Online Appendix

Arrival of Young Talent: The Send-down Movement and Rural Education in China

Yi Chen Ziying Fan Xiaomin Gu Li-An Zhou

Appendix A: Additional Figures and Tables

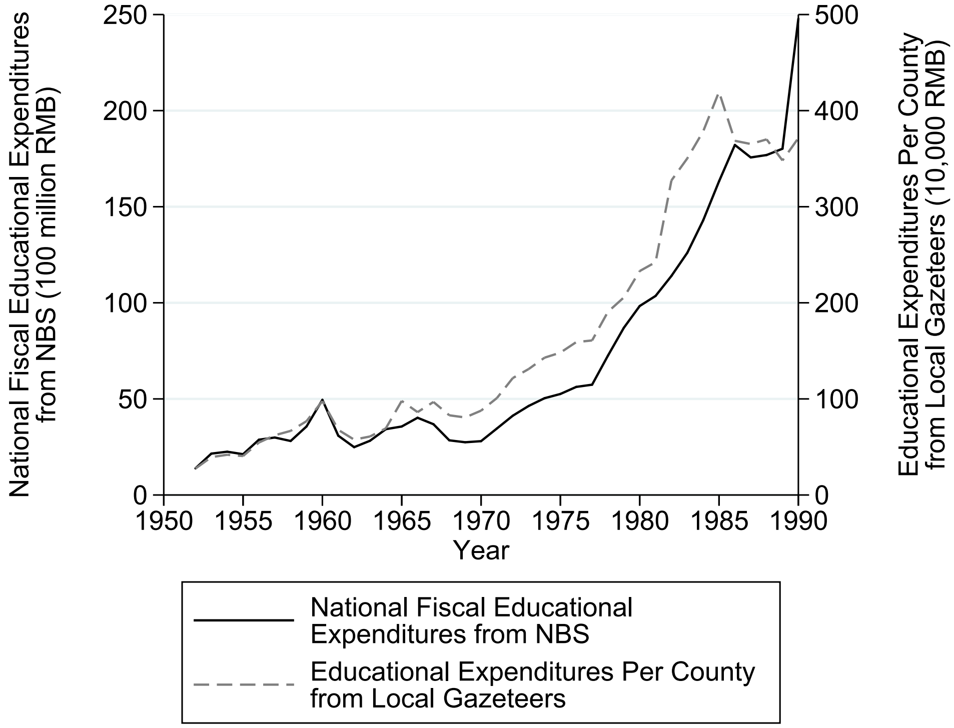
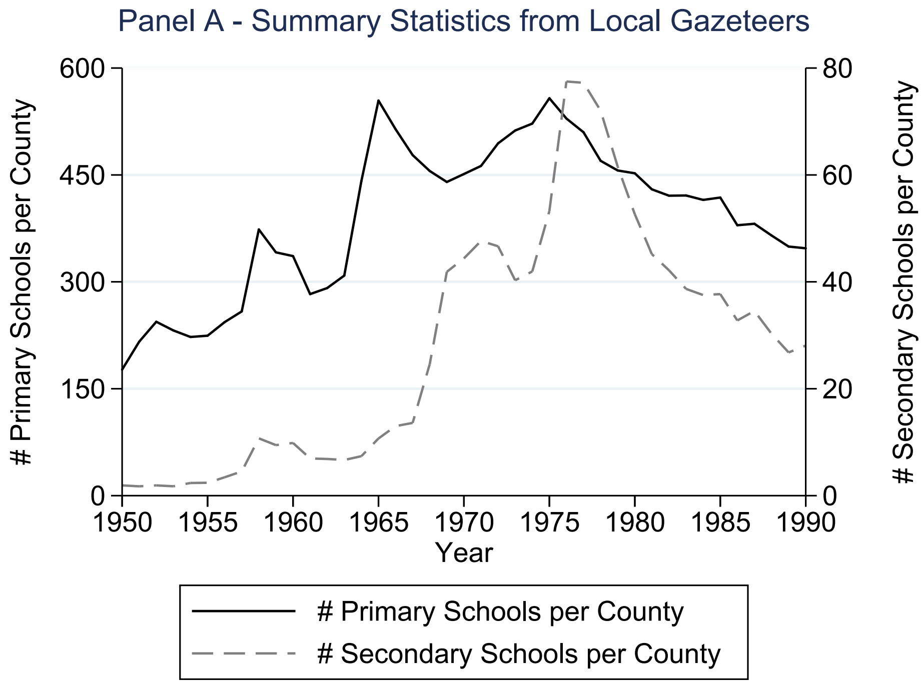


Figure A1: Trends of Real Educational Expenditures in Local Gazetteers



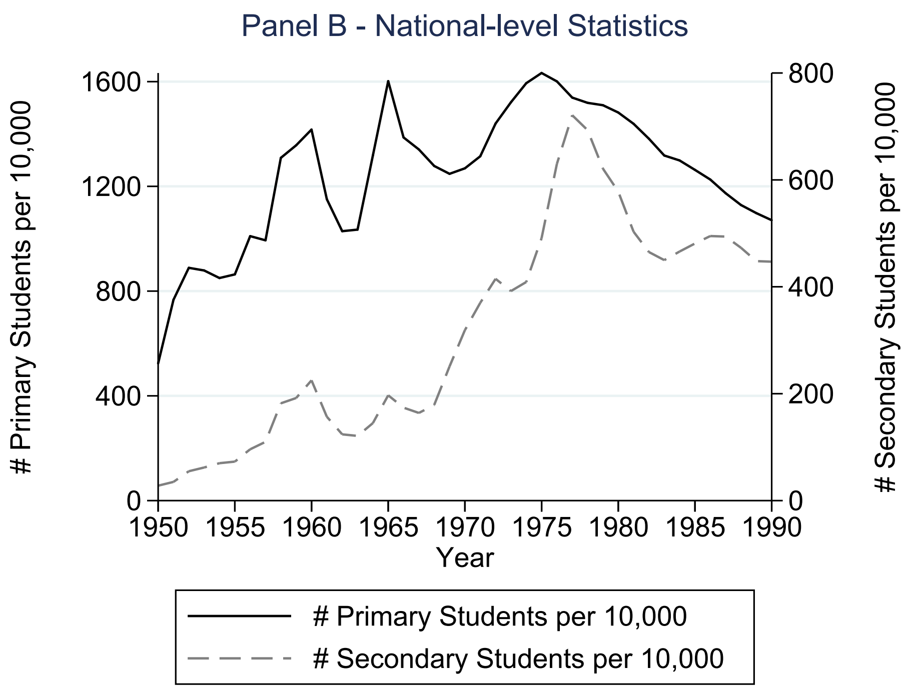


Figure A2: The Process of China's Secondary Education Expansion since the Late 1960s

Table A1: Education Gap and the Effect of SDYs (1990 Census)

Dependent Variables	Years of Education		
	(1)	(2)	(3)
Local Density of Received SDYs	3.215	3.989	4.720
*Affected Cohorts (1956–1969)	(0.800)	(0.796)	(0.900)
Density of SDYs*Affected Cohorts	0.937		
*Edu Gap from County Urban Mean[a]	(0.530)		
Density of SDYs*Affected Cohorts		1.397	
*Edu Gap from Prefecture Urban Mean[a]		(0.628)	
Density of SDYs*Affected Cohorts			2.076
*Edu Gap from Province Urban Mean[a]			(0.708)
Observations	2,594,679	2,773,142	2,775,858
R-squared	0.290	0.293	0.293
Individual Controls	✓	✓	✓
County FE	✓	✓	✓
Province-cohort FE	✓	✓	✓
Base Education×cohort FE	✓	✓	✓

Note: Only rural sample is used. Standard errors are clustered at the county level. Individual controls include gender and ethnicity. Local density of received SDYs is computed by dividing the number of received SDYs by the county population in 1964. Base education is calculated as the primary and junior high graduation rates of the control group.

a. The educational gap and SDY density in the triple-interaction term use the deviation from the sample mean. The double-interaction terms betwee affected cohorts and educational gap are included in the regressions but not repored.

Table A2: Questions for Generating the LOC Index

Variables	Survey Questions: How much do you agree with the following statement: 1 (strongly disagree)–5(strongly agree)
Internal Local of Control	
Education	The higher level of education one receives, the higher the probability of his/her future success.
Talent	The most important factor affecting one's future success is his/her talent.
Effort	The most important factor affecting one's future success is his/her effort.
Hard Work	In today's society, hard work is rewarded.
Intellect	In today's society, intellect is rewarded.
External Local of Control	
Family Socioeconomic Status	The higher a family's social status is, the greater the child's future achievement will be; the lower a family's social status is, the smaller the child's future achievement will be.
Family Wealth	A child from a rich family has a better chance of succeeding in the future; a child from a poor family has a worse chance of succeeding in the future.
Family Connection	The most important factor affecting one's future success is whether his/her family has "connection."
Luck	The most important factor affecting one's future success is his/her luck.
Connection	In today's society, having social connections is more important than having individual capability.

Table A3: The Effect of SDYs on Occupational Choice (1990 Census)

Dependent Variables	Legislators, senior officials, managers	Professionals	Technicians, associate professionals	Clerks	Service workers, shop/market sales workers	Skilled agricultural/fishery workers	Crafts and related trades workers	Plant/machine operators and assemblers	Elementary occupations
ISCO-88 skill level	-	4th	3rd	2nd	2nd	2nd	2nd	2nd	1st
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Local Density of Received SDYs *Affected Cohorts (1956–1969)	-0.0510	0.00155	0.0627	-0.00289	0.0115	-0.214	0.150	0.0371	0.00567
	(0.0121)	(0.00624)	(0.0146)	(0.00873)	(0.0169)	(0.0758)	(0.0628)	(0.0203)	(0.0132)
Observations	2,645,751	2,645,751	2,645,751	2,645,751	2,645,751	2,645,751	2,645,751	2,645,751	2,645,751
R-squared	0.015	0.008	0.009	0.015	0.025	0.208	0.158	0.038	0.025
$\overline{Y}$ of Control Group	0.00851	0.00235	0.0160	0.00478	0.0144	0.893	0.0414	0.0130	0.00669
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Province-cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Base Education×cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Only rural sample is used. Standard errors are clustered at the county level. Individual controls include gender and ethnicity. Local density of received SDYs is computed by dividing the number of received SDYs by the county population in 1964. Base education is calculated as the primary and junior high graduation rates of the control group.

Appendix B: Missing Values in Local Gazetteers

Our empirical analysis involves different pieces of information from local gazetteers. Aside from the key dependent variable—the number of received SDYs, we also collect other information for a complementary analysis, including the number of primary/secondary school teachers, grain production in 1965, educational expenditures, number of primary/secondary schools, and victims during the Cultural Revolution. Because each county compiles its own local gazetteer, the way of recording history is likely to be different across counties; therefore, missing values are inevitable. In this appendix, we show the patterns of missing values and estimate how the missing pattern is related to some county characteristics.

Two factors can affect whether a certain piece of information is available in one gazetteer. The first is political sensitivity. The issue can be especially relevant to some political events (e.g., the Cultural Revolution). The second factor is the power of the local government. Compiling a local encyclopedia is no easy task and requires both manpower and material resources. Therefore, a more powerful local government may be able to document a more detailed history of the locality. Note that even if the missing pattern in local gazetteers is nonrandom, it does not necessarily threaten our identification that relies on the parallel-trend assumption.

The county-level information used in our analysis falls into two categories. In the first category, there is only one number for each county (number of received SDYs, grain production in 1965, school construction rate, victims during the Cultural Revolution). Defining information availability is straightforward for this category—a dummy variable indicating whether that piece of information is available or not. In the second category, there is a time series for each county (number of teachers and education expenditures). While some counties' records cover an entire period, other counties only have statistics in several year increments. We define the information availability of those pieces of information as the shares of available years in a given period (1955–1977). The number equals one if and only if a county has a complete time series.

Panel B of Table B1 presents the availability of various information for our core sample (1,773 counties). Most of them record victims during the Cultural Revolution (1,760 counties) and grain production in 1965 (1,252 counties). A smaller share (806) can impute the school construction rate during the period of the Cultural Revolution. Only 592 (776) counties kept some records on the number of teachers (educational expenditures). Conditional on having at least one year's record, the time series is 47.5%–65.5% complete. Table B2 examines the correlation between information availability and various county characteristics: densities of SDYs, severity of the Great Famine,[1] educational level (primary and junior high school graduation rates computed from the 1990 census), and share of minorities in 1990. We also report in Panel B of Table 6 the effect of SDYs conditional on the availability of various county-level information—all the estimates are highly significant.

[1]Because we use a relative cohort size approach to construct the severity, we can compute this number for all counties.

Table B1: Count of Number of Counties

Panel A: Core Counties	
Counties in 1990 census	2521
Exclude Beijing, Tianjin, Shanghai	2469
Exclude city-governed districts	2039
Exclude counties w/o information on SDYs	1843
Exclude counties w/o population in 1964	1773
Panel B: Counties with Other Information	
Counties with at least one year's information on teachers	592
Share of years with information (1955–1977)	47.5%
With information on grain production in 1965	1252
Counties with at least one year's information on educational expenditures	776
Share of years with information (1955–1977)	65.5%
Counties with primary and secondary schools construction rate	806
With information on the Cultural Revolution	1760

Note: The share of years with information on teachers (1955–1977) is calculated as (county-year observations on primary school teachers + county-year observations on secondary school teachers)/(2*23*(number counties with at least one observation on either primary school or secondary school teachers)). Shares of years with information on the educational expenditures (1955–1977) are computed in a similar fashion.

Table B2: Correlation between County-level Information Availability and County Characteristics

Dependent Variables	SDYs Density	SDYs Density	Info on Teacher Numbers (1955–1977)	Info on Grain Production in 1965	Info on Educational Expenditures (1955–1977)	Info on School Construction Rate	Info on Cultural Revolution
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SDYs Density			-0.278	-0.765	-0.788	0.0473	-0.114
			(0.206)	(0.436)	(0.269)	(0.489)	(0.123)
Local Severity of Cultural Revolution		0.0827					
		(0.0298)					
Local Severity of Great Famine	0.00104	0.00127	-0.0437	0.0159	0.00417	0.00550	-0.00458
	(0.00381)	(0.00381)	(0.0347)	(0.0505)	(0.0427)	(0.0536)	(0.00764)
Primary Graduation Rate	-0.00195	-0.000401	-0.0942	-0.0948	-0.0407	-0.00919	-0.0308
(Control Group)	(0.00514)	(0.00503)	(0.0644)	(0.0914)	(0.0789)	(0.101)	(0.0201)
Junior High Graduation Rate	0.0238	0.0227	0.0369	0.119	-0.000886	-0.0712	0.0202
(Control Group)	(0.0122)	(0.0122)	(0.116)	(0.161)	(0.146)	(0.180)	(0.0310)
Share of Minority (in 1990)	-0.00330	-0.00289	0.0499	-0.0509	-0.0112	0.00147	-0.000793
	(0.00216)	(0.00212)	(0.0318)	(0.0495)	(0.0430)	(0.0526)	(0.00688)
Observations	1,763	1,750	1,763	1,763	1,763	1,763	1,763
R-squared	0.257	0.266	0.054	0.066	0.091	0.053	0.042
Province FE	✓	✓	✓	✓	✓	✓	✓

Note: Robust standard errors are in parentheses.

Appendix C: Data Quality in Local Gazetteers

Scholarly concerns may be raised regarding the accuracy of the numbers of SDYs in local gazetteers. We argue that the numbers are generally trustworthy for the following reasons. When the local gazetteers were compiled, the send-down movement, as opposed to the Great Famine and the Cultural Revolution, was not a politically sensitive issue. Even to this day, there are no official statistics from the Chinese government about the fatalities for either the Great Famine or the Cultural Revolution. By contrast, numerous official documents about the send-down movement are publicly available (see Gu (2009) for a comprehensive collection). More importantly, under Chairman Mao's instructions in 1968 that "the comrades in the countryside should welcome them," receiving those SDYs became an important political task. Therefore, local governments usually kept a good record of how many urban youths they received during the movement.

To provide further evidence on the quality of information related to SDYs, we use Benford's Law to detect the possible statistical anomalies in our county-level data (Benford, 1938). The Benford Law is an empirical observation that the probability of a number having a particular non-zero first digit d is roughly $\log_{10}(1 + 1/d)$. This statistical tool has been widely applied to test survey data quality in developing countries as well as in the U.S. (Judge and Schechter, 2009; Michalski and Stoltz, 2013; Holz, 2014). In Figure C1, we plot the distribution of the first digits of received SDYs numbers and compare to the distribution suggested by Benford's Law. Our county-level data on SDYs nicely conform to the law. Note that each county only contributes one observation to our dataset. It is unlikely that county governments are aware of the statistical law and coordinate together in falsifying those numbers.

We also made comparisons among three sources of data: our county-level data, national reports documented in Gu (2009), and estimates from the 2010 wave of the CFPS.[2] Figure C2 compares the time series of nationally reported numbers and those of CFPS estimates. It is evident that the CFPS not only matches well the size of the sent-down population but also mimics the temporal fluctuations. This result lends credibility to national reports on the send-down movement—the CFPS was carried out by an academic institute (Peking University) forty years after the send-down movement.

In Table C1, we compare our county-aggregation at the province level to those from a national report, which is documented in Gu (2009, pp.259–260). The ratio varies from 60%–80% for most provinces. Three reasons can account for the gap. First, national statistics cover a longer time span than the county-aggregate (1962–1979 versus 1968–1977, respectively). If we consider the differential length in time, the ratio becomes 72%–96%. Second, the local gazetteers only cover SDYs received by the rural villages and collective farms, which means that 2.9 million out of 17.7 million SDYs who were sent to state farms during the movement were not recorded in local gazetteers. We verified that the SDYs sent to state farms were separately recorded in the farm chronicles (*nongchangzhi*). State farms did not distribute evenly across the country. The largest state farm organization was the Xinjiang Production and Construction Army Group (*Xinjiang Jianshe Bingtuan*). Heilongjiang also had many state farms because of China's intense relationship with the Soviet Union during the 1960s. These factors explain why the ratios are especially low in Xinjiang (39.4%) and Heilongjiang (26.6%). The third reason is that we excluded more-developed city-governed districts. We also did a robustness check that excludes five provinces that do not perform well in matching county-aggregation and national statistics (ratio below 50%): Shanxi, Heilongjiang, Yunan, Ningxia, and Xinjiang. Following appendix provides more details.

[2]The CFPS provides sample weight, which can be interpreted as the number of populations each observation can represent to allow for the national representativeness. Therefore, summing up the sample weight of the observations who report being sent down yields an estimate for the number of SDYs.

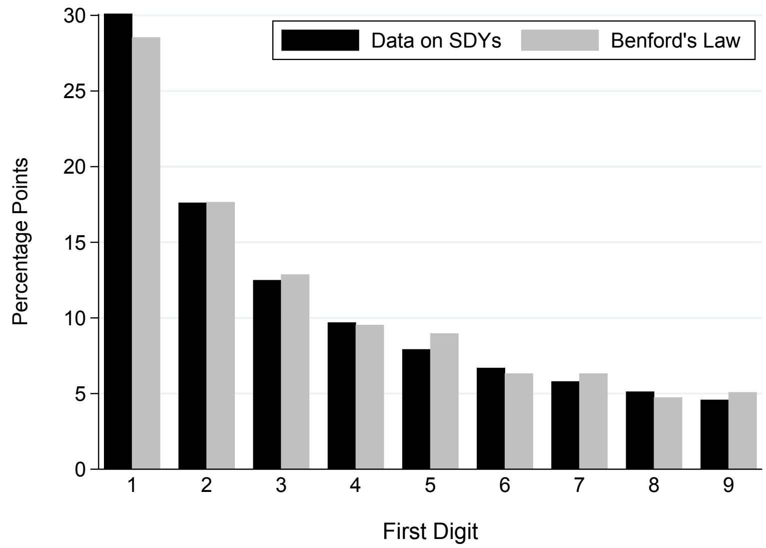


Figure C1: Benford's Law and Data Quality on SDYs ($\chi^2\,(9) = 6.96$, p-value $= 0.5413$)

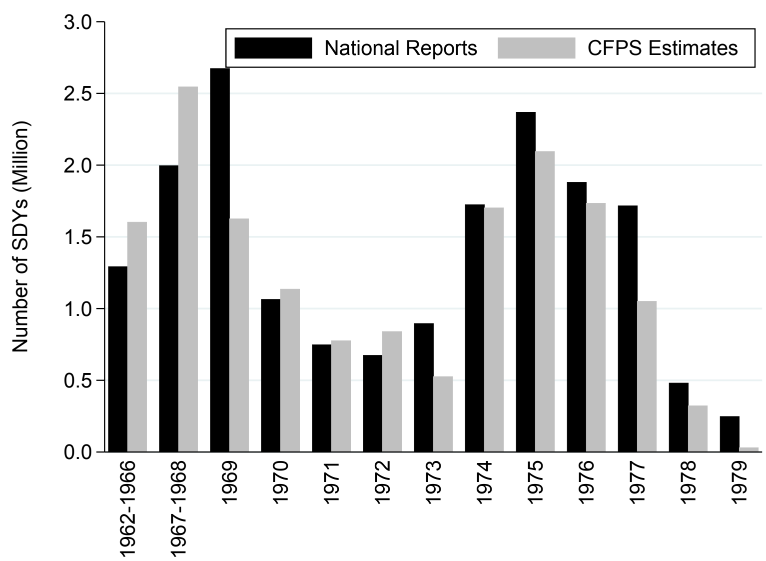


Figure C2: Number of SDYs Estimated from CFPS 2010

Table C1: Comparing the Number of Received SDYs from County-aggregate with that from National Report in Each Province

Province	SDY Received (Thousand)		
	County Aggregate 1968–1977	National Report 1962–1979	Ratio (%)
Hebei	297.2	510.5	58.2
Shanxi	141.9	312.9	45.3
Inner Mongolia	319.0	299.3	106.6
Liaoning	1385.4	2018.0	68.6
Jilin	628.7	1052.6	59.7
Heilongjiang	512.1	1922.2	26.6
Jiangsu	494.6	861.2	57.4
Zhejiang	435.9	595.9	73.1
Anhui	511.6	725.5	70.5
Fujian	323.3	372.3	86.8
Jiangxi	383.4	622.5	61.6
Shandong	397.5	492.7	80.7
Henan	430.9	673.0	64.0
Hubei	654.0	878.6	74.4
Hunan	583.0	635.8	91.7
Guangdong	519.1	973.2	53.3
Guangxi	293.6	434.8	67.5
Sichuan	1243.7	1427.4	87.1
Guizhou	157.6	224.1	70.3
Yunnan	165.0	339.1	48.7
Shaanxi	376.9	490.3	76.9
Gansu	162.7	264.3	61.6
Qinghai	32.2	51.0	63.1
Ningxia	20.7	57.5	36.0
Xinjiang	163.9	416.6	39.4
Total	10,634.0	16,651.3	63.9

Note: The county aggregate numbers are computed based on authors' collection of data from local gazetteers. Numbers from the national report are the same as those in the first column of Table 1.

Appendix D: Robustness Check

Different Bandwidths of Treated Cohorts

Our main specification focuses on cohorts born between 1946 and 1969, and cohorts 1956–1969 are defined as the treatment group. That is, our bandwidth of the treatment cohort is 14 years. The rationale of such choice is that cohort 1956 was receiving their last year of primary education at the beginning of the massive rustication movement, while cohort 1969 had just started their primary school when the movement came to an end. This approach assumes a standard procedure of school attendance: children started primary school at the age of seven and spent six years in primary school. However, the compliance may be incomplete. Columns (1)–(3) in Table D1 replicates the results from Table 3 using different bandwidths for treated cohorts (10 years, 7 years, and 4 years). In terms of statistical significance, Table D1 yields exactly the same results regardless of the choice of bandwidth. In terms of the magnitude of the coefficients, the results also match the findings from Figure 3, which displays a reversed-U-shaped pattern. Exclusion of the last few SDY-affected cohorts (bandwidth = 10, 7 years) makes the coefficients a bit larger (from 3.237 in Table 3 to 3.260 and 3.351) because those cohorts were exposed to SDYs only for several years at the start of primary education. Following the same logic, only focusing on the first few affected cohorts (bandwidth = 4 years) makes the coefficients smaller (from 3.237 to 2.654).

Alternative Denominator for the Densities of SDYs

Our main specification uses the total county population in 1964 as the denominator to compute the density of SDYs. One may think the population of schooling-age children can be a sharper and better denominator. However, micro-level data from the 1964 census are unavailable, and we only have access to aggregate population figures at the county level. We adopt this idea in a slightly different way. We compute the size of the affected cohorts in each county using the 1% sample of the 1990 census. Note that this approximation is subject to three potential biases: sampling error (only a 1% sample is available), migration selection (fortunately not very likely at that time), and mortality attrition. Using the alternative denominator still yields highly significant gains in education as a result of the SDYs' arrival (Column (4) in Table D1).

Using a Continuous Treatment Variable

We use a dummy treatment variable in our cohort DID specification. Figure 3 suggests that the treatment effect is about proportional to the overlapped years with the SDYs during one's primary education period. Column (5) in Table D1 uses the overlapped schooling years as an alternative (continuous) measure of the exposure to SDYs. The estimated effects are still highly significant. However, we would still prefer the cohort-dummy specification as our baseline because it is more flexible and requires fewer assumptions on the functional form.

Excluding Cohorts 1953–1955

Cohorts 1953–1955 were junior high age when the SDYs arrived. Whether we should put them in the treatment or control group may be a subject of debate. On the one hand, they were still receiving education and could benefit from the arrival of urban youths. On the other hand, as discussed in footnote 22 in the main text, the effects of SDYs on those older cohorts are relatively small for a set of reasons. Nevertheless, we dropped cohorts 1953–1955 to form a more conservative control group in Column (6) of Table D1. The result does not change at all.

Allowing SDYs to Affect Junior High Education

Following previous discussion, an alternative approach is to move cohorts 1953–1955 to the treatment group (instead of dropping them). Currently, we define exposure to the SDYs according to whether rural children's years in primary school overlap with the massive send-down movement. The justification is based on the fact that most rural children in China during that period received at most primary-level education. Still, about one-fourth went to junior high school. If we allow the flow of SDYs to affect not only primary school but also junior high school, the treatment cohorts would be extended by three years and become 1953–1969. Columns (7) and (8) in Table D1 report the results using this alternative definition of exposure, with Column (8) using a continuous measure of overlapped years. The coefficients stay positively significant but become smaller, which should be expected because a smaller share of rural residents would attend junior high school.

Excluding Nine Provinces in the "Third Front" Construction Region

Starting in 1964, China constructed hundreds of large manufacturing plants in its southwest and northwest hinterland known as the "third front region" (TF region hereafter). The construction was a response to the deteriorated relationship with the Soviet Union and the escalated Vietnam War, and its goal was to establish self-sustaining industrial clusters defending against potential military attacks. The TF regions received massive investment (mostly in manufacturing sectors) between 1964 and the mid-1970s (Fan and Zou, 2019). To the best of our knowledge, the plants were separate from rural villages. Nevertheless, local spillovers might exist. Because the TF construction concentrated in the southwestern and northwestern part of China, one simple way to examine the robustness of our main results against the possible influence of TF construction is to exclude nine provinces in the TF region (Sichuan, Guizhou, Yunan, Gansu, Ningxia, Shaanxi, Guangxi, Hubei, and Hunan). Columns (1) and (2) in Table D2 show that excluding those nine provinces barely changes our estimation.

Excluding Five Provinces—Shanxi, Heilongjiang, Yunnan, Ningxia, and Xinjiang

We document in Table C1 that the county-level aggregates from local gazettes account for a relatively small share of received SDYs from national statistics in five provinces—Shanxi (45.3%), Heilongjiang (26.6%), Yunnan (48.7%) Ningxia (36.0%), and Xinjiang (39.4%). The small share can be partially rationalized by state farms (especially for Heilongjiang and Xinjiang). The SDYs sent to state farms were intentionally not recorded in local gazettes but recorded in the national statistics. Nevertheless, we examine the robustness of our results by excluding those five provinces in Columns (3) and (4) of Table D2. The coefficient of SDYs drops slightly from 3.237 to 2.673, which should be expected because the excluded provinces were less developed areas and benefited more from SDYs according to our heterogeneity analysis in Table 4.

Stronger Assumption on Migration

Our empirical analysis combines the census 1990 with the historical data on the SDYs' flow during the period 1968–1977. One implicit assumption is that people lived in the same county at those two points in time. Although China's population census in 2000 suggests that 86% of the sample in this study lived in the same county as their birthplace, we cannot rule out the possibility that the remaining small proportion has an important influence on our results. In our main analysis, we exclude the sample whose residence county differs from their registration county/prefecture. In this appendix, we impose an even stronger assumption: people resided in the current locality on July 1st, 1985. If migration is a real issue, the extra restriction should have an important influence on our results. Columns (5) and (6) in Table D2 give almost identical results as those in Table 3. With the additional assumption, we lose less than 1.0% of the sample. This is not surprising because migration in China was still of limited scope before 1990.

Hukou Changes from Rural to Urban

Similar to the concerns about migration, we may also be concerned with the changes of *hukou* from rural to urban because our analysis focuses on the rural sample. However, China's household registration system means that people cannot change their *hukou* status at will. They must first satisfy certain requirements. One way for rural residents to change to urban *hukou* is to become sufficiently well-educated. Graduates from technical secondary schools, colleges, and universities satisfy this requirement. Thus, junior high graduates and regular senior high graduates are ineligible. Columns (7) and (8) in Table D2 exclude the sample of sufficiently educated to change their *hukou* status and yield identical results, suggesting a limited impact of rural-to-urban *hukou* switches. These results arose because the overall level of education in rural China was still quite low, and few rural children could reach the required level of education.

Measurement Errors in the Number of Received SDYs

One final issue is measurement error in our key independent variable—numbers of received SDYs in each county. We take those numbers from local gazetteers and therefore implicitly assume that the records are accurate. However, measurement errors and recall biases are inevitable in historical documents. To evaluate the possible consequences of measurement errors, we use the following idea: if the compilers of local gazetteers did not have any specific records and had to "guess" the numbers, those numbers are more likely to end with zero. Those counties should account for 10% of the sample if the last digit is randomly distributed. This share is 18% in our data, suggesting the existence of measurement error. Columns (9) and (10) in Table D2 drop those counties whose numbers of received SDYs end with zero. The impact of SDYs actually becomes larger in this case. It is a classic econometric result that measurement errors make the coefficients downward biased. Therefore, the true magnitude of the treatment effect should be even larger if historical documentation could be more accurate.

