Educated Youth Should Go to the Rural Areas: A Tale of Education, Employment and Social Values*

Yang You†

Harvard University

This draft: July 2018

Abstract

I use a quasi-random urban-dweller allocation in rural areas during Mao's Mass Rustication Movement to identify human capital externalities in education, employment, and social values. First, rural residents acquired an additional 0.1-0.2 years of education from a 1% increase in the density of sent-down youth measured by the number of sent-down youth in 1969 over the population size in 1982. Second, as economic outcomes, people educated during the rustication period suffered from less non-agricultural employment in 1990. Conversely, in 2000, they enjoyed increased hiring in all non-agricultural occupations and lower unemployment. Third, sent-down youth changed the social values of rural residents who reported higher levels of trust, enhanced subjective well-being, altered trust from traditional Chinese medicine to Western medicine, and shifted job attitudes from objective cognitive assessments to affective job satisfaction. To explore the mechanism, I document that sent-down youth served as rural teachers with two new county-level datasets.

Keywords: Human Capital Externality, Sent-down Youth, Rural Educational Development, Employment Dynamics, Social Values, Culture

JEL: A13, N95, O15, I31, I25, I26

---

* This paper was previously titled and circulated, "Does living near urban dwellers make you smarter" in 2017 and "The golden era of Chinese rural education: evidence from Mao's Mass Rustication Movement 1968-1980" in 2015. I am grateful to Richard Freeman, Edward Glaeser, Claudia Goldin, Wei Huang, Lawrence Katz, Lingsheng Meng, Nathan Nunn, Min Ouyang, Andrei Shleifer, and participants at the Harvard Economic History Lunch Seminar, Harvard Development Economics Lunch Seminar, and Harvard China Economy Seminar, for their helpful comments. I also thank Yihan Zhao, Alice (Nan) Wu, Xinwei Xu, Jialin Li, Yu Luo, Sixue Liu, Vivian (Siwei) Xu for excellent research assistance work and archival data collection.

† Corresponding Author: yangyou@g.harvard.edu Harvard University Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138.

Educated youth should go to the rural areas and be re-educated by peasants. It is necessary!

--------Zedong Mao, December 22, 1968

1. Introduction

Nearly 16.5 million urban youths—about 10 percent of the entire urban population of China—were forcibly sent to Chinese rural areas from 1968 to 1980, a period known as Mao's Mass Rustication Movement (MacFarquhar and Schoenhals 2009). The rustication relocation is the largest urban-to-rural migration known in human history. I study the event of moving a population of more-educated urban youth into Chinese rural areas and estimate the effect of that relocation on the education, labor outcomes, and social values of youth in the rural areas.

Using archival data from 61 rural counties of Heilongjiang Province, I construct the density of "sent-down" urban youth, as they were called. I define the sent-down youth density as the ratio of the number of sent-down youth in 1969[1] to the population size of each rural county in the 1982 Chinese Census[2]. Two birth cohort exposure dummies are constructed based on two watershed events—the onset of the Mass Sent-down movement in 1968, and the "Returning Cities" rehabilitation in 1980. The variables indicate whether rural youths were or were not exposed during their school years to the sent-down urban youth. I then estimate human capital externalities with a difference-in-difference approach by interacting the sent-down youth density with the two exposure dummies.

I find that rural youths received more education if their county was exposed to a higher density of sent-down youth. As opposed to the Mao's quote, urban people became teachers and educated rural people. Compared with low-density counties, the six highest-density counties experienced an extra 0.8 years of educational improvement in cohