indexes event time relative to the exposure cohort G_c. We define event time relative to the county's first provincial top scorer. Event time k is indexed by birth cohorts, with $k = -1$ denoting the top scorer's own birth cohort, and $k = 0$ denoting the first cohort potentially exposed to the top-scorer signal. Estimates for leads ($k < 0$) provide a diagnostic for pre-trends and anticipation, while lags ($k \geq 0$) trace the evolution of exposure effects after a county produces a top scorer.

In our baseline construction, the county index c refers to an individual's county of current residence as recorded in the 2010 population census. We assign individuals to this county and aggregate outcomes to the county-by-birth-cohort level, so Y_{ct} captures educational attainment of cohort t among residents of county c in the census year. Because residence may reflect post-education migration, we conduct a key robustness check restricting the sample to "non-movers," defined as individuals whose registered *hukou* county matches their current county of residence. Results in Section 6.7 are consistent with the baseline estimates, suggesting that selective migration is unlikely to drive our findings.

Throughout, estimation and aggregation follow the Callaway and Sant'Anna (2021) CSDID procedure to accommodate staggered adoption and heterogeneous effects. Inference is based on county-clustered standard errors, and we assess robustness to alternative event windows, comparison-group definitions, and weighting schemes.

5. Results

5.1. Main Result

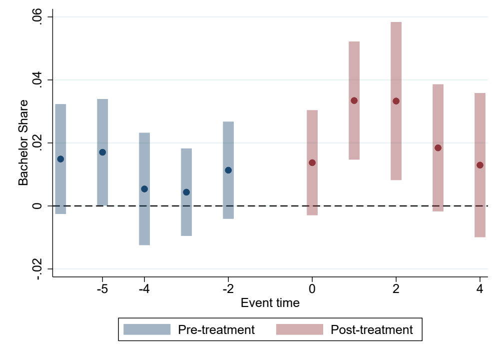
Figure 2 shows how educational attainment evolves around the first provincial top scorer in a county. Prior to exposure, the estimates hover around zero across outcomes, and the average lead effects reported in Table 1 are close to zero, indicating no clear pre-existing divergence (all four joint tests for the lead coefficients yield p-values above 0.9). After exposure, educational attainment rises. For the first post-exposure cohort ($k = 0$), years of schooling increase by about 0.18 years and high school completion rises by 3.4 percentage points. For the second post-exposure cohort ($k = 1$), the gains are larger: years of schooling increases by 0.354 years, about four additional months of education, while high school completion rises by 5.6 percentage points. Tertiary attainment also increases: the share completing any higher education (including 4-year bachelor's programs and 3-year junior-college programs) rises by about 4.2 percentage points, and the share completing a four-year university degree rises by about 3.9 percentage points. At more distant post-exposure cohorts, the estimated effects attenuate. A natural interpretation is that the College Entrance Exam top-scorer signal is most relevant for cohorts that are only one to two years younger than the top scorer—students who are still in senior high school when the event occurs and can adjust effort and expectations at the margin. In contrast, for substantially younger cohorts, the treatment effect decrease because the information shock arrives too early relative to key schooling decisions and any behavioral response may dissipate before it can translate into measurable attainment outcomes.

One notable pattern is that the estimate at $k = 2$ is lower than those for subsequent post-exposure cohorts. This dip is plausibly consistent with our mechanism. Cohorts at $k = 2$ are plausibly in the summer immediately after the high-stakes middle-school exit exam when top-scorer results are publicized: they have already finished junior secondary school and cannot adjust effort to change the high school they will attend, and this transition summer is typically the lowest-pressure period in the secondary-school pipeline. As a result, the salience and motivational content of the signal should be weakest precisely for this cohort, leading to a smaller response.

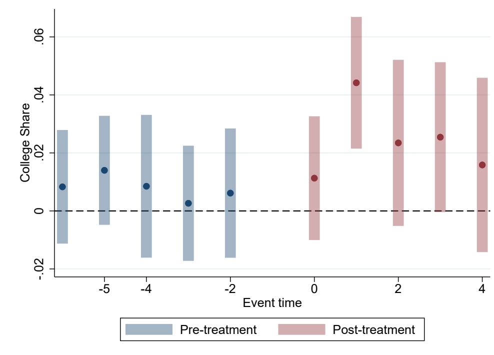
These magnitudes are meaningful because they reflect changes at distinct continuation margins and are concentrated among cohorts that are plausibly able to respond. The increases in college and four-year university degree completion suggest that the top-scorer signal primarily affects students who are already academically prepared and close to the higher-education admission threshold: the achievement provides a locally credible benchmark that can raise effort and persistence, helping some near-margin students convert relatively small performance gains into actual admission and eventual degree completion. At the same time, the rise in high school completion is more naturally interpreted as reduced dropout among students already enrolled in regular academic high school, especially those in Grade 10 and Grade 11 when the top scorer emerges, rather than as increased entry into high school, since the affected cohorts are already past the high-school entry margin. Consistent with this mechanism, the event-study estimates attenuate for substantially younger cohorts who are not yet in high school when the signal arrives, for whom the information shock is temporally distant from the key education decisions it is most likely to influence.

Moreover, the event-study dynamics help alleviate concerns that our estimates are driven by subsequent shifts in local education resources targeted to the top-scorer school or county. If the effects were primarily due to county- or higher-level resource reallocation, we would expect more persistent gains across a wider set of cohorts, rather than impacts that disappear quickly within two cohorts.

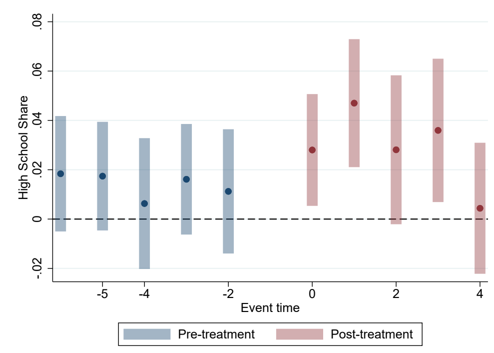
Figure 2: Event-Study Estimates of the Top-Scorer Exposure Effect



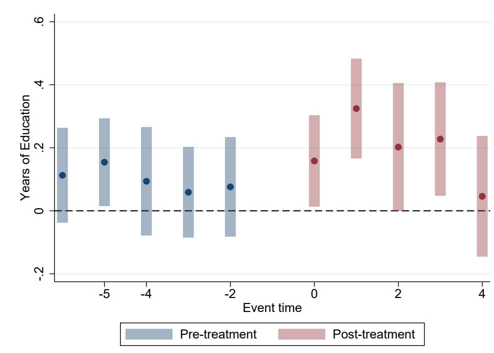
(a) Bachelor's degree



(b) Any higher education



(c) High school completion



(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with county-clustered standard errors. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

Table 1: Dynamic Effects of Exposure to a Provincial Top Scorer on Educational Attainment

Event time t	Schooling (yrs)	HS graduate	Any college	Bachelor's
$t = -5$	0.010	-0.001	-0.000	0.002
	(0.104)	(0.016)	(0.016)	(0.015)
$t = -4$	-0.037	-0.004	-0.007	-0.006
	(0.106)	(0.015)	(0.016)	(0.014)
$t = -3$	-0.019	0.010	-0.003	-0.004
	(0.091)	(0.014)	(0.015)	(0.010)
$t = -2$	0.001	0.000	-0.003	0.003
	(0.090)	(0.014)	(0.013)	(0.008)
$t = -1$	0.008	-0.001	0.007	-0.009
	(0.105)	(0.017)	(0.014)	(0.009)
$t = 0$	0.180**	0.034***	0.013	0.019*
	(0.082)	(0.013)	(0.013)	(0.010)
$t = 1$	0.354***	0.056***	0.042***	0.039***
	(0.097)	(0.017)	(0.014)	(0.012)
Short-run ATT ($t = 0, 1$)	0.267***	0.045***	0.027**	0.029***
	(0.075)	(0.012)	(0.012)	(0.009)
Pre-treatment mean	10.367	0.374	0.185	0.088
County-Cohort Obs	5093	5093	5093	5093
Number of individuals	225,183	225,183	225,183	225,183
Joint test F-value	0.26	0.64	0.57	1.43

Notes: The table reports CSDID event-study estimates for four outcomes: years of schooling, high school completion share, any higher education share (including 4-year bachelor's programs and 3-year junior-college programs), and 4-year bachelor's programs share. Standard errors (clustered at the county level) are in parentheses. The pre-treatment mean is computed using the same sample as the main specification—treated and not-yet-treated counties only—and is averaged over all pre-exposure cohorts ($k < 0$) at the county-by-cohort level. The event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

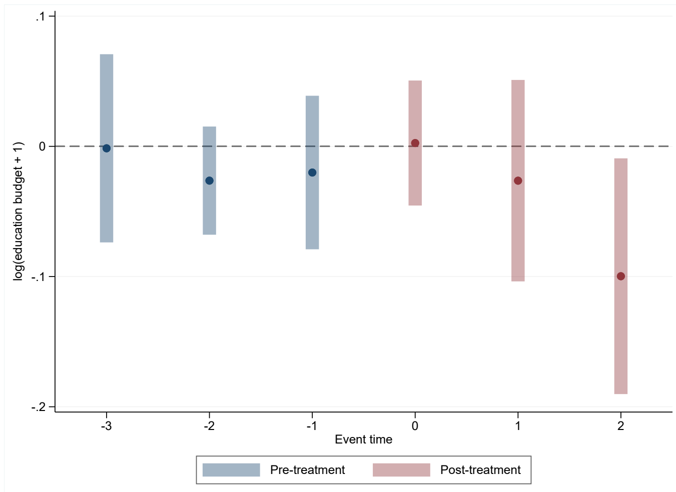
6. Robustness Checks

6.1. Robustness: Ruling Out Time-Varying Local Shocks and Fiscal Confounders

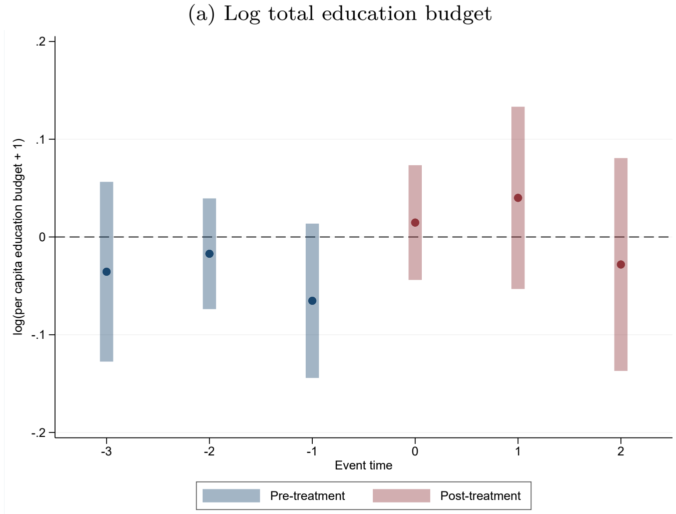
A natural concern is that provincial top scorers may themselves reflect underlying time-varying changes in local educational environments, rather than constituting an exogenous informational shock. For example, one might worry that improvements in school quality, resources, or broader educational investments could simultaneously raise student outcomes and increase the likelihood that a county produces a top scorer. Several features of our main results make this interpretation unlikely. In particular, across all four educational outcomes—high school completion, years of schooling, any college completion, and bachelor's degree completion—we find no immediate effect in the top-scorer year. Instead, outcomes rise only for cohorts that are one to two years younger than the top scorer, consistent with a delayed behavioral response rather than a contemporaneous shift in local educational conditions. If the emergence of a top scorer were driven by a sudden improvement in the local educational landscape, one would expect educational attainment to increase at the same time or even prior to the top-scorer year, which we do not observe.

To further address this concern, we directly examine whether local fiscal investments in education change around the appearance of a top scorer. Using county-level education expenditure data from the *China Education Expenditure Statistical Yearbook* (*Zhongguo Jiaoyu Jingfei Tongji Nianjian*), published by the Ministry of Education, which reports harmonized county-level fiscal measures over 1996–2002, we conduct an event-study analysis of education spending around the first top-scorer year (Department of Finance, Ministry of Education of the People's Republic of China and Department of Social, Science and Cultural Statistics, National Bureau of Statistics of China, 2002). Figure 3 shows that neither total education budgets nor per-capita education budgets exhibit systematic increases in the years preceding the top-scorer event. Moreover, we find no evidence of sustained increases in education spending following the appearance of a top scorer. These patterns rule out the possibility that top scorers are generated by prior increases in local educational investment, or that governments respond to the emergence of a top scorer by subsequently expanding education budgets. Taken together, these results suggest that the effects documented in the main analysis are unlikely to be driven by time-varying fiscal shocks at the county level.

To be clear, our fiscal event-study does not mechanically rule out all school-level supply responses, such as changes in tracking, elite classes, or teacher allocation, nor does it fully capture informal channels (donations, parental spending, or the tutoring market). However, the institutional context of secondary education in China makes a supply-driven interpretation unlikely. Senior high schools during our period were overwhelmingly public, with staffing, teacher compensation, and capacity largely determined by local education bureaus and financed through public budgets (Ministry of Education of the People's Republic of China, 2002). Consistent with this institutional setting, in our top-scorer dataset all observed provincial top scorers graduated from public senior high schools. Systematic improvements in instructional inputs—such as expanding elite-track capacity or recruiting additional teachers—would therefore be difficult to implement at scale without commensurate fiscal support, and would plausibly show up in county-level education expenditures.



(a) Log total education budget



(b) Log education budget per capita

Figure 3: Fiscal Responses Around the First Top-Scorer Event

Notes: This figure examines whether counties adjust education-related fiscal inputs around the first appearance of a provincial Gaokao top scorer. Panel (a) reports event-study estimates for the logarithm of total education budget (log(education budget + 1)), while Panel (b) shows corresponding estimates for the logarithm of per-capita education budget (log(education budget per capita + 1)). Estimates are obtained using the Callaway and Sant'Anna (2021) difference-in-differences estimator with not-yet-treated counties serving as controls. Event time is defined relative to the first year in which a county produces a top scorer ($t = 0$). Shaded areas denote 95% confidence intervals.

6.2. Robustness: Mitigating Bias from Non-random "High-Ability" Counties

A concern for interpretation is that counties producing multiple top scorers may be systematically different—for example, they may have unusually strong schools, sustained resource investments, or other persistent advantages, so that our baseline estimates could partly reflect these underlying differences rather than the incremental role-model shock of a first top-scorer event. To address this concern, we conduct a robustness exercise in which we exclude all counties that produce two or more provincial top scorers during our sample period and retain only counties that experience exactly one top-scorer event within the 16-year window. We then re-estimate the same CSDID event-study specifications, identifying effects solely from variation in the timing of this single exposure across counties. This restriction mitigates endogeneity concerns stemming from persistent, unobserved county-level advantages that could jointly generate repeated top-scorer realizations and higher subsequent attainment.

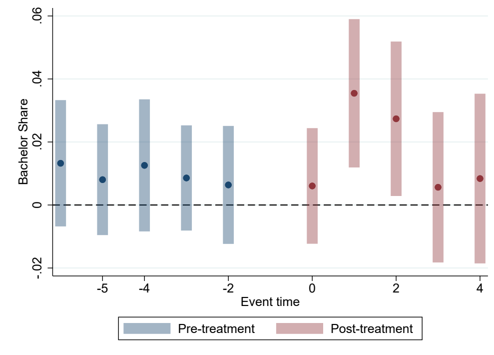
Figure A1 reports the resulting event-study estimates. The dynamic patterns and magnitudes closely track the baseline results: pre-exposure coefficients remain small and statistically indistinguishable from zero, while post-exposure effects exhibit a similar evolution across outcomes. Restricting the sample to "one-time" top-scorer counties therefore yields estimates that are highly consistent with the main findings, suggesting that repeated top-scorer counties are not driving our baseline results. While this exercise cannot fully rule out transitory, one-time supply-side shocks that coincide with the emergence of a first top scorer, the absence of pre-trends and the cohort-specific timing and attenuation of the effects are more consistent with a localized information or belief-updating shock than with sustained changes in educational inputs.

6.3. Robustness to Hukou-Specific Unobservables

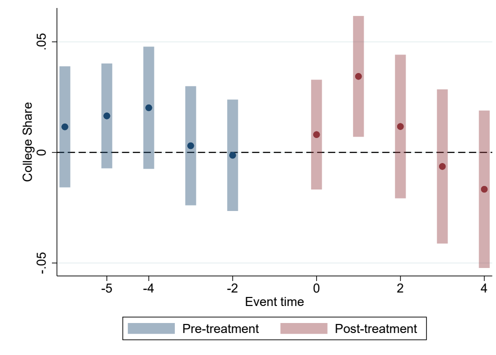
In our baseline analysis, we construct a county-by-cohort pseudo-panel and control for the rural population share to account for demographic differences across regions. However, one might be concerned that educational trajectories and responses to local role models systematically differ between agricultural and non-agricultural populations due to unobserved, time-invariant institutional or cultural factors within a county.

To address this concern and rule out the possibility that our findings are driven by underlying compositional changes in the hukou population, we construct a more granular pseudo-panel at the county-by-hukou-by-cohort level. This specification allows us to explicitly incorporate county-by-hukou fixed effects, thereby absorbing any time-invariant unobservables specific to the rural or urban populations within each county.

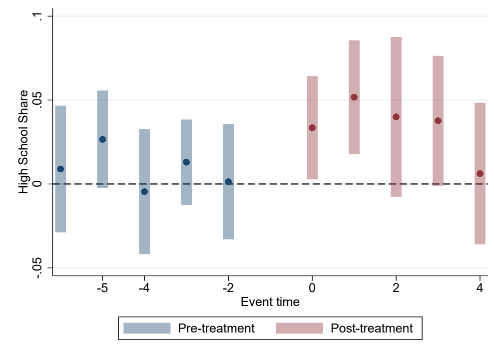
Figure 4 presents the event-study estimates from this alternative specification. Reassuringly, the results remain remarkably robust to the inclusion of the hukou dimension. The pre-treatment coefficients remain statistically indistinguishable from zero, confirming that the parallel trends assumption holds within hukou-specific populations. More importantly, the post-treatment dynamics for years of schooling, high school completion, and higher education are highly consistent with our baseline findings both in magnitude and statistical significance. This confirms that our main conclusions are not driven by underlying time-invariant differences across hukou types.



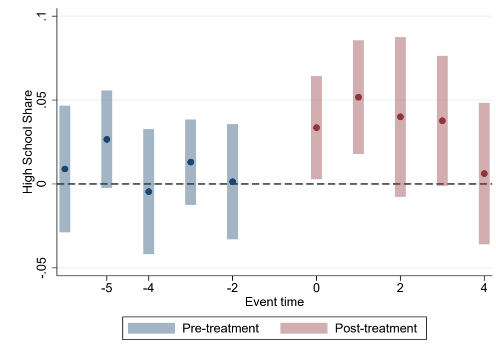
(a) Bachelor's degree



(b) Any higher education



(c) High school completion



(d) Years of schooling

Figure 4: Event-Study Estimates with County-by-Hukou Fixed Effects

Notes: This figure plots event-study estimates from the CSDID framework using an alternative county-by-hukou-by-cohort pseudo-panel. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county-by-hukou fixed effects, cohort fixed effects, and control for the Han ethnicity rate, with standard errors clustered at the county level. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

6.4. Robustness: Addressing Incomplete Early Coverage of Top-Scorer Records

A practical concern with the top-scorer record is that coverage is less complete in earlier years. For some provinces, archival sources are more difficult to retrieve, and top-scorer records from the early 1990s (especially 1990–1995) may be missing due to data availability rather than true absence of top scorers. To assess whether our main findings are sensitive to this potential measurement issue, we conduct a robustness check that restricts the top-scorer sample to the period 1995–2005, when reporting is substantially more consistent and coverage is higher across provinces. We reconstruct county-level exposure timing using only top-scorer events in 1995–2005 and re-estimate the same CSDID event-study specifications on the county-by-cohort pseudo-panel.

Figure A2 presents the resulting event-study estimates. The estimated dynamic patterns and magnitudes are broadly similar to the baseline results, with no evidence of meaningful pre-trends and comparable post-exposure responses across outcomes. Overall, restricting attention to the 1995–2005 top-scorer records yields results that are consistent with our main findings, suggesting that incomplete coverage in the early 1990s is unlikely to be driving the baseline estimates.

6.5. Robustness: Including Never-Treated Counties in the Control Group

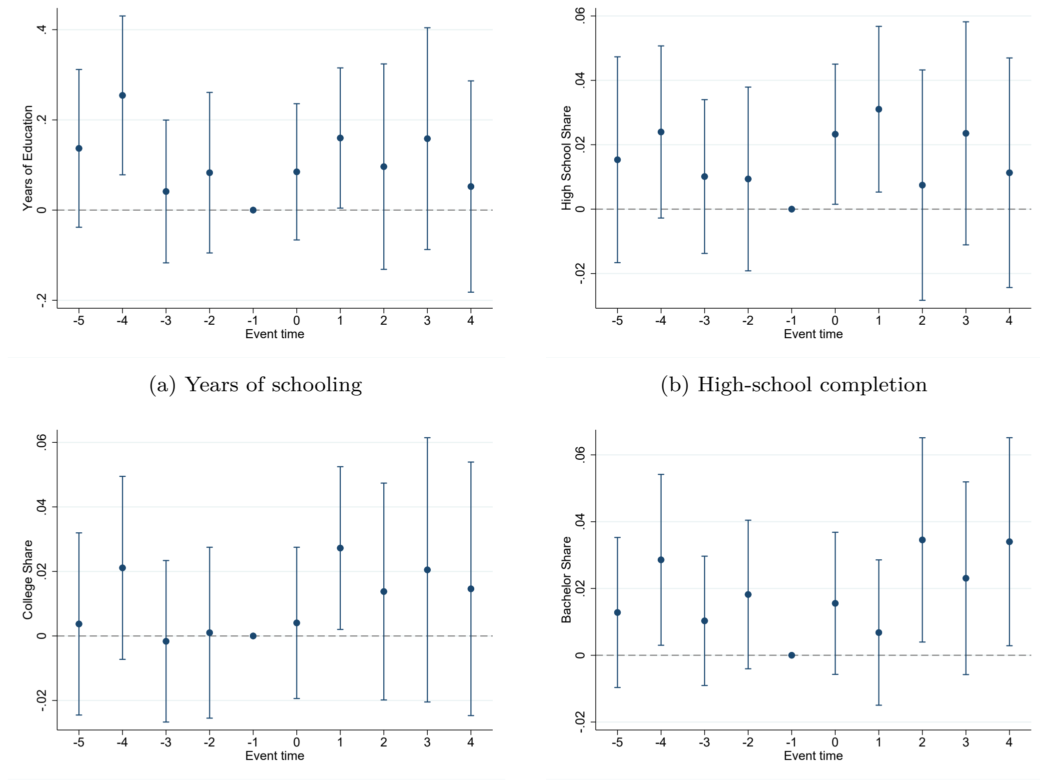
Our baseline CSDID estimates primarily use not-yet-treated counties as the comparison group, which ensures that treated and control units are drawn from the same set of eventually treated counties at each cohort. As an additional robustness check, we expand the control group to include counties that never produce a provincial top scorer during our sample period. Intuitively, this exercise asks whether the estimated exposure effects are sensitive to using a broader set of untreated counties as counterfactuals.

Figure A3 reports the corresponding event-study estimates. Relative to the baseline specification, the estimated post-exposure effects are larger in magnitude across outcomes, while pre-exposure coefficients remain close to zero. A natural interpretation is that never-treated counties differ systematically from eventually treated counties in baseline educational resources or long-run attainment trends, so including them as controls increases the contrast between exposed counties and the comparison group. Importantly, the qualitative time profile of the estimates remains similar, and the larger magnitudes reinforce the conclusion that exposure to a top-scorer event is associated with meaningful gains in educational attainment for subsequent cohorts.

6.6. Robustness: Within-City Comparison Using City-by-Cohort Fixed Effects

A natural concern is that our staggered DID estimates may capture province-level or prefectural-level shocks that happen to coincide with the timing of top-scorer events. To address this, we re-estimate the event study after adding city-by-cohort fixed effects to the baseline specification. Identification then comes from comparing students in counties first exposed to a top scorer to students in other counties within the same prefecture-level city in the same birth cohort, eliminating any variation common to a city in a given cohort. Two practical changes follow from this design. First, we move from the Callaway-Sant'Anna estimator to a TWFE event-study estimator, because the Callaway-Sant'Anna procedure does not accommodate the additional city-by-cohort fixed effects. Second, we retain all never-treated counties in the comparison group at every event time rather than restricting to not-yet-treated comparisons, which substantially enlarges the within-city pool of untreated counties available for identification.

Figure 5 reports the event-study coefficients for the four outcomes. Pre-treatment coefficients are statistically insignificant for all four outcomes. Post-treatment effects track the baseline closely: years of schooling rise across the post-treatment window, with the contemporaneous and short-horizon coefficients statistically significant; high-school completion and college completion rises immediately following treatment. The within-city specification therefore reproduces the main result and rules out city-level shocks coincident with top-scorer timing as the source of the baseline estimates.



(a) Years of schooling

(b) High-school completion

(c) College completion

(d) Four-year college completion

Figure 5: Event study with county and city-by-cohort fixed effects, by outcome. Sample restricted to cohorts born between 1968 and 1985, excluding Henan, Shandong, Guangdong, Hunan, and Jiangxi. Vertical bars are 95% confidence intervals based on standard errors clustered at the county level.

6.7. Robustness: Addressing Exposure Misclassification from Migration and Schooling Mobility

To assess whether selective migration biases our county-by-cohort estimates, we restrict the census sample to individuals whose hukou remains in their original county and re-estimate the CSDID event-study specifications. Figure A4 shows that the dynamic profiles are broadly similar to the baseline results, suggesting that differential migration is unlikely to drive our findings. This exercise also addresses potential exposure misclassification arising when census county identifiers (residence or hukou) do not perfectly coincide with the county of senior-high-school attendance. Cross-county high-school attendance may occur, but it was not widespread at the county level for cohorts in the 1990s–2005 period. Moreover, even if a top-scorer county attracts students from other counties, those students are enumerated in the census in their own hukou/residence counties rather than in the destination county, so such reallocation cannot mechanically raise measured educational attainment in the top-scorer county.

6.8. Robustness: Sensitivity to Clustering of Standard Errors

In the baseline analysis, we report standard errors clustered at the county level, which is the unit of treatment assignment and the level at which the county-by-cohort outcomes are constructed. A remaining concern is that shocks to education outcomes may be correlated across counties within the same prefecture-level city or province, for example through common education policies, labor-market conditions, or migration networks, so county-level clustering could understate sampling uncertainty.

To assess the sensitivity of inference to alternative correlation structures, we re-compute standard errors clustering at broader geographic levels. Specifically, we re-estimate the event-study specifications using (i) city-level clustering (Figure A5) and (ii) province-level clustering (Figure A6). The resulting confidence intervals and statistical conclusions are essentially unchanged relative to the county-clustered baseline, and the event-study profiles remain qualitatively identical across clustering choices. Overall, our main findings are not sensitive to the level at which standard errors are clustered.

7. Heterogeneity

7.1. Heterogeneity by Top-Scorer and Cohort Gender

We next ask whether the educational response to a provincial top-scorer event depends on the gender of the role model and on the gender of the affected cohort. This is a meaningful robustness exercise for two reasons.

First, two of the closest comparison papers in our literature, Beaman et al. (2012) and Riise et al. (2022), document role-model effects that are concentrated in same-gender pairings: female leaders raise the aspirations of adolescent girls, and female doctors raise the choice of medical and other STEMM occupations among female students. If our setting operated through a similar same-gender channel, we would expect the post-treatment response to be larger when the gender of the affected cohort matches the gender of the top scorer. Second, finding that the response varies meaningfully across the four scorer-by-cohort cells would speak directly to the role-model framing of the paper, since a purely informational story about returns to schooling has no obvious reason to produce gender-differentiated effects.

To run this exercise we need scorer-level gender information. The hand-collected database used elsewhere in the paper records gender for a subset of scorers, primarily those documented in standardized provincial summary lists. For the remaining scorers we infer gender from given-name characters through a two-track procedure that combines manual classification by the authors with an independent classification by a large language model, accepting an assignment only when both tracks agree and only when the distinguishing characters carry strong gender signal in modern Chinese given names. Names whose given characters are gender-neutral or where the two tracks disagree are left as missing rather than forced into a category.

Table 2 reports the resulting simple ATT estimations. Two patterns are visible in the post-treatment estimates. First, when we stratify within male-scorer counties by the cohort's gender, the response is concentrated in the female cohort: the high-school completion effect rises to 4.4 percentage points and the four-year college completion effect rises to 4.6 percentage points, both significant at conventional levels, while the male-cohort estimates are smaller and reach only marginal significance for high-school completion. Second, in female-scorer counties the male-cohort estimates are essentially zero or even slightly negative, while the female-cohort estimates are positive and marginally significant for any-college and four-year college completion at 5.3 and 4.3 percentage points respectively.

The picture that emerges is that the responsive group in our setting is the female cohort, and that the gender of the role model matters less than whether the affected cohort responds at all. Women appear to absorb the salient academic-success signal in both same-gender and opposite-gender cases, while men in this context show essentially no measurable response to either kind of role model. This finding is consistent with the Beaman et al. and Riise et al. evidence in the sense that locally identifiable academic achievement moves girls' attainment, and it is consistent with the broader literature on the asymmetry of educational expectations during this period in China, where girls had more headroom on the upper-tier credentials and were therefore better positioned to convert a salience shock into completed education.