Table D1: Robustness Check with Different Specifications (1990 Census)

Dependent Variables	Years of Education							
Robustness Checks	Affected Cohorts=1956–1956+N			Alternative Density Measure, Treatment Measure, and Control Group			Allow Junior High to be Affected	
	N=10	N=7	N=4					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Rural Sample								
Local Density of Received SDYs	3.260	3.351	2.654					
*Affected Cohorts	(0.711)	(0.726)	(0.661)					
Alternative Density of Received SDYs[a]				0.750				
*Affected Cohorts (1956–1969)				(0.193)				
Local Density of Received SDYs					0.628			
*Overlapped Years in Primary School					(0.135)			
Local Density of Received SDYs						3.084		
*Affected Cohorts (cohorts 1953–1955 excluded)						(0.753)		
Local Density of Received SDYs							2.366	
*Affected Cohorts (junior high affected)[b]							(0.674)	
Local Density of Received SDYs								0.459
*Overlapped Years in Primary & Junior High[b]								(0.109)
Observations	2,285,274	1,814,427	1,437,195	2,775,858	2,775,858	2,432,867	2,964,596	2,964,596
R-squared	0.302	0.305	0.288	0.293	0.293	0.293	0.302	0.302
Panel B: Urban Sample								
Corresponding Coefficients to Panel A	0.297	0.594	0.680	-0.00831	0.0527	-0.00365	-0.234	0.0459
	(0.515)	(0.525)	(0.519)	(0.273)	(0.0942)	(0.661)	(0.682)	(0.0865)
Observations	354,271	287,480	226,630	417,883	417,883	367,763	451,299	451,299
R-squared	0.223	0.223	0.200	0.225	0.225	0.232	0.229	0.229
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Province-cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Base Education×cohort FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: Standard errors are clustered at the county level. Individual controls include gender and ethnicity. Local density of received SDYs is computed from dividing the number of received SDYs by the county population in 1964. Base education is calculated as the primary and junior high graduation rates of the control group.

a. Alternative density of received SDYs uses the population size of the treatment cohort in each county (calculated using the 1990 census) as the denominator.

b. The treatment group become cohorts 1953–1969, and the control group are cohorts 1943–1952. Both groups are extended by three cohorts because we allow junior high education to be affected.

Table D2: Other Robustness Checks (1990 Census)

Dependent Variables	Years of Education									
Robustness	Exclude 9 Provinces in the "Third Front" Construction Region		Exclude Shanxi, Heilongjiang, Yunnan, Ningxia and Xinjiang		Stayed in the County/Prefecture for at Least Five Years		Exclude Graduates from Technical Secondary School/from College		Exclude Counties whose Last Digit of SDY number is Zero	
Sample	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Local Density of Received SDYs	2.871	0.745	2.673	-0.0817	3.184	0.325	3.194	0.546	3.692	0.712
*Affected Cohorts (1956–1969)	(0.771)	(0.519)	(0.751)	(0.574)	(0.708)	(0.503)	(0.699)	(0.487)	(0.831)	(0.656)
Observations	1,791,139	269,636	2,513,043	367,548	2,750,293	381,249	2,770,356	345,583	2,301,166	344,217
R-squared	0.272	0.232	0.290	0.230	0.294	0.199	0.294	0.259	0.296	0.231
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Province-cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Base Education×cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Standard errors are clustered at the county level. Individual controls include gender and ethnicity. Local density of received SDYs is computed by dividing the number of received SDYs by county population in 1964. Base education is calculated as the primary and junior high graduation rates of the control group.

Appendix E: Synthetic Control Method

The possible heterogeneous cohort trends in the absence of treatment is a crucial concern in our identification strategy. To alleviate this concern, we first control for a comprehensive set of (non-linear) cohort trends, including the province-cohort fixed effects and cohort dummies interacted with county base education. We also provide evidence that there are no pre-existing trends (Figure 3), and that our results are robust to a wide range of contemporaneous events (Table 6). We cannot exhaust all differential trends because many county characteristics are unobservable. We therefore present in this appendix an alternative empirical method—the Abadie synthetic control (SC) method (Abadie and Gardeazabal, 2003; Abadie, Diamond and Hainmueller, 2010), which does not require the parallel-trend assumption.

The SC method was originally designed to estimate the treatment effect of interventions that are implemented at an aggregate level that affect a small number of large units (such as cities, regions, or countries) (Abadie, forthcoming). Our empirical setting is different from a standard Abadie SC setting in two ways. First, we use micro-level census data instead of aggregate data. Second, the treatment (SDYs' arrival) affected the entire country, and almost all rural counties received some SDYs during the movement. The variations rest on the *intensity* of treatment, whereas the treatment in typical SCs is a dummy variable.

To apply the SC method, we need to aggregate the micro-level census data into county-by-cohort data. Because we need an accurate estimate for the average number of years of education for each county-cohort cell, we restrict our sample to large counties that can guarantee a minimum of 30 observations for each cohort born between 1946 and 1969. We ended up with 439 large counties with complete information. We then bisected the sample according to the local density of SDYs into a treatment pool and a donor pool. This allows us to extend the SC method from a single treatment to multiple treatments, following Cavallo et al. (2013).

Here we briefly describe the procedures of the extended SC method.[3] In the first step, we take one county from the treatment pool and construct the SC county using the donor pool, which is a standard Abadie SC approach.[4] In the second step, we repeat the first step for every county in the treatment pool and compute the average treatment effect. Lastly, we compute p-values using the placebo treatment effects for inference purposes.[5]

Figure E1 depicts the SC results. The gap in average education is small between the treatment group and the SC counties before the arrival of the SDYs. The gap starts to emerge and gradually expands afterwards. The pattern of Abadie SC estimates mimics that of our main results in Figure 3.

[3]Cavallo et al. (2013) and Galiani and Quistorff (2017) provide further details.

[4]When choosing the county characteristics used to match counties before the SDYs' arrival, we try to exhaust all available information that does not lead to a significant reduction in the sample size. For example, the number of teachers is not a good candidate because we have this information for only one-third of the counties. The characteristics include the average years of education for each control cohort (1946–1955), primary and junior high school graduation rates of the control group, share of urban population in 1964, share of minority population (imputed with 1990 census data), grain output in 1965 (scaled by county population in 1964), and local intensity of the Great Famine (see Section IV.E for details).

[5]See Cavallo et al. (2013) for technical details.

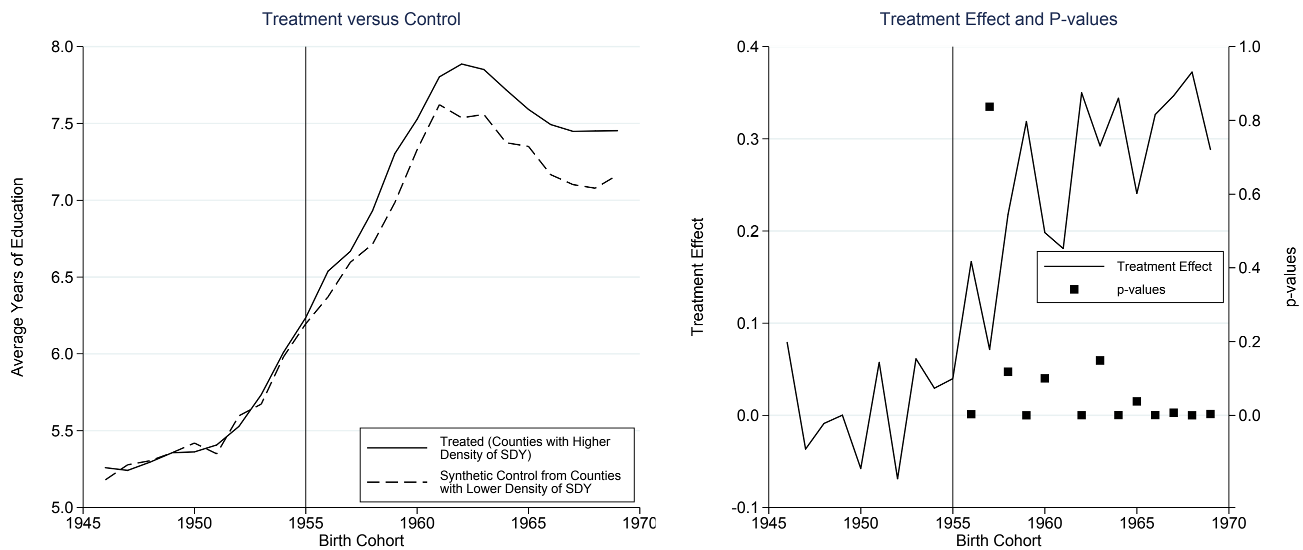


Figure E1: Estimating the Effect of SDYs using the Synthetic Control Method (1990 Census)

References

Abadie, Alberto. n.d.. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature* (forthcoming).

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493–505.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*, 93(1): 113–132.

Benford, Frank. 1938. "The law of Anomalous Numbers." *Proceedings of the American Philosophical Society*, 78(4): 551–572.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *Review of Economics and Statistics*, 95(5): 1549–1561.

Fan, Jingting, and Ben Zou. 2019. "Industrialization from Scratch: The "Construction of Third Front" and Local Economic Development in China's Hinterland." Working Paper.

Galiani, Sebastian, and Brian Quistorff. 2017. "The Synth_Runner Package: Utilities to Automate Synthetic Control Estimation Using Synth." *The Stata Journal*, 17(4): 834–849.

Gu, Hongzhang. 2009. *Chinese Educated City Youth: The Whole Story (zhongguo zhishi qingnian shangshan xiaxiang shimo, in Chinese).* People's Daily Publishing House.

Holz, Carsten A. 2014. "The Quality of China's GDP Statistics." *China Economic Review*, 30: 309–338.

Judge, George, and Laura Schechter. 2009. "Detecting Problems in Survey Data using Benford's Law." *Journal of Human Resources*, 44(1): 1–24.

Michalski, Tomasz, and Gilles Stoltz. 2013. "Do Countries Falsify Economic Data Strategically? Some Evidence that They Might." *Review of Economics and Statistics*, 95(2): 591–616.