Table 2: Heterogeneous Effects by Top-Scorer Gender and Cohort Gender

		Years of Schooling	High School Completion	Any College ($\geq$ 15 years)	Four-Year College ($\geq$ 16 years)
Panel A: Male top-scorer counties					
Male cohorts	Pre	−0.011	−0.004	0.004	−0.004
		(0.034)	(0.005)	(0.008)	(0.006)
	ATT	0.097	0.034*	0.007	0.009
		(0.123)	(0.019)	(0.019)	(0.014)
Female cohorts	Pre	−0.010	−0.002	−0.001	−0.004
		(0.034)	(0.005)	(0.005)	(0.003)
	ATT	0.214	0.044**	0.021	0.046***
		(0.140)	(0.022)	(0.021)	(0.016)
Panel B: Female top-scorer counties					
Male cohorts	Pre	0.068	0.005	0.014	−0.001
		(0.112)	(0.015)	(0.025)	(0.012)
	ATT	−0.102	−0.002	−0.040	0.021
		(0.233)	(0.025)	(0.044)	(0.029)
Female cohorts	Pre	−0.130	−0.016	−0.022	−0.008
		(0.116)	(0.020)	(0.016)	(0.010)
	ATT	0.290	0.010	0.053*	0.043*
		(0.229)	(0.041)	(0.032)	(0.026)

Notes: The table reports the simple overall ATT from the Callaway–Sant'Anna staggered DiD estimator, with pre-treatment averages from the same event-study reported as parallel-trends evidence. Counties are classified by the gender of their earliest provincial top scorer, with mixed-gender and unassigned counties excluded. Cohorts are classified by individual gender from the 2010 Census. The all-cohorts row within Panel B is omitted because the small number of female-scorer counties combined with the male-share covariate produces unstable estimates in the combined specification. All regressions include county-level covariates; standard errors clustered at the county level are reported in parentheses. $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

7.2. Heterogeneity by Baseline Educational Attainment

We next examine whether the effects of provincial top-scorer exposure vary with counties' pre-existing educational attainment. Baseline attainment is measured using cohorts who reached Gaokao age prior to 1990, whose educational outcomes were fully realized before the start of the sample period.

Counties are split into low- and high-baseline groups according to whether their pre-1990 high school completion rate lies below or above the median.

Table 3 reports the average pre-treatment and post-treatment coefficients from the CSDID event-study specification, estimated separately for each baseline group and each outcome. The pre-treatment averages are statistically indistinguishable from zero across all eight specifications, indicating that parallel trends are broadly satisfied within each subsample.

The post-treatment averages reveal two distinct patterns. In counties with high baseline attainment, the effect concentrates at the top of the educational distribution: four-year college completion rises by 5.4 percentage points and is significant at the one-percent level, while effects at lower margins are positive but not statistically distinguishable from zero. In counties with low baseline attainment, effects appear at both college and four-year college completion: any-college attainment rises by 3.4 percentage points and is significant at the ten-percent level, and four-year college completion rises by 2.4 percentage points and is significant at the five-percent level, with smaller and less precise effects at earlier margins.

A natural interpretation is that the role-model channel operates differently depending on local capacity to convert the response into completed attainment. In counties with stronger pre-existing educational infrastructure, the salience of a top-scorer event is most easily translated into the highest credential, namely a four-year university degree. In counties starting from a lower base, the signal raises attainment across multiple tiers of higher education, but magnitudes at the top tier are smaller, consistent with binding capacity constraints on the marginal student. The fact that earlier margins are positive in both groups but only weakly identified suggests that the response operates predominantly through tertiary rather than secondary schooling decisions, in line with the schooling-lifecycle interpretation in Section 4.1.

Table 3: Heterogeneous Effects by Baseline Bachelor's Degree Completion

	Years of Schooling	High School Completion	Any College	Four-Year College
Panel A: High-baseline counties				
Post	0.203	0.019	0.018	0.054***
	(0.138)	(0.019)	(0.020)	(0.019)
Panel B: Low-baseline counties				
Post	0.142	0.028	0.034*	0.024**
	(0.129)	(0.020)	(0.019)	(0.009)

Notes: The table reports averages of pre-treatment and post-treatment coefficients from the CSDID event-study specification, estimated separately for each subsample. Counties are classified as high- or low-baseline based on whether their pre-1990 high school completion rate, lies above or below the median. All specifications include county-level covariates, and standard errors clustered at the county level are reported in parentheses. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

7.3. Heterogeneity by Comparison Group

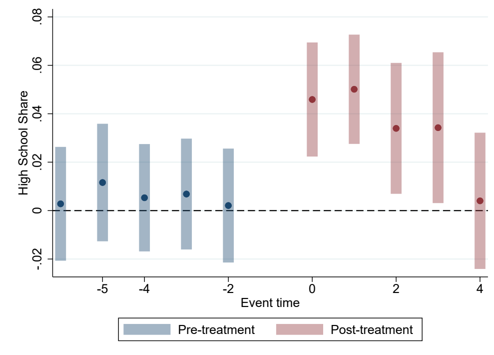
We also present a complementary specification that uses a different comparison group: rather than all not-yet-treated counties, we use only counties in the same prefecture-level city as a treated county. This exercise serves three purposes. It documents that the estimates hold when the comparison group is geographically and administratively proximate to the treated counties; it tests whether treated counties differ systematically from far-away peers, since replacing distant untreated peers with same-city untreated peers would shift the estimates if such selection were operative; and it probes for spatial spillovers within the city, since positive spillovers from a top-scorer realization onto untreated peers in the same city would attenuate the within-prefecture estimate relative to the baseline.

We emphasize that this restriction does not, by itself, absorb city-level shocks in the strict statistical sense, because the specification continues to use our baseline staggered DiD estimator without city-by-cohort fixed effects;

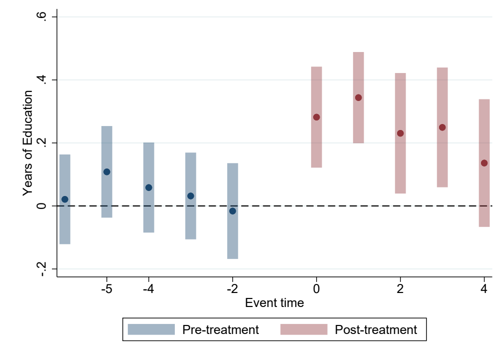
that exercise is reported separately in Section 6.6. What the restriction accomplishes is the use of within-city variation in the timing of top-scorer realizations to identify the effect, so that the identifying variation comes from differential treatment timing within a relatively homogeneous regional environment rather than from cross-region differences in trajectories.

Figure 6 reports the resulting event-study estimates. Pre-treatment coefficients remain statistically indistinguishable from zero, and post-treatment estimates remain positive and similar in magnitude to the baseline. The absence of meaningful attenuation relative to the baseline indicates that our estimates are not produced by selection of treated counties from regions with systematically different educational trajectories, and is consistent with a localized response that does not propagate to nearby untreated counties within the same prefecture-level city.

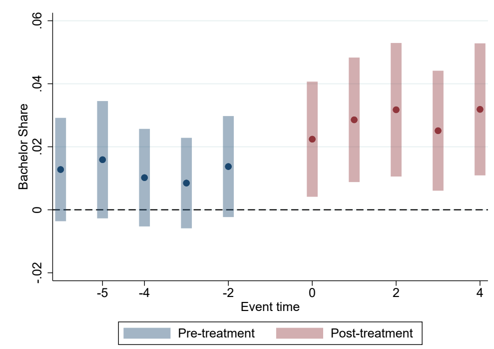
Figure 6: Event-Study Estimates with Within-Prefecture Comparison Group



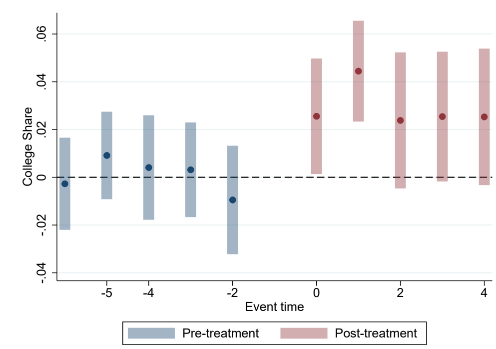
(a) High school completion



(b) Years of schooling



(c) Bachelor's degree



(d) Any higher education

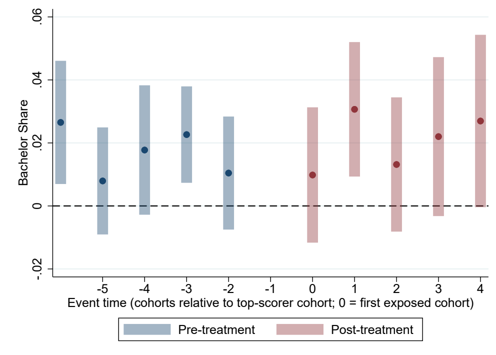
Notes: Each panel plots event-study estimates from the CSDID framework, estimated on the restricted sample of counties in prefecture-level cities that produced at least one provincial top scorer during the sample period. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with standard errors clustered at the county level. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county; $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after the top scorer.

7.4. Heterogeneity by Track: Humanities-Track Top Scorers

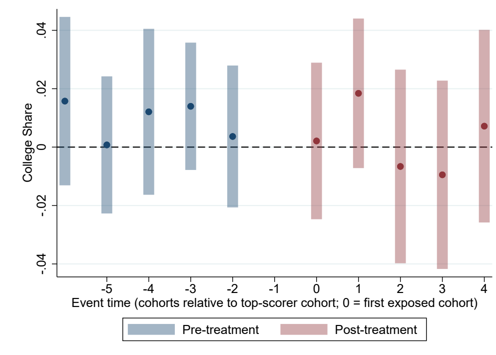
To examine heterogeneity by academic track, we re-define treatment using humanities-track top scorers only. We reconstruct the county-level exposure year as the first year in which a county produces a humanities-track top scorer, map it into an exposure cohort using the same age-based rule, and re-estimate the event-study specifications within the same CSDID framework. Figure 7 reports the resulting event-study estimates. In contrast to the baseline results, the estimated effects associated with humanities-track top-scorer exposure are close to zero and statistically insignificant across event times.

This pattern is consistent with institutional features of China's high school tracking environment that can generate substantively weaker spillovers from humanities-track top scorers. First, humanities-track top scorers typically receive substantially less media attention and public celebration than science-track top scorers, implying a less widely disseminated information signal at the county level. Second, the humanities track accounts for a relatively small share of students in regular academic high schools (often well below one quarter), and in practice it disproportionately includes students with weaker academic preparation who sort into the humanities track under competitive pressure. Together, these features imply that even if a humanities-track top scorer shifts aspirations within a narrow subgroup, the aggregate county-by-cohort outcomes we study need not respond detectably. The contrast between the baseline estimates and the humanities-specific estimates therefore suggests that the top-scorer effect is heterogeneous across tracks and depends on both the visibility of the achievement and the size and composition of the potentially exposed peer group.

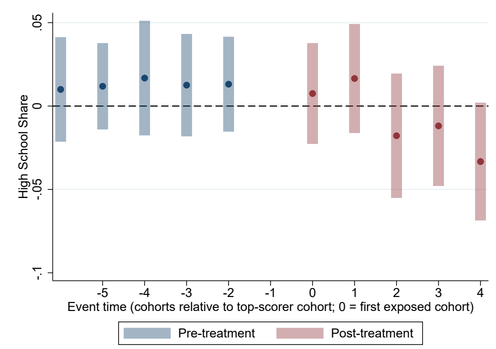
Figure 7: Event-Study Estimates of the Humanities-track Top-Scorer Exposure Effect



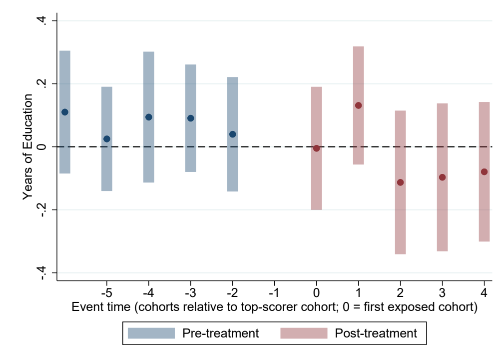
(a) Bachelor's degree



(b) Any higher education



(c) High school completion



(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with county-clustered standard errors. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

8. Discussion

This paper speaks to a recurring question in development economics and the economics of education: whether low-cost, recognition-based instruments can meaningfully affect human-capital accumulation, and how they compare to conventional input- or incentive-based approaches. Our estimates indicate that the emergence of a locally celebrated academic achievement, namely a provincial Gaokao top scorer, is associated with sizable changes in schooling outcomes for nearby cohorts, especially those close to key continuation margins. These patterns map naturally to a class of policy instruments that operate through public recognition and information disclosure, including the publicizing of academic rankings, honor rolls, and exemplary student cases. Read together with the existing literature on locally identified role models (Beaman et al., 2012; Riley, 2024; Riise et al., 2022), our findings are consistent with the view that recognition-based policies can complement traditional educational investments. They do not, however, settle the question of which behavioral channel is doing the work: our design identifies a reduced-form effect rather than the underlying mechanism.

The magnitude of the estimated effects underscores their potential policy relevance. Averaging over the first two post-exposure cohorts, exposure increases completed schooling by 0.267 years (Table 1). A conservative back-of-the-envelope translation using a 6–10% return to schooling per additional year implies a 1.6–2.7% increase in lifetime earnings for exposed individuals, based on global estimates of private returns to schooling reported by Montenegro and Patrinos (2014). Because education generates fiscal and social externalities, these private earnings gains plausibly translate into public returns through higher tax revenues and social-insurance contributions, even abstracting from other non-pecuniary benefits. Importantly, the underlying intervention in our setting, namely public recognition and dissemination of a top-scorer outcome, requires very limited direct public spending relative to standard education policies that rely on school construction, staffing, or financial incentives. Even under conservative assumptions, the implied benefit–cost ratios are therefore likely to be substantial. At the same time, our heterogeneity results indicate that these gains are larger in areas with stronger baseline educational capacity, suggesting that information-based policies are most effective when paired with institutions capable of converting increased effort into completed attainment.

Our evidence is informative about mechanisms, although indirectly. Several patterns are difficult to reconcile with a pure supply-side explanation. First, we find no detectable pre-treatment trends in the event-study leads, and the effects concentrate sharply among cohorts one to two years younger than the top scorer. This timing pattern is more consistent with a behavioral adjustment by students still able to alter effort and persistence than with gradual changes in school inputs. Second, the rapid attenuation of effects for more distant cohorts contrasts with what would be expected from sustained improvements in school quality or resources. Third, within-city comparisons and robustness exercises that exclude counties with repeated top-scorer realizations further reduce the scope for persistent supply-side advantages to drive the results. Taken together, these features are more consistent with a role-model or aspirations channel than with durable changes in educational production. Distinguishing among the specific behavioral mechanisms—belief updating about returns to schooling, aspiration-based shifts in perceived attainability, or peer-based emulation—would require data that we do not have.

An additional public-economics concern is displacement in a rank-based admission system. Because Gaokao admissions are largely determined by within-province rank, increases in college or bachelor's completion in winning counties could partly crowd out students from other counties, so our earnings-based welfare calculations should be interpreted as an upper bound in the presence of such competition. By contrast, our estimates for high school completion and years of schooling capture net increases in educational persistence and human-capital accumulation: higher high school completion reflects real skill acquisition with independent economic value, not a mechanical reallocation of fixed college slots.

A central limitation of our analysis is that we cannot directly observe the precise pathways through which the achievement is transmitted to nearby students. We are unable to distinguish whether the response operates primarily through local media coverage, teacher communication, school-level announcements, peer networks, or informal social interactions, nor whether the underlying cognitive shift is best characterized as a change in beliefs about returns to schooling or as an aspiration-based shift in perceived attainability. Our estimates should therefore be interpreted as the reduced-form effect of a bundled exposure environment rather than the impact of any single channel. Future work that links administrative education records with measures of media exposure, school communication practices, or peer networks would be valuable in isolating the most effective mechanisms and in designing more targeted recognition policies.

9. Conclusion

This paper asks whether locally celebrated academic achievement can shape human-capital investment. We exploit the staggered, rare emergence of provincial Gaokao top scorers across Chinese counties as a quasi-idiosyncratic local exposure event and link exposure timing to educational outcomes constructed from census-based county-by-cohort panels. Event-study and staggered difference-in-differences estimates show that cohorts exposed in the relevant schooling window experience meaningful gains in educational attainment: dropout declines, years of schooling increase, and completion of upper-secondary and tertiary education rises. The effects are concentrated among cohorts that are close enough to the event to plausibly adjust effort and persistence, and the increase in four-year university degree completion is particularly robust relative to shorter-lived gains in overall attainment.

The evidence is consistent with a role-model channel through which a salient, publicly identified achievement shifts the schooling decisions of near-margin students. Two patterns strengthen this interpretation. First, effects are larger in counties with stronger pre-existing educational capacity, suggesting complementarity between the exposure event and the local capacity to convert behavioral responses into completed attainment. Second, redefining treatment using humanities-track top scorers yields little detectable response, consistent with lower salience and a narrower exposed peer group rather than a general change in county-level education policy or resources. Our design does not, however, identify the underlying behavioral mechanism: distinguishing between belief updating about returns to schooling, aspiration-based shifts in perceived attainability, and peer-based emulation would require data we do not have.

From a public-economics perspective, the findings indicate that socially salient signals can generate spillovers in human-capital investment even in the absence of formal policy reform. At the same time, the capacity complementarity implies that recognition-based interventions, such as the public dissemination of credible success stories, are unlikely to substitute for investments in educational infrastructure and may deliver uneven benefits across places. An important direction for future research is to measure the exposure channel more directly (for example, through local media intensity, parental beliefs, and student effort), to assess potential general-equilibrium responses by schools and local governments, and to evaluate longer-run and distributional consequences, including whether such events narrow or widen spatial and socioeconomic gaps in educational attainment.

References

Acemoglu, D. and Angrist, J. (2000). How large are human-capital externalities? evidence from compulsory schooling laws. *NBER macroeconomics annual*, 15:9–59.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *science*, 335(6068):582–586.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

China Central Television (CCTV) (2004). This year's national college entrance examination uses 15 provincial test versions. https://idm.cctv.com/education/20040607/100272.shtml. Reports official statistics on arts (wenke) and science (like) gaokao examinees.

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and aspirations failure. *The Economic Journal*, 126(590):165–188.

Davey, G., De Lian, C., and Higgins, L. (2007). The university entrance examination system in china. *Journal of further and Higher Education*, 31(4):385–396.

Department of Finance, Ministry of Education of the People's Republic of China and Department of Social, Science and Cultural Statistics, National Bureau of Statistics of China, editors (2002). *China Education Expenditure Statistical Yearbook*. County-level tables used for 1996–2002.

Duflo, E., Dupas, P., and 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.

Genicot, G. and Ray, D. (2017). Aspirations and inequality. *Econometrica*, 85(2):489–519.

Guangming Daily (2001). Starting from developed countries' shift toward "valuing humanities over science". https://www.gmw.cn/01wzb/2001-10/14/06-B2C7D4ADDFFF130648256AE500099776.htm. Discusses the widespread perception of zhongli qingwen (valuing science over humanities) in China.

Hoxby, C., Turner, S., et al. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, 12(014):7.

Lochner, L. and Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American economic review*, 94(1):155–189.

Ministry of Education of the People's Republic of China (2001). Notice on the admission of recommended students to regular higher education institutions in 2001. http://www.moe.gov.cn/jyb_xxgk/gk_gbgg/moe_0/moe_7/moe_12/tnull_5934.html. Specifies Olympiad-based and other merit-based channels for recommended (baosong) admissions.

Ministry of Education of the People's Republic of China (2002). Educational development statistical bulletin. Official statistical reports. 1996–2002.

Ministry of Education of the People's Republic of China (2004). Outline for education development and reform in china (1990–2000). https://cernet.edu.cn/edu/zheng_ce_gs_gui/zheng_ce_wen_jian/fa_zhan/zong_he/200603/t20060323_110458.shtml. Originally issued in 1993; archived version.

Ministry of Education of the People's Republic of China (2005). Educational development statistical bulletin. http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/. Official statistics on national education development.

Montenegro, C. E. and Patrinos, H. A. (2014). Comparable estimates of returns to schooling around the world. *World Bank policy research working paper*, (7020).

Moretti, E. (2004). Human capital externalities in cities. In *Handbook of regional and urban economics*, volume 4, pages 2243–2291. Elsevier.

National Bureau of Statistics of China (2010). Population census of the people's republic of china 2010: Anonymized microdata sample. Microdata. Anonymized individual-level census microdata produced by the National Bureau of Statistics of China.

National Bureau of Statistics of China (2012). *Tabulation on the 2010 Population Census of the People's Republic of China*. China Statistics Press, Beijing.

National Bureau of Statistics of China (2015). The compendium of township-level administrative divisions of the people's republic of china (2015). https://www.stats.gov.cn/zs/tjwh/tjkw/tjzl/202302/t20230215_1907832.html. Official publication by the National Bureau of Statistics.

Riise, J., Willage, B., and Willén, A. (2022). Can female doctors cure the gender stemm gap? evidence from exogenously assigned general practitioners. *The Review of Economics and Statistics*, 104(4):621–635.

Riley, E. (2024). Role models in movies: The impact of *Queen of Katwe* on students' educational attainment. *The Review of Economics and Statistics*, 106(2):334–349.

Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.

China Daily (2010). Top scorers as instant celebrities: banners and press conferences in china's secondary schools. News article. http://www.chinadaily.com.cn/.

China Daily (2018). Top scorers as a school-reputation tool. News article. http://www.chinadaily.com.cn/.

Global Times (2024). Reports flag continued "over-exposure" of gaokao top scorers. News article. https://www.globaltimes.cn/.

Ministry of Education of the People's Republic of China (2017). Reply to npc proposal: long-running policy treatment of gaokao top-scorer publicity. Official document.

Ministry of Education of the People's Republic of China (2020). Plan for deepening the reform of evaluation in education in the new era: prohibition on publicizing gaokao top-scorer rankings. Official document.

Ministry of Education of the People's Republic of China (2024). Annual notice on regular higher-education enrollment work: reiteration of the prohibition on publicizing gaokao top scorers. Official document.

State Council of the People's Republic of China (2024). Statistics on the 2024 national college entrance examination. Official statistics.

The China Project (2019). Front-page coverage of provincial gaokao top scorers in chinese local newspapers. Magazine article.

Azoulay, P., Graff Zivin, J. S., and Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics*, 125(2):549–589.

Bursztyn, L. and Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics*, 130(3):1329–1367.

Carrell, S. E., Page, M. E., and West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, 125(3):1101–1144.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly Journal of Economics*, 116(2):681–704.

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

Online Appendix

1. Documenting the Visibility of Provincial Top Scorers

This appendix collects four sets of facts that motivate our interpretation of the first appearance of a provincial Gaokao top scorer in a county as a salient, locally anchored event during the 1990–2005 sample period. The Gaokao is China's national college entrance examination, administered once a year at the end of senior high school and used as the primary basis for university admission throughout the country. A provincial top scorer is the student with the highest examination score within a given province, academic track, and year, and is publicly identified by name within days of the results release. We document the rarity of this outcome relative to the size of the examinee population, the channels through which the achievement was publicized in central and provincial press, the school-level reputational dynamics it triggered together with the post-1999 university recruitment competition, and the regulatory response that this publicity has eventually generated.

Rarity and scale.. The Gaokao is taken by an extraordinarily large number of students. During our sample period the total number of examinees ranged from approximately 2.8 million in 1990 to 3.8 million in 2000 (Ministry of Education of the People's Republic of China, 2005), and in recent years roughly 13.4 million students have sat for the examination annually (State Council, 2024). China had approximately 13,991 regular senior high schools in 1995 (Ministry of Education of the People's Republic of China, 2005) and around 2,862 county-level administrative units in 2005 (National Bureau of Statistics of China, 2015), implying roughly five senior high schools per county-level unit on average. Within this system each province typically identified one provincial top scorer per academic track per year, so the outcome is an extreme realization both at the national and at the local level.

Press coverage during the sample period.. Coverage of provincial top scorers throughout 1990–2005 was intensive and geographically anchored. The country's major national dailies, including *People's Daily*, the official organ of the Communist Party of China, *Guangming Daily*, oriented toward intellectuals and educators, and *China Youth Daily*, aimed at younger readers, carried annual summer features that named the year's provincial top scorers, identified their high schools and counties of origin, and frequently profiled their families and study habits. The Chinese senior high curriculum is divided into a science track and a humanities track at the start of the second year, so each province typically produces one provincial top scorer in each track per year. Major regional dailies based in large cities, including *Beijing Youth Daily* in Beijing, *Wenhui Daily* in Shanghai, *Yangcheng Evening News* in Guangzhou, *Chutian Metropolis Daily* in Wuhan, and *West China Metropolis Daily* in Chengdu, ran parallel coverage focused on local top scorers and their schools, typically on the front page of summer issues, and this coverage was reproduced and amplified by school newsletters, county education bulletins, and alumni networks throughout the sample window (The China Project, 2019; China Daily, 2010).

School-level reputation and the post-1999 recruitment competition.. Within schools, recognition of the provincial top scorer took the form of public banners, photographs displayed at the school entrance, and press conferences hosted by the school administration in the days immediately after results were released. The most visible cases, such as Huanggang High School in Hubei province in central China, built nationally recognized reputations through the 1990s on the basis of repeated provincial top-scorer outcomes. Coverage in central outlets and feature segments on China Central Television, the national state broadcaster, treated the school as a case study of what was then known as the "Huanggang phenomenon," and a parallel publishing wave produced widely circulated standardized practice books, marketed under the school's name and used by senior high students nationwide as Gaokao preparation during the late 1990s and early 2000s. Top-scorer publicity intensified within the sample period rather than diminished after the higher-education enrollment expansion of 1999, a central government policy that nearly doubled incoming undergraduate cohorts within a few years. As national university admission rates rose from approximately one-third of high-school graduates in 1998 to over half by the early 2000s (Ministry of Education of the People's Republic of China, 2005), the country's most selective universities, including Tsinghua University, Peking University, and Fudan University, deployed visible recruitment teams that competed publicly for each province's top scorers in the days immediately after results were released, and individual high schools used their top scorer to boost reputations and to attract applicants in subsequent admission cycles (China Daily, 2018; Global Times, 2024).

Official recognition and the regulatory response.. The publicity attached to provincial top scorers was pervasive enough to generate a sustained regulatory response from the central government. As part of a broader 2020 evaluation reform, the Ministry of Education, the central government department responsible for primary, secondary, and higher education in China, explicitly prohibited schools, local governments, and media from publicizing or hyping "top scorer" rankings (Ministry of Education, 2020); the prohibition has been reiterated in subsequent annual admissions notices (Ministry of Education, 2024), and the Ministry has formally acknowledged top-scorer hype as a longstanding policy target in its responses to the National People's Congress, China's national legislature (Ministry of Education, 2017). The fact that a regulatory ban was eventually deemed necessary, and that policy documents describe the underlying behavior as longstanding, is itself indirect evidence that the publicity surrounding top-scorer realizations was both pervasive and consequential throughout the 1990–2005 window covered by our analysis.

Implications for our empirical design.. Two features of our identification strategy are supported by these facts. First, the rarity of the outcome at the county level, where roughly one in a hundred counties produces a top scorer in any given year and the majority of counties produce none over a fifteen-year window, means that the realization of a top scorer is an unusual event for most students rather than a routine feature of the local environment. Second, the multiple mutually reinforcing publicity channels of central press, provincial and metropolitan press, school-level amplification, and post-1999 university recruitment competition ensure that the realization is widely observed within the home county, supporting the assumption that the event constitutes a salient information shock to nearby students throughout our sample period.

2. Arts–Science Tracking ("Wenli Fenke") in China, 1990–2005

From roughly 1990 to 2005, China's upper-secondary tracking system operated de facto as *wenli fenke* (arts–science streaming), even though national reform blueprints articulated a goal of weakening or abolishing the split; in practice, the organization of teaching and exam preparation remained tightly coupled to the Gaokao's track-specific subject bundles (Ministry of Education of the People's Republic of China, 2004). Consistent with this institutional reality, publicly reported Gaokao statistics in the early 2000s show that the science track was quantitatively dominant: among about 6.65 million national test-takers in 2004, arts candidates were about 2.19 million (32.9%) versus science candidates about 3.57 million (53.7%); and in 2005, among roughly 8.00 million national test-takers, arts candidates were about 2.78 million (35%) versus science candidates about 4.23 million (53%) (China Central Television (CCTV), 2004). Importantly, this was not only a "preference" phenomenon but also an opportunity-structure phenomenon: in Beijing in 2004, science registrants (53,465) substantially exceeded arts registrants (31,608), and the planned admissions quota likewise tilted toward science (45,978) relative to arts (17,797), reinforcing the perception that the science track offered broader or more secure pathways (China Central Television (CCTV), 2004). The contemporaneous social discourse explicitly labeled this pattern *zhongli qingwen* (valuing science over the humanities), often summarized by the widely circulated maxim *xuehao shulihua, zou bian tianxia dou bu pa* ("master math, physics, and chemistry and you can go anywhere"), indicating that science excellence carried disproportionate symbolic status and public esteem (Guangming Daily, 2001). Finally, elite "star" pipelines that generate highly visible adolescent role models—most notably Olympiad-based recognition and admissions advantages—were institutionally concentrated in mathematics, physics, chemistry, biology, and informatics, further amplifying the relative salience of science top scorers as role models compared with Humanities-track counterparts in this era (Ministry of Education of the People's Republic of China, 2001).

3. China's Administrative Geography

China's administrative system is multi-tiered. For our purposes, the relevant hierarchy is: *province* → *prefecture-level city* → *county-level units* (National Bureau of Statistics of China, 2015).

Provinces constitute the main subnational jurisdictions. The intermediate tier is typically the *prefecture-level city*, which, despite the name, is an administrative region that usually combines an urban core with surrounding suburban and rural areas.

Below the prefecture level, local jurisdictions are organized into *county-level units*. Two categories are most common in our data. The first is the *urban district* (*shixiaqu*), which governs the city proper and adjacent built-up areas. The second includes *counties* and *county-level cities*, which are generally located outside the central districts and often cover peri-urban towns and rural townships. Consequently, a single prefecture-level "city" frequently contains both urban districts and surrounding counties.

This distinction matters for interpretation because many public services and policy-relevant administrative functions are implemented at the county level (counties and districts). Accordingly, we treat county-level units as the local jurisdiction throughout the paper and cluster statistical inference at that level.

4. Conceptual Framework

This section presents a stylized belief-updating framework to rationalize the empirical patterns. The model is intentionally reduced-form, designed to map a local "top-scorer" shock into aggregate cohort-level changes in educational attainment.

Environment and decisions. Consider a student i in county j and cohort c. The student first chooses whether to stay in high school ($D = 0$) or drop out ($D = 1$). Conditional on staying, she chooses study effort $e \geq 0$ with convex cost $k(e) = e^2/(2\kappa)$ with $\kappa > 0$. Effort translates into a Gaokao score s according to the production function:

$$s_i = a_i + q_j \theta_j e_i + \varepsilon_i, \tag{D.1}$$

where a_i denotes individual ability, and ε_i is a mean-zero idiosyncratic shock. The term $q_j\theta_j$ captures the return to effort, composed of two components: observable local educational capacity q_j and an unobservable county-specific return parameter θ_j. Students know q_j but have imperfect information about θ_j.

Postsecondary outcomes are determined by score cutoffs τ_A (Associate/College) and τ_B (Bachelor), with $\tau_B > \tau_A$. The net present value of attaining each level is $V_B > V_A > V_N$, where N denotes no college. A student chooses effort to maximize expected utility:

$$U_S(a_i;\mu,q_j)=\max_{e\geq 0}\left\{-\frac{e^2}{2\kappa}+\beta\sum_{k\in\{N,A,B\}}\Pr(\text{Outcome}=k\mid e,\mu)\cdot V_k\right\}, \tag{D.2}$$

where expectations are taken over ε_i and beliefs $\mu\equiv\mathbb{E}[\theta_j\mid\mathcal{I}]$. Students stay in school if $U_S(a_i;\mu,q_j)$ exceeds the outside option $U_D(a_i)$.

Role-model shock as an informative signal. Let $Z_{jt}=1$ denote the emergence of a provincial top-scorer in county j. We conceptualize this rare event as a salient signal about the local return to effort. Specifically, the probability of producing a top-scorer, $\Pr(Z_{jt}=1\mid\theta_j)$, is strictly increasing in θ_j.

Given a prior $\theta_j\sim\mathcal{N}(\mu_0,\sigma_0^2)$, observing $Z_{jt}=1$ triggers a Bayesian update, shifting the posterior mean to $\mu_1>\mu_0$. This positive update is stronger for students in the same county due to the signal's local salience.

Implications for empirical patterns. The model yields three key testable implications that match our CSDID results:

(1) Reduction in dropout rates. A higher belief μ increases the marginal return to effort, raising optimal effort e^* and the expected value of schooling U_S. Consequently, students at the dropout margin (where $U_S\approx U_D$) are induced to stay in school, lowering the dropout rate.

(2) Disproportionate rise in Bachelor attainment (The Upgrade Effect). The shock shifts the score distribution to the right. The effect on "College or above" ($s\geq\tau_A$) versus "Bachelor or above" ($s\geq\tau_B$) depends on the density of marginal students.

$$\begin{aligned}\Delta\Pr(\text{BA+})&=\underbrace{\Pr(\text{N}\to\text{B})}_{\text{New Entrants}}+\underbrace{\Pr(\text{A}\to\text{B})}_{\text{Intensive Margin Upgrade}}\\ \Delta\Pr(\text{College+})&=\Pr(\text{N}\to\text{B})+\Pr(\text{N}\to\text{A}).\end{aligned} \tag{D.3}$$

If the belief shock primarily motivates high-performing students to cross the tier-one threshold τ_B (the "A to B" upgrade), we will observe a large increase in four-year university degree attainment with a potentially smaller net change in total college completion. This rationalizes why the BA effect is particularly robust and persistent in our data.

(3) Amplification by local capacity (Complementarity). The cross-derivative of the score function is positive: $\frac{\partial^2 s}{\partial e \partial \theta} = q_j$. This implies that the same belief shock ($\Delta\mu$) translates into larger score gains in regions with better educational infrastructure (q_j). High q_j regions have the resources to effectively convert the increased motivation into actual score improvements.

5. Additional Summary Statistics

Counties that first produce a provincial top scorer earlier tend to have stronger pre-period educational foundations than counties that experience their first top-scorer realization later or never. As shown in Table A1, earlier-treated counties exhibit higher average educational attainment and substantially higher rates of high-school and college completion among cohorts that reached Gaokao age prior to 1990, while later-treated and never-treated counties are more rural on average. These patterns are not surprising: the emergence of an extreme upper-tail outcome such as a provincial top scorer is more likely in places with a stronger pre-existing human-capital base.

Importantly, these baseline differences do not undermine our identification strategy. Our empirical design does not require treated and untreated counties to be identical in levels; instead, it relies on comparisons over time (across birth cohorts) within counties and on the use of not-yet-treated counties as contemporaneous controls under a parallel-trends assumption. We therefore treat Table A1 as descriptive evidence documenting the staggered and spatially dispersed nature of first top-scorer realizations and the heterogeneity in pre-period county conditions. The credibility of the design is assessed directly through event-study diagnostics, pre-trend tests, and robustness to alternative estimators and control-group definitions in subsequent sections.

Table A1: Baseline County Characteristics by Timing of First Provincial Top Scorer

	1990–1994	1995–1999	2000–2005	Never
Number of counties	76	106	95	2266
Share of counties	0.030	0.042	0.037	0.891
Years of schooling	10.70	10.09	9.82	8.79
High school share	0.43	0.34	0.31	0.20
Any higher educ share	0.22	0.16	0.14	0.07
4-yr higher educ share	0.11	0.08	0.06	0.03
Male share	0.51	0.51	0.51	0.51
Rural hukou share	0.45	0.57	0.59	0.71
Han ethnicity share	0.96	0.96	0.93	0.86

Notes: Counties are grouped by the calendar year in which they first produced a provincial top scorer during 1990–2005. Each county enters exactly one timing group based on its first appearance; subsequent top-scorer realizations within the same county are ignored. Baseline characteristics are constructed using cohorts that completed the Gaokao prior to 1990 (birth cohorts 1967–1972), whose educational outcomes were fully realized before the sample period. The table documents the staggered and spatially dispersed timing of first top-scorer realizations and is not intended to assess covariate balance.

6. Provincial Gaokao Top Scorers: County and School Attribution (1990—2005)

Table A2: Provincial Top Scorers by County and School, 1990–1995

Year	County	School
1990	Xicheng District	Beijing No. 4 High School
1990	Hexi District	Tianjin Experimental High School
1990	Chang'an District	Shijiazhuang No. 1 High School
1990	Yingze District	Jinshan High School
1990	Chaoyang District	Changchun Foreign Languages School

Continued on next page

Table A2 continued

Year	County	School
1990	Yangpu District	SUST Affiliated High School
1990	Qinhuai District	Nanjing No. 3 High School
1990	Zixi County	Zixi No. 1 High School
1990	Yuzhou City	Yuzhou No. 1 High School
1990	Zhecheng County	Zhecheng No. 1 High School
1990	Huaibin County	Huaibin High School
1990	Yicheng District	Zhumadian High School
1990	Jingzhou District	Jingzhou High School
1990	Huangzhou District	Huanggang High School
1990	Lengshuijiang City	Lengshuijiang No. 1 High School
1990	Xincheng District	Nanning No. 2 High School
1990	Lingchuan County	Guangxi Normal University Affiliated High School
1990	Quanzhou County	Quanzhou Senior High School
1990	Shapingba District	Chongqing No. 8 High School
1991	Hedong District	Tianjin No. 100 High School
1991	Qiaoxi District	Shijiazhuang No. 17 High School
1991	Mancheng District	Mancheng High School
1991	Fuping County	Fuping High School
1991	Yingze District	Chengcheng High School
1991	Xinghualing District	Taiyuan No. 12 High School
1991	Xiyang County	Xiyang High School
1991	Tonghua County	Tonghua County No. 7 High School
1991	Meihekou City	Meihekou No. 5 High School
1991	Suining County	Suining High School
1991	Donghu District	Jiangxi Normal University Affiliated High School
1991	Jimo District	Jimo No. 1 High School
1991	Bincheng District	Beizhen High School
1991	Suburban District	Luoyang No. 1 High School
1991	Dengzhou City	Dengzhou No. 1 High School
1991	Xiangcheng District	Xiangyang No. 4 High School
1991	Shaodong City	Shaodong No. 3 High School
1991	Xincheng District	Nanning No. 2 High School

Continued on next page

Table A2 continued

Year	County	School
1991	Xiufeng District	Guilin High School
1991	Xing'an County	Xing'an High School
1991	Beibei District	Southwest Normal University Affiliated High School
1991	Honghuagang District	Zunyi No. 4 High School
1991	Sinan County	Sinan High School
1991	Pengyang County	Pengyang No. 2 High School
1992	Chang'an District	Shijiazhuang No. 1 High School
1992	Xinhua District	Shijiazhuang No. 2 High School
1992	Yingze District	Taiyuan No. 5 High School
1992	Chaoyang District	Northeast Normal University Affiliated High School
1992	Chaoyang District	Changchun Foreign Languages School
1992	Hai'an City	Hai'an Senior High School
1992	Xinjian District	Wanli No. 1 High School
1992	Honggutan District	Nanchang No. 2 High School
1992	Shizhong District	Zaozhuang No. 3 High School
1992	Yanshi District	Yanshi High School
1992	Wuchang District	Wuhan University Affiliated High School
1992	Yueyang County	Yueyang County No. 1 High School
1992	Miluo City	Miluo No. 1 High School
1992	Xincheng District	Nanning No. 2 High School
1992	Haicheng District	Beihai High School
1992	Gangbei District	Guigang Senior High School
1992	Jinkouhe District	Yanfeng High School
1992	Bozhou District	Zunyi Nanbai High School
1992	Xingqing District	Yinchuan No. 9 High School
1993	Dongcheng District	Dongzhimen High School
1993	Xinhua District	Shijiazhuang No. 2 High School
1993	Yingze District	Taiyuan No. 5 High School
1993	Yaodu District	Linfen No. 3 High School
1993	Nanguan District	Changchun Experimental High School
1993	Meihekou City	Meihekou No. 11 High School
1993	Songjiang District	Songjiang No. 2 High School

Continued on next page

Table A2 continued

Year	County	School
1993	Donghu District	Jiangxi Normal University Affiliated High School
1993	Jizhou District	Ji'an No. 1 High School
1993	Xinzhou District	Shangrao No. 1 High School
1993	Suburban District	Luoyang No. 1 High School
1993	Wugang City	Wugang No. 1 High School
1993	Hongqi District	Xinxiang No. 1 High School
1993	Qingshan District	Wugang No. 3 High School
1993	Huangzhou District	Huanggang High School
1993	Suining County	Suining No. 1 High School
1993	Guiyang County	Guiyang No. 3 High School
1993	Nanshan District	Shenzhen Experimental High School
1993	Chengzhong District	Liuzhou Senior High School
1993	Xiufeng District	Guilin High School
1993	Pannan County	Pannan County High School
1993	Kaizhou District	Kaixian High School
1993	Guanshanhu District	Guiyang No. 1 High School
1993	Xincheng District	Shaanxi Construction No. 2 High School
1993	Weiyang District	Xi'an High School
1993	Shihezi City	Shihezi No. 5 High School
1993	Anshun City	Pingshui School
1994	Xicheng District	Beijing Normal University Second Affiliated High School
1994	Haidian District	Peking University Affiliated High School
1994	Nankai District	Tianjin Nankai High School
1994	Xinhua District	Shijiazhuang No. 2 High School
1994	Xiaodian District	Shanxi Provincial Experimental High School
1994	Shahekou District	Dalian No. 8 High School
1994	Ji'an City	Ji'an No. 1 High School
1994	Hunchun City	Hunchun No. 3 High School
1994	Pudong New Area	Shangnan High School
1994	Pudong New Area	Gaoqiao High School
1994	Hai'an City	Hai'an Senior High School
1994	Urban District	Yancheng High School

Continued on next page

Table A2 continued

Year	County	School
1994	Shizhong District	Bengbu No. 3 High School
1994	Tunxi District	Tunxi No. 1 High School
1994	Huoshan County	Huoshan High School
1994	Donghu District	Jiangxi Normal University Affiliated High School
1994	Xunyang District	Jiujiang No. 2 High School
1994	Zhongyuan District	Zhengzhou Foreign Languages School
1994	Jili District	Mengjin No. 1 High School
1994	Hongqi District	Xinxiang No. 1 High School
1994	Qingfeng County	Qingfeng No. 1 High School
1994	Shihe District	Xinyang High School
1994	Qingshan District	Wugang No. 3 High School
1994	Dangyang City	Dangyang No. 1 High School
1994	Liling City	Liling No. 1 High School
1994	Liling City	Liling No. 1 High School
1994	Yueyang County	Yueyang County No. 1 High School
1994	Xincheng District	Nanning No. 2 High School
1994	Lingchuan County	Guangxi Normal University Affiliated High School
1994	Lingshan County	Pubei High School
1994	Pannan County	Pannan County High School
1994	Jiangbei District	Chongqing No. 18 High School
1994	Shapingba District	Chongqing No. 1 High School
1994	Baiyun District	Baiyun No. 7 High School
1994	Hongta District	Yuxi No. 1 High School
1994	Weiyang District	Xi'an High School
1994	Xingqing District	Yinchuan No. 9 High School
1995	Dongcheng District	Huiwen High School
1995	Xicheng District	Beijing Normal University Affiliated Experimental High School
1995	Xicheng District	Beijing No. 8 High School
1995	Xicheng District	Beijing No. 4 High School
1995	Chang'an District	Shijiazhuang No. 1 High School
1995	Qiaoxi District	Shijiazhuang No. 17 High School

Continued on next page

Table A2 continued

Year	County	School
1995	Tianzhen County	Tianzhen No. 1 High School
1995	Xiangyuan County	Xiangyuan No. 1 High School
1995	Beining City	Beining High School
1995	Xuhui District	Shanghai No. 2 High School
1995	Songjiang District	Songjiang No. 2 High School
1995	Urban District	Nantong High School
1995	Urban District	Nantong No. 1 High School
1995	Tianjiaan District	Huainan No. 2 High School
1995	Tongguan District	Tongling No. 1 High School
1995	Honggutan District	Nanchang No. 2 High School
1995	Honggutan District	Nanchang No. 2 High School
1995	Zhanggong District	Ganzhou No. 1 High School
1995	Taihe County	Taihe High School
1995	Laiyang City	Laiyang No. 1 High School
1995	Dongping County	Dongping Senior High School
1995	Suburban District	Luoyang No. 1 High School
1995	Wen County	Wen County No. 1 High School
1995	Weidu District	Xuchang No. 2 High School
1995	Linying County	Linying No. 1 High School
1995	Guangshan County	Guangshan No. 2 High School
1995	Laohekou City	Laohekou No. 1 High School
1995	Jingzhou District	Jingzhou High School
1995	Yueyang County	Yueyang County No. 1 High School
1995	Lianyuan City	Lianyuan No. 1 High School
1995	Tianhe District	South China Normal University Affiliated High School
1995	Binyang County	Binyang High School
1995	Xiufeng District	Guilin High School
1995	Pidu District	Chengdu Foreign Languages School
1995	Youxian District	Mianyang Science City No. 1 High School
1995	Zhijin County	Zhijin No. 2 High School
1995	Beilin District	Xi'an Tieyi High School
1995	Xingqing District	Yinchuan No. 9 High School
1995	Kuitun City	Kuitun No. 1 High School

Table A3: Provincial Top Scorers by County and School, 1996–2000

Year	County	School
1996	Dongcheng District	Huiwen High School
1996	Haidian District	Renmin University of China Affiliated High School
1996	Chang'an District	Shijiazhuang No. 1 High School
1996	Xinji City	Xinji High School
1996	Mining District	Yangquan No. 15 High School
1996	Pingyao County	Pingyao High School
1996	Yanhu District	Kangjie High School
1996	Huanggu District	Shenyang Korean No. 1 High School
1996	Tiedong District	Anshan No. 1 High School
1996	Jiutai District	Jiutai No. 1 High School
1996	Gongzhuling City	Gongzhuling No. 1 High School
1996	Yangpu District	Fudan University Affiliated High School
1996	Baoshan District	Xingzhi High School
1996	Feng County	Feng County High School
1996	Urban District	Nantong No. 1 High School
1996	Wuyi County	Wuyi No. 1 High School
1996	Yiwu City	Yiwu High School
1996	Suburban District	Hefei No. 1 High School
1996	Jiujiang District	Wuhu No. 1 High School
1996	She County	She County High School
1996	Zhushan District	Jingdezhen No. 1 High School
1996	Jizhou District	Bailuzhou High School
1996	Pingdu City	Pingdu No. 9 High School
1996	Changyi City	Changyi Wenshan High School
1996	Rongcheng City	Rongcheng No. 2 High School
1996	Wulian County	Wulian No. 1 High School
1996	Yishui County	Yishui No. 2 High School
1996	Xingyang City	Xinmi Experimental High School
1996	Shunhe Hui District	Henan University Affiliated High School
1996	Jun County	Jun County No. 1 High School
1996	Pingqiao District	Xinyang No. 4 High School
1996	Jiang'an District	Wuhan No. 6 High School
1996	Zhicheng City	Yidu No. 1 High School

Continued on next page

Table A.3 continued

Year	County	School
1996	Beiqu District	Changsha No. 1 High School
1996	Xinhua County	Xinhua No. 1 High School
1996	Suburban District	Shantou No. 1 High School
1996	Xincheng District	Nanning No. 2 High School
1996	Chengzhong District	Liuzhou Senior High School
1996	Tongliang District	Tongliang High School
1996	Pidu District	Chengdu Foreign Languages School
1996	Yunyan District	Guiyang No. 6 High School
1996	Wuhua District	Yunnan Normal University Affiliated High School
1996	Kaiyuan City	Kaiyuan No. 1 High School
1996	Weiyang District	Xi'an Senior High School
1996	Qindu District	Xianyang Weicheng High School
1996	Xigu District	Lanzhou Refinery No. 1 High School
1996	Wulan County	Qinghai Oilfield No. 1 High School
1996	Pingluo County	Hengli High School
1997	Xicheng District	Beijing Normal University Second Affiliated High School
1997	Shunyi District	Niulanshan No. 1 High School
1997	Heping District	Tianjin No. 55 High School
1997	Binhai New Area	Dagang Oilfield Experimental High School
1997	Haigang District	Qinhuangdao No. 1 High School
1997	Qiaoxi District	Zhangjiakou No. 1 High School
1997	Linyi County	Linyi High School
1997	Huanggu District	Shenyang Korean No. 1 High School
1997	Xinchengzi District	Shenyang No. 83 High School
1997	Chaoyang District	Changchun Foreign Languages School
1997	Xuhui District	Shanghai High School
1997	Tianning District	Changzhou Senior High School
1997	Urban District	Nantong High School
1997	Tongxiang City	Tongxiang Senior High School
1997	Lanxi City	Lanxi No. 1 High School
1997	Huashan District	Ma'anshan No. 2 High School
1997	Tongguan District	Tongling No. 1 High School

Continued on next page

Table A.3 continued

Year	County	School
1997	Jin'an District	Lu'an No. 1 High School
1997	Donghu District	Nanchang No. 10 High School
1997	Xunyang District	Jiujiang No. 1 High School
1997	Laiwu District	Laiwu No. 1 High School
1997	Laiyang City	Laiyang No. 1 High School
1997	Wucheng County	Wucheng No. 2 High School
1997	Xinzheng City	Xinzheng No. 1 High School
1997	Lushan County	Lushan No. 1 Senior High School
1997	Urban Area	Puyang Oilfield No. 1 High School
1997	Xiangcheng District	Xiangyang No. 4 High School
1997	Shashi District	Shashi High School
1997	Beihu District	Chenzhou No. 1 High School
1997	Lengshuijiang City	Lengshuijiang No. 1 High School
1997	Liwan District	Guangya High School
1997	Liujiang District	Liuzhou Prefecture Ethnic High School
1997	Shapingba District	Chongqing No. 1 High School
1997	Jiulongpo District	Chongqing Foreign Languages School
1997	Jingyang District	Deyang No. 3 High School
1997	Xixiu District	Anshun No. 2 High School
1997	Chenggong District	Kunming No. 3 High School
1997	Weibin District	Baoji High School
1997	Chengguan District	Lanzhou No. 1 High School
1997	Xingqing District	Yinchuan No. 9 High School
1997	Yizhou District	Hami No. 2 High School
1998	Xicheng District	Beijing No. 4 High School
1998	Chaoyang District	Beijing No. 80 High School
1998	Heping District	Tianjin No. 20 High School
1998	Heping District	Yaohua High School
1998	Xinhua District	Shijiazhuang No. 2 High School
1998	Congtai District	Handan No. 1 High School
1998	Taocheng District	Hengshui High School
1998	Xiyang County	Xiyang High School
1998	Huanggu District	Liaoning Provincial Experimental High School
1998	Shahekou District	Liaoning Normal University Affiliated High School

Continued on next page

Table A.3 continued

Year	County	School
1998	Chaoyang District	Jilin Provincial Experimental High School
1998	Huinan County	Huinan No. 6 High School
1998	Nangang District	Harbin No. 3 High School
1998	Jintan District	Jintan Hua Luogeng High School
1998	Rudong County	Rudong Senior High School
1998	Cixi City	Cixi High School
1998	Tongxiang City	Tongxiang Senior High School
1998	Dangtu County	Dangtu No. 1 High School
1998	Dongzhi County	Dongzhi No. 3 High School
1998	Qingyunpu District	Hongdu High School
1998	Boyang County	Poyang No. 1 High School
1998	Huangdao District	Jiaonan No. 2 High School
1998	Huangdao District	Jiaonan No. 2 High School
1998	Pingdu City	Pingdu No. 1 High School
1998	Pingdu City	Pingdu No. 9 High School
1998	Zhangdian District	Zibo Experimental High School
1998	Boshan District	Zibo No. 1 High School
1998	Yiyuan County	Yiyuan No. 1 High School
1998	Laiyang City	Laiyang No. 1 High School
1998	Gaomi City	Gaomi No. 1 High School
1998	Hongqi District	Xinxiang No. 1 High School
1998	Changge City	Changge No. 1 Senior High School
1998	Shihe District	Xinyang High School
1998	Huangchuan County	Huangchuan High School
1998	Suiping County	Suiping No. 1 Senior High School
1998	Zhongxiang City	Huji High School
1998	Huangzhou District	Huanggang High School
1998	Heshan District	Yiyang Zhenyan High School
1998	Taojiang County	Taojiang No. 1 High School
1998	Suburban District	Shantou No. 1 High School
1998	Chengzhong District	Liuzhou Senior High School
1998	Haicheng District	Beihai High School
1998	Wuhou District	Chengdu Yulin High School
1998	Yanjiang District	Ziyang High School

Continued on next page

Table A.3 continued

Year	County	School
1998	Shuicheng Special District	Shuicheng Mining No. 1 High School
1998	Kaili City	Kaili No. 1 High School
1998	Xiaguan City	Xiaguan No. 1 High School
1998	Xincheng District	Xi'an Aichi High School
1998	Lianhu District	Xi'an No. 1 High School
1998	Qindu District	Xianyang Daobei Railway High School
1998	Xingqing District	Yinchuan No. 2 High School
1998	Korla City	Bayannur No. 2 High School
1999	Dongcheng District	Huiwen High School
1999	Haidian District	Beijing No. 101 High School
1999	Heping District	Yaohua High School
1999	Heping District	Tianjin No. 1 High School
1999	Chang'an District	Shijiazhuang No. 1 High School
1999	Gucheng County	Gucheng Zhengkou High School
1999	Yanhu District	Kangjie High School
1999	Yanhu District	Kangjie High School
1999	Kangping County	Kangping High School
1999	Ganjingzi District	Dalian No. 23 High School
1999	Tiedong District	Anshan No. 1 High School
1999	Chaoyang District	Changchun Foreign Languages School
1999	Hunchun City	Hunchun No. 2 High School
1999	Changning District	Shanghai Yan'an High School
1999	Baoshan District	Luodian High School
1999	Jiangyin City	Nanjing Senior High School
1999	Tianning District	Changzhou Senior High School
1999	Rudong County	Rudong Senior High School
1999	Jiangdu District	Jiangdu High School
1999	Yiwu City	Yiwu High School
1999	Dongyang City	Dongyang High School
1999	Jingning She Autonomous County	Jingning High School
1999	Dangtu County	Dangtu No. 1 High School
1999	Zongyang County	Zongyang Fushan High School
1999	Jin'an District	Lu'an No. 2 High School

Continued on next page

Table A.3 continued

Year	County	School
1999	Fenyi County	Fenyi High School
1999	Nancheng County	Nancheng No. 1 High School
1999	Nanguan District	Kaifeng High School
1999	Shilong District	Pingdingshan Experimental High School
1999	Neihuang County	Neihuang No. 1 High School
1999	Hongqi District	Xinxiang No. 1 High School
1999	Neixiang County	Neixiang High School
1999	Zhecheng County	Zhecheng No. 1 Senior High School
1999	Hongshan District	Central China Normal University No. 1 Affiliated High School
1999	Danjiangkou City	Danjiangkou No. 1 High School
1999	Zhicheng City	Yidu No. 1 High School
1999	Suburban District	Zhuzhou No. 2 High School
1999	Beihu District	Chenzhou No. 1 High School
1999	Chengzhong District	Liuzhou Senior High School
1999	Yuanjiang District	Wuzhou Senior High School
1999	Guiping City	Xunzhou Senior High School
1999	Qingyang District	Shude High School
1999	Wenjiang District	Chengdu Experimental Foreign Languages School
1999	Huaxi District	Guiyang Tsinghua High School
1999	Qian County	Qian County No. 1 High School
1999	Qian County	Qian County No. 1 High School
1999	Baiyin District	Baiyin Company No. 2 High School
1999	Pingchuan District	Jingyuan Mining Bureau No. 1 High School
1999	Chengbei District	Xining No. 5 High School
1999	Yuanzhou District	Guyuan No. 1 High School
2000	Chaoyang District	Beijing No. 80 High School
2000	Shunyi District	Shunyi No. 1 High School
2000	Heping District	Tianjin No. 20 High School
2000	Heping District	Yaohua High School
2000	Nankai District	Tianjin Nankai High School
2000	Chang'an District	Shijiazhuang No. 1 High School
2000	Lianchi District	Baoding No. 17 High School

Continued on next page

Table A.3 continued

Year	County	School
2000	Yunhe District	Cangzhou No. 1 High School
2000	Taigu District	Taigu High School
2000	Pingyao County	Pingyao High School
2000	Shahekou District	Dalian Yuming High School
2000	Tiexi District	Anshan Korean High School
2000	Chaoyang District	Jilin Provincial Experimental High School
2000	Chaoyang District	Northeast Normal University Affiliated High School
2000	Jing'an District	Shixi High School
2000	Hongkou District	East China Normal University No. 1 Affiliated High School
2000	Gusu District	Suzhou High School
2000	Rugao City	Rugao No. 1 High School
2000	Urban District	Yancheng High School
2000	Jiashan County	Jiashan High School
2000	Tongxiang City	Tongxiang Senior High School
2000	Zhuji City	Zhuji High School
2000	Lujiang County	Lujiang High School
2000	Huashan District	Ma'anshan No. 2 High School
2000	Jinzhai County	Jinzhai No. 1 High School
2000	Zhushan District	Jingdezhen No. 2 High School
2000	Yuanzhou District	Yichun High School
2000	Jimo District	Jimo No. 4 High School
2000	Yicheng District	Zaozhuang No. 1 High School
2000	Longkou City	Longkou No. 1 High School
2000	Dongping County	Dongping Senior High School
2000	Donggang District	Rizhao No. 1 High School
2000	Mudan District	Heze No. 1 High School
2000	Jianxi District	Luoyang University of Technology Affiliated High School
2000	Hongqi District	Xinxiang No. 1 High School
2000	Huixian City	Huixian No. 1 High School
2000	Puyang County	Puyang County No. 3 High School
2000	Dengzhou City	Dengzhou No. 1 Senior High School

Continued on next page

Table A.3 continued

Year	County	School
2000	Xiangcheng City	Xiangcheng No. 1 Senior High School
2000	Jianghan District	Wuhan Foreign Languages School
2000	Tianmen City	Tianmen High School
2000	Nan District	Changjun High School
2000	Yuhu District	Xiangtan University Affiliated School
2000	Lixian County	Lixian No. 1 High School
2000	Cili County	Cili No. 1 High School
2000	Chengzhong District	Liuzhou Senior High School
2000	Lingchuan County	Guangxi Normal University Affiliated High School
2000	Xing'an County	Xing'an High School
2000	Yuanjiang District	Wuzhou Senior High School
2000	Wuhou District	Chengdu No. 7 High School
2000	Wenjiang District	Chengdu Experimental Foreign Languages School
2000	Yunyan District	Guiyang No. 6 High School
2000	Guanshanhu District	Guiyang No. 1 High School
2000	Huichuan District	Zunyi No. 1 High School
2000	Hongta District	Yuxi No. 1 High School
2000	Chuxiong City	Chuxiong No. 1 High School
2000	Weiyang District	Xi'an High School
2000	Qindu District	Xianyang Daobei Railway High School
2000	Chengguan District	Lanzhou Railway No. 1 High School
2000	Huining County	Huining No. 1 High School
2000	Chengxi District	Huangchuan High School
2000	Wulan County	Qinghai Oilfield No. 1 High School
2000	Xingqing District	Yinchuan No. 2 High School
2000	Xingqing District	Tanglai Hui High School
2000	Tianshan District	Xinjiang Experimental High School
2000	Toutunhe District	Urumqi No. 1 High School
2000	Karamay District	Karamay No. 6 High School

Table A4: Provincial Top Scorers by County and School, 2001–2005

Year	County	School
2001	Xicheng District	Beijing No. 15 High School
2001	Chaoyang District	Beijing No. 80 High School
2001	Hexi District	Tianjin Foreign Languages School
2001	Nankai District	Tianjin Nankai High School
2001	Zunhua City	Zunhua No. 1 High School
2001	Haigang District	Qinhuangdao Private Bohai High School
2001	Xinghualing District	Taiyuan Foreign Languages School
2001	Yanhu District	Kangjie High School
2001	Shenhe District	Shenyang No. 27 High School
2001	Huanggu District	Shenyang Korean No. 1 High School
2001	Shahekou District	Dalian No. 8 High School
2001	Tiedong District	Anshan No. 1 High School
2001	Chaoyang District	Northeast Normal University Affiliated High School
2001	Chaoyang District	Northeast Normal University Affiliated High School
2001	Yanji City	Yanbian No. 1 High School
2001	Jing'an District	Shixi High School
2001	Pudong New Area	East China Normal University Second Affiliated High School
2001	Changshu City	Changshu High School
2001	Urban District	Nantong High School
2001	Jiangdu District	Jiangdu High School
2001	Urban District	Wenzhou No. 15 High School
2001	Shengzhou City	Shengzhou No. 2 High School
2001	Shengzhou City	Shengzhou No. 1 High School
2001	Feidong County	Feidong No. 1 High School
2001	Xiangshan District	Huaibei No. 1 High School
2001	Donghu District	Nanchang No. 10 High School
2001	Xunyang District	Jiujiang No. 1 High School
2001	Shanggao County	Shanggao No. 2 High School
2001	Wendeng District	Wendeng No. 2 High School
2001	Decheng District	Dezhou No. 2 High School

Continued on next page

Table A.4 continued

Year	County	School
2001	Urban Area	Puyang Oilfield No. 1 High School
2001	Wancheng District	Nanyang No. 2 High School
2001	Shihe District	Xinyang High School
2001	Xiangcheng District	Xiangyang No. 4 High School
2001	Macheng City	Macheng No. 1 High School
2001	Zhuhui District	Hengyang No. 8 High School
2001	Hengdong County	Hengdong No. 1 High School
2001	Xincheng District	Nanning No. 2 High School
2001	Yuanjiang District	Wuzhou Senior High School
2001	Yuanjiang District	Wuzhou Senior High School
2001	Qingyang District	Shishi High School
2001	Shizhong District	Neijiang No. 6 High School
2001	Guanshanhu District	Guiyang No. 1 High School
2001	Guanshanhu District	Guiyang No. 1 High School
2001	Kaili City	Kaili No. 1 High School
2001	Wuhua District	Yunnan Normal University Affiliated High School
2001	Gejiu City	Gejiu No. 1 High School
2001	Beilin District	Northwestern Polytechnical University Affiliated High School
2001	Beilin District	Northwestern Polytechnical University Affiliated High School
2001	Weiyang District	Xi'an High School
2001	Yanta District	Xidian University Affiliated High School
2001	Chengguan District	Lanzhou No. 1 High School
2001	Jiayuguan City	Jiugang No. 3 High School
2001	Chengxi District	Huangchuan High School
2001	Chengxi District	Huangchuan High School
2001	Xingqing District	Tanglai Hui High School
2001	Xingqing District	Yinchuan No. 2 High School
2001	Tianshan District	Xinjiang Experimental High School
2001	Karamay District	Karamay No. 13 High School
2002	Xicheng District	Beijing No. 4 High School
2002	Chaoyang District	Beijing No. 80 High School

Continued on next page

Table A.4 continued

Year	County	School
2002	Heping District	Tianjin No. 1 High School
2002	Nankai District	Tianjin Nankai High School
2002	Lianchi District	Baoding No. 17 High School
2002	Laiyuan County	Laiyuan No. 1 High School
2002	Shenzhou City	Shenzhou High School
2002	Xiaodian District	Shanxi Provincial Experimental High School
2002	Yangcheng County	Yangcheng No. 1 High School
2002	Huanren County	Huanren Korean High School
2002	Xinqiu District	Fuxin High School
2002	Chaoyang District	Changchun Foreign Languages School
2002	Lishu County	Lishu No. 1 High School
2002	Putuo District	Caoyang No. 2 High School
2002	Yangpu District	Kongjiang High School
2002	Tianning District	Changzhou Senior High School
2002	Zhenhai District	Ningbo Zhenhai High School
2002	Xiangshan County	Xiangshan High School
2002	Yuyao City	Yuyao High School
2002	Jinghu District	Anhui Normal University Affiliated High School
2002	Jinghu District	Anhui Normal University Affiliated High School
2002	Zhushan District	Jingdezhen No. 1 High School
2002	Xinzhou District	Shangrao No. 1 High School
2002	Qufu City	Qufu Normal University Affiliated High School
2002	Feicheng City	Feicheng Taixi High School
2002	Wulian County	Wulian No. 1 High School
2002	Nanguan District	Kaifeng High School
2002	Sheqi County	Sheqi No. 1 Senior High School
2002	Jianghan District	Wuhan Foreign Languages School
2002	Qianjiang City	Qianjiang High School
2002	Xiangxiang City	Xiangxiang No. 1 High School
2002	Huarong County	Huarong No. 1 High School
2002	Lianyuan City	Lianyuan No. 1 High School
2002	Xincheng District	Nanning No. 2 High School
2002	Xincheng District	Nanning No. 2 High School
2002	Liunan District	Liutie No. 1 High School

Continued on next page

Table A.4 continued

Year	County	School
2002	Shapingba District	Chongqing No. 1 High School
2002	Qingyang District	Shude High School
2002	Ziliujing District	Zigong Shuguang High School
2002	Pingshan County	Pingshan High School
2002	Nanming District	Guiyang Defen Tutoring School
2002	Kaili City	Kaili No. 1 High School
2002	Wuhua District	Yunnan Normal University Affiliated High School
2002	Chuxiong City	Chuxiong No. 1 High School
2002	Yanta District	Xi'an Jiaotong University Affiliated High School
2002	Yanta District	Shaanxi Normal University Affiliated High School
2002	Weibin District	Baoji High School
2002	Jiayuguan City	Jiugang No. 3 High School
2002	Baiyin District	Baiyin No. 10 High School
2002	Xingqing District	Yinchuan No. 2 High School
2002	Xingqing District	Tanglai Hui High School
2002	Xingqing District	Yinchuan No. 9 High School
2002	Dawukou District	Shizuishan Mining Bureau No. 1 High School
2002	Toutunhe District	Urumqi No. 1 High School
2002	Korla City	Nong'er Shi Huashan High School
2002	Kashgar City	Nongsan Shi High School
2003	Xicheng District	Beijing Normal University Second Affiliated High School
2003	Haidian District	Peking University Affiliated High School
2003	Binhai New Area	Dagang Oilfield Experimental High School
2003	Binhai New Area	Tanggu No. 1 High School
2003	Chang'an District	Shijiazhuang No. 1 High School
2003	Zunhua City	Zunhua No. 1 High School
2003	Yingze District	Chengcheng High School
2003	Yu County	Yu County No. 1 High School
2003	Pingshan District	Benxi High School
2003	Nong'an County	Nong'an Experimental High School

Continued on next page

Table A.4 continued

Year	County	School
2003	Fengman District	Jilin No. 1 High School
2003	Fengman District	Jilin No. 1 High School
2003	Huangpu District	Shanghai Datong High School
2003	Jing'an District	Shixi High School
2003	Gusu District	Suzhou High School
2003	Urban District	Nantong No. 1 High School
2003	Xihu District	Hangzhou Foreign Languages School
2003	Pujiang County	Pujiang High School
2003	Daguan District	Anqing No. 1 High School
2003	Jin'an District	Lu'an No. 1 High School
2003	Ruichang City	Ruichang No. 1 High School
2003	Wannian County	Wannian High School
2003	Gaomi City	Gaomi No. 1 High School
2003	Shijiao District	Jining Yucai High School
2003	Luozhuang District	Linyi Shuangyueyuan School
2003	Erqi District	Zhengzhou No. 101 High School
2003	Qi County	Qi County High School
2003	Shashi District	Shashi High School
2003	Xiantao City	Mianzhou High School
2003	Xiangtan County	Xiangtan County No. 1 High School
2003	Taoyuan County	Taoyuan No. 1 High School
2003	Lianyuan City	Lianyuan Xingzhi High School
2003	Binyang County	Binyang High School
2003	Yufeng District	Liuzhou No. 1 High School
2003	Yufeng District	Liuzhou No. 1 High School
2003	Wenjiang District	Chengdu Experimental Foreign Languages School
2003	Fushun County	Fushun No. 2 High School
2003	Yunyan District	Guiyang No. 6 High School
2003	Guanshanhu District	Guiyang No. 1 High School
2003	Chenggong District	Kunming No. 3 High School
2003	Qilin District	Qujing No. 1 High School
2003	Beilin District	Xi'an No. 86 High School

Continued on next page

Table A.4 continued

Year	County	School
2003	Yanta District	Xi'an Jiaotong University Affiliated High School
2003	Chengguan District	Lanzhou No. 1 High School
2003	Yongdeng County	Gaolan No. 1 High School
2003	Chengxi District	Huangchuan High School
2003	Chengxi District	Huangchuan High School
2003	Xingqing District	Yinchuan No. 1 High School
2003	Xingqing District	Yinchuan No. 9 High School
2003	Haiyuan County	Haiyuan High School
2003	Toutunhe District	Urumqi No. 1 High School
2003	Yizhou District	Hami No. 2 High School
2004	Xicheng District	Beijing Normal University Affiliated High School
2004	Haidian District	Renmin University of China Affiliated High School
2004	Heping District	Yaohua High School
2004	Chang'an District	Shijiazhuang No. 1 High School
2004	Xinhua District	Shijiazhuang No. 2 High School
2004	Zhengding County	Zhengding High School
2004	Yingze District	Taiyuan No. 5 High School
2004	Xinghualing District	Taiyuan No. 18 High School
2004	Shahekou District	Dalian Yuming High School
2004	Zhuanghe City	Zhuanghe High School
2004	Chaoyang District	Northeast Normal University Affiliated High School
2004	Chaoyang District	Northeast Normal University Affiliated High School
2004	Putuo District	Yichuan High School
2004	Yangpu District	Kongjiang High School
2004	Jiangyan District	Jiangyan High School
2004	Haining City	Haining Hongda High School
2004	Keqiao District	Shaoxing Luxun High School
2004	Suburban District	Hefei No. 1 High School
2004	Suburban District	Hefei No. 1 High School

Continued on next page

Table A.4 continued

Year	County	School
2004	Nanling County	Nanling High School
2004	Shizhong District	Bengbu No. 2 High School
2004	Yuehu District	Yingtan No. 1 High School
2004	Ningdu County	Ningdu High School
2004	Zhifu District	Yantai No. 2 High School
2004	Laizhou City	Laizhou No. 1 High School
2004	Shihe District	Xinyang High School
2004	Xiangcheng City	Xiangcheng No. 1 Senior High School
2004	Jianghan District	Wuhan Foreign Languages School
2004	Hongshan District	Central China Normal University No. 1 Affiliated High School
2004	Shashi District	Shashi High School
2004	Nan District	Changjun High School
2004	Sangzhi County	Sangzhi No. 1 High School
2004	Xincheng District	Nanning No. 2 High School
2004	Liujiang District	Liuzhou Prefecture Ethnic High School
2004	Gangbei District	Guigang Senior High School
2004	Youjiang District	Baise Senior High School
2004	Wenjiang District	Chengdu Experimental Foreign Languages School
2004	Pidu District	Chengdu Foreign Languages School
2004	Shifang City	Shifang High School
2004	Guanshanhu District	Guiyang No. 1 High School
2004	Xixiu District	Anshun No. 2 High School
2004	Hongta District	Yuxi No. 1 High School
2004	Shuifu City	Yuntianhua High School
2004	Beilin District	Northwestern Polytechnical University Affiliated High School
2004	Zhen'an County	Zhen'an High School
2004	Ganzhou District	Zhangye High School
2004	Lintao County	Lintao High School
2004	Huzhu Tu Autonomous County	Huzhu No. 1 High School
2004	Xingqing District	Yinchuan No. 1 High School

Continued on next page

Table A.4 continued

Year	County	School
2004	Xingqing District	Yinchuan No. 2 High School
2004	Korla City	Nong'er Shi Huashan High School
2004	Kuitun City	Kuitun No. 1 High School
2005	Dongcheng District	Dongzhimen High School
2005	Xicheng District	Beijing Normal University Second Affiliated High School
2005	Haidian District	Renmin University of China Affiliated High School
2005	Hexi District	Tianjin Nankai Xiangyu School
2005	Nankai District	Tianjin Nankai High School
2005	Xinhua District	Shijiazhuang No. 2 High School
2005	Congtai District	Handan No. 1 High School
2005	Yanhu District	Kangjie High School
2005	Hejin City	Hejin Yongmin High School
2005	Huanggu District	Liaoning Provincial Experimental High School
2005	Xinglongtai District	Panjin Senior High School
2005	Shuangta District	Chaoyang No. 2 Senior High School
2005	Chaoyang District	Northeast Normal University Affiliated High School
2005	Chaoyang District	Northeast Normal University Affiliated High School
2005	Jing'an District	Pengpu High School
2005	Yangpu District	Yangpu High School
2005	Songjiang District	Songjiang No. 2 High School
2005	Gusu District	Suzhou High School
2005	Ruian City	Ruian High School
2005	Kecheng District	Quzhou No. 2 High School
2005	Tianjiaan District	Huainan No. 1 High School
2005	Lingbi County	Lingbi No. 1 High School
2005	Guichi District	Guichi No. 1 High School
2005	Yushui District	Xinyu No. 1 High School
2005	Jizhou District	Ji'an No. 1 High School
2005	Lichuan County	Lichuan No. 1 High School
2005	Xuecheng District	Zaozhuang No. 8 High School

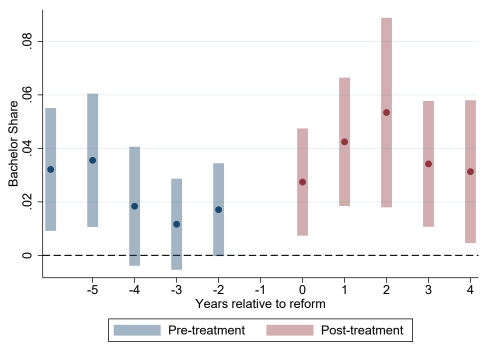
Continued on next page

Table A.4 continued

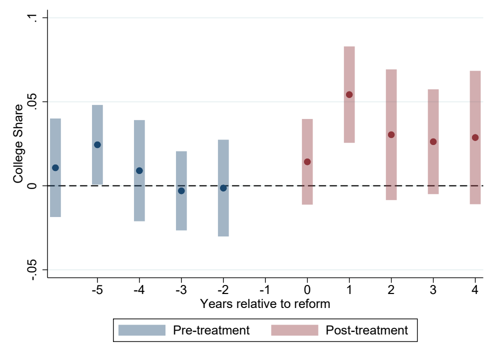
Year	County	School
2005	Zhifu District	Yantai No. 2 High School
2005	Laizhou City	Laizhou No. 1 High School
2005	Laizhou City	Laizhou No. 1 High School
2005	Jinshui District	Henan Provincial Experimental High School
2005	Yicheng District	Zhumadian High School
2005	Jianghan District	Wuhan Foreign Languages School
2005	Xiangcheng District	Xiangyang No. 4 High School
2005	Zhuhui District	Hengyang Experimental High School
2005	Changning City	Changning No. 1 High School
2005	Jishou City	Xiangxi Ethnic High School
2005	Chengzhong District	Liuzhou Senior High School
2005	Chengzhong District	Liuzhou Senior High School
2005	Guiping City	Xunzhou Senior High School
2005	Youjiang District	Qifu Senior High School
2005	Shizhong District	Neijiang No. 6 High School
2005	Shizhong District	Neijiang No. 6 High School
2005	Tongzi County	Tongzi No. 1 High School
2005	Xingyi City	Xingyi No. 1 High School
2005	Wuhua District	Kunming No. 1 High School
2005	Qilin District	Qujing No. 1 High School
2005	Beilin District	Northwestern Polytechnical University Affiliated High School
2005	Yanta District	Xi'an Bodi School
2005	Chengguan District	Lanzhou No. 2 High School
2005	Qingcheng County	Longdong High School
2005	Chengxi District	Huangchuan High School
2005	Chengxi District	Huangchuan High School
2005	Xingqing District	Yinchuan No. 2 High School
2005	Xingqing District	Tanglai Hui High School
2005	Tianshan District	Xinjiang Experimental High School
2005	Yizhou District	Hami No. 2 High School
2005	Aletai City	Aletai No. 2 High School

7. Robustness Checks: Additional Figures

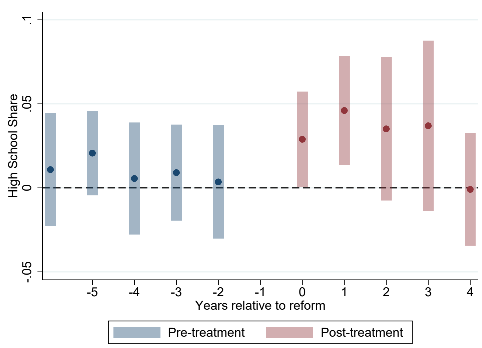
Figure A1: Event-Study Estimates without Multiple Top-Scorer Counties



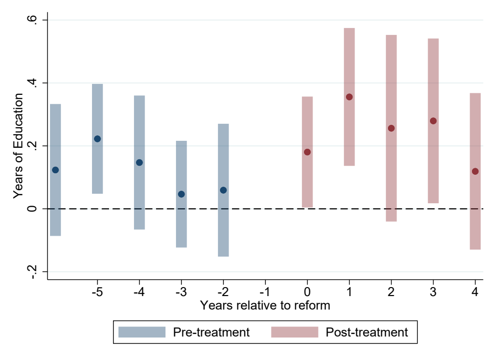
(a) Bachelor's degree



(b) Any higher education



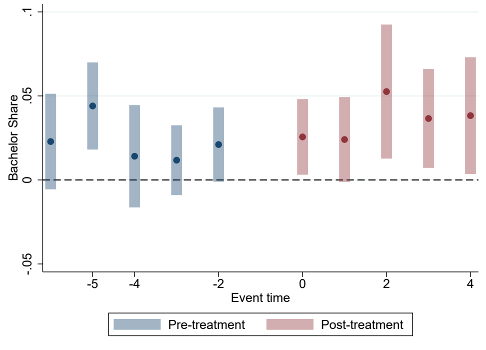
(c) High school completion



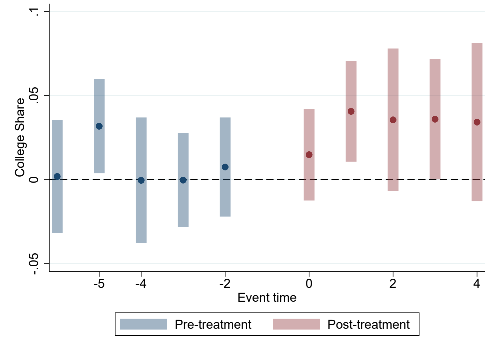
(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with county-clustered standard errors. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

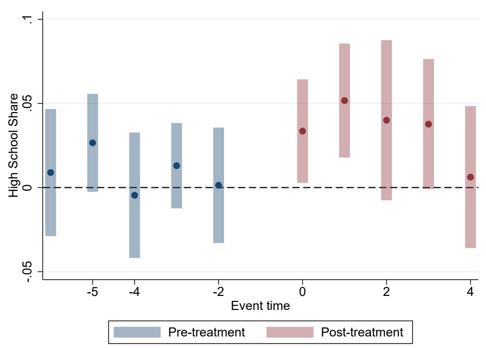
Figure A2: Event-Study Estimates with Smaller Top-Scorer Sample



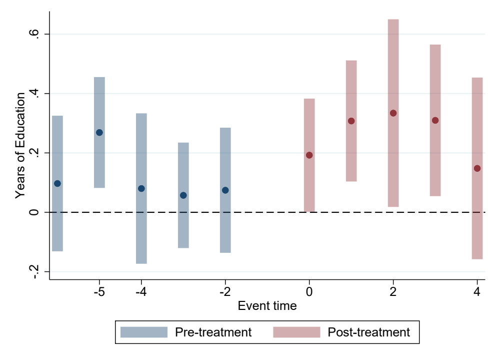
(a) Bachelor's degree



(b) Any higher education



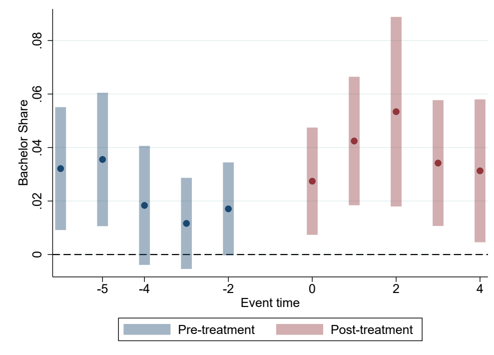
(c) High school completion



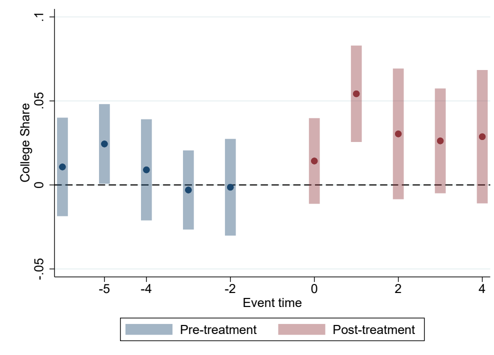
(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with county-clustered standard errors. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

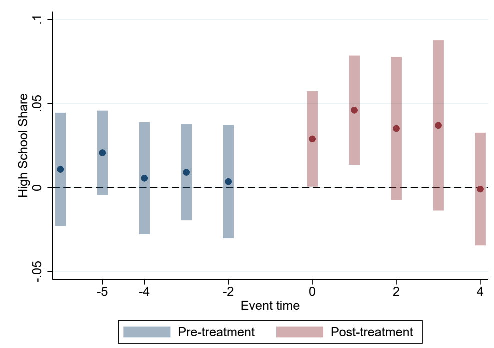
Figure A3: Event-Study Estimates with Never-Treated Counties



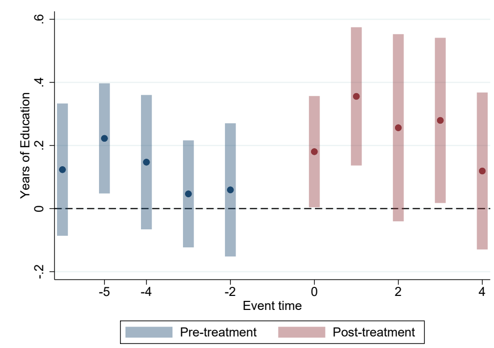
(a) Bachelor's degree



(b) Any higher education



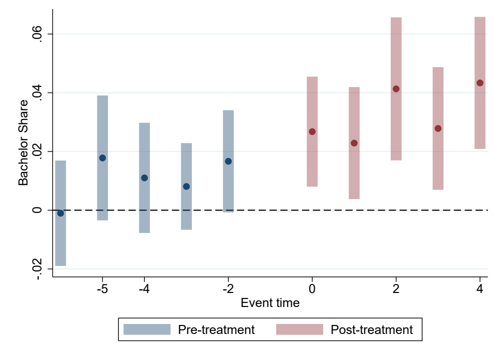
(c) High school completion



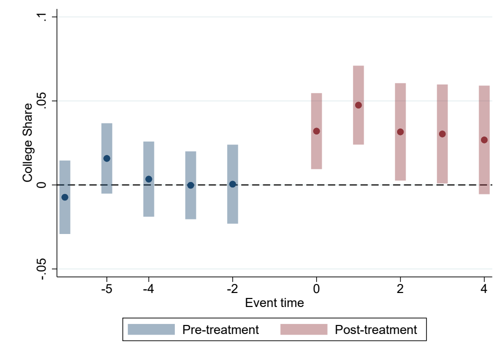
(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with county-clustered standard errors. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

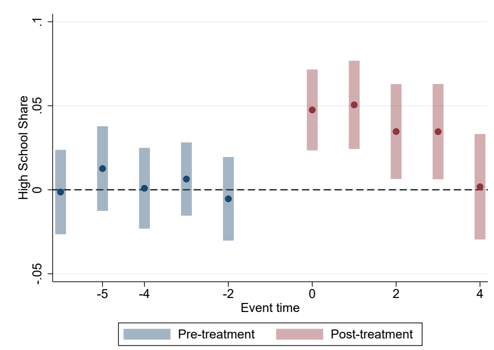
Figure A4: Event-Study Estimates (Migration Robustness: Restricting to Non-movers)



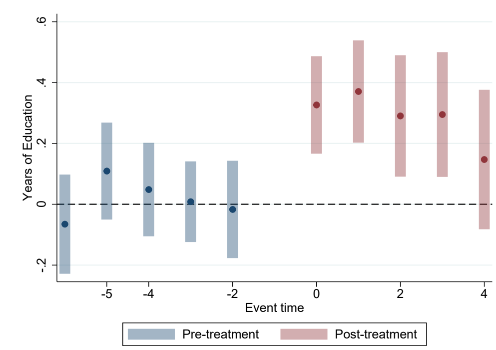
(a) Bachelor's degree



(b) Any higher education



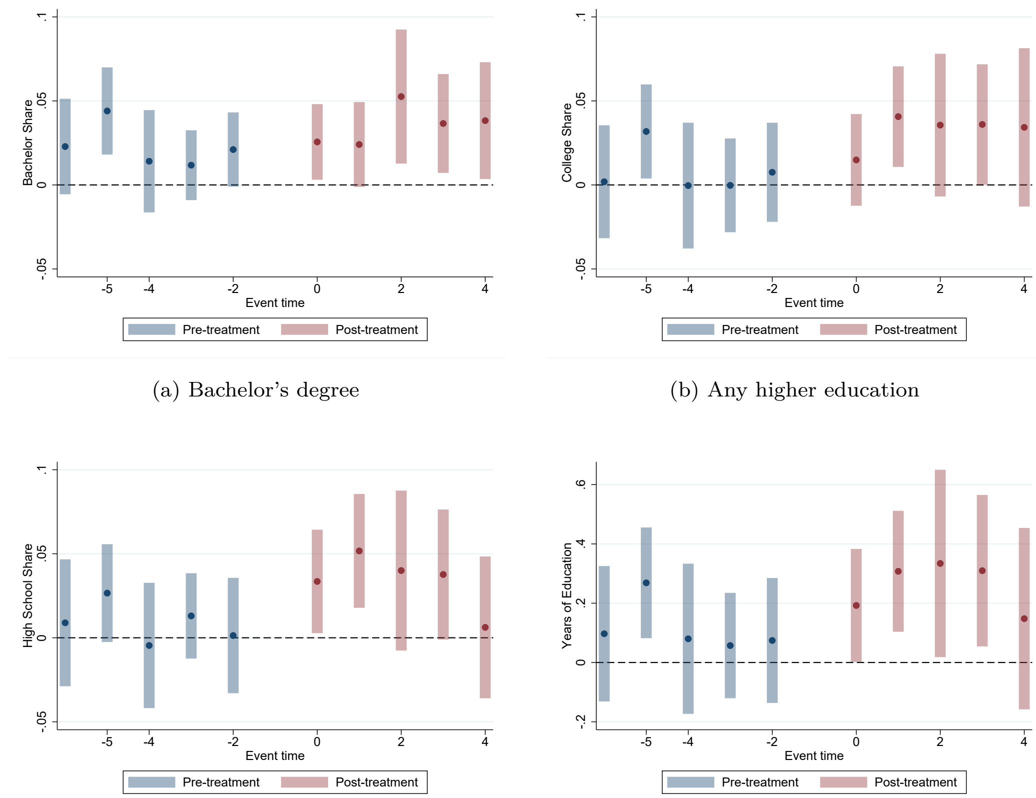
(c) High school completion



(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. The sample is restricted to individuals whose hukou remains in the same county. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with standard errors clustered at the county level. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

Figure A5: Event-Study Estimates with City-Level Clustering



(a) Bachelor's degree

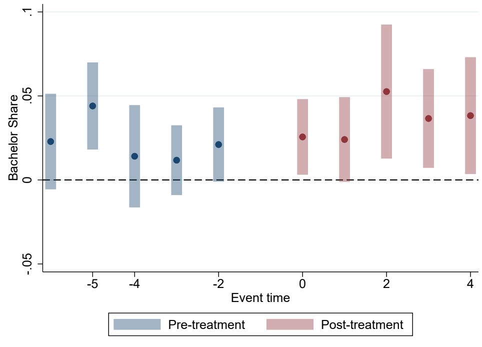
(b) Any higher education

(c) High school completion

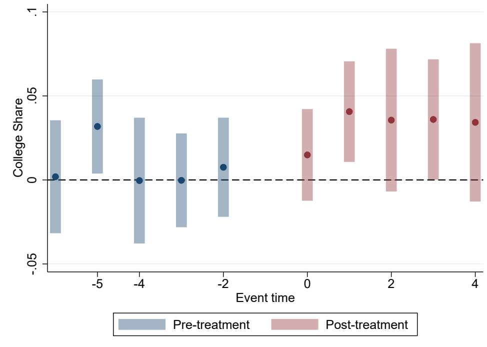
(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with standard errors clustered at the city level. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.

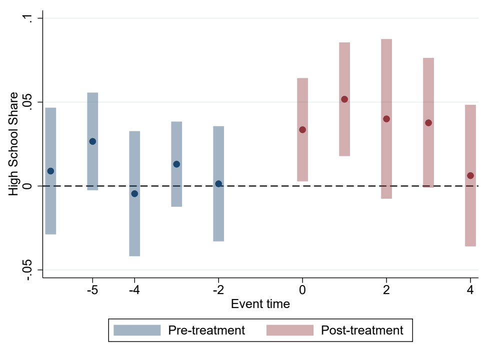
Figure A6: Event-Study Estimates with Province-Level Clustering



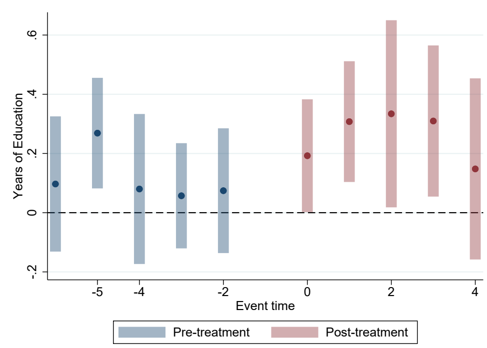
(a) Bachelor's degree



(b) Any higher education



(c) High school completion



(d) Years of schooling

Notes: Each panel plots event-study estimates from the CSDID framework. Points are estimates and vertical bars denote 95% confidence intervals. All specifications include county and cohort fixed effects and county-by-cohort controls, with standard errors clustered at the province level. The horizontal axis is event time defined as the birth cohort's distance from the top scorer's birth cohort within the same county: $t = -1$ denotes the top scorer's own birth cohort, and $t = 0$ denotes the cohort born one year after (i.e., one year younger than) the top scorer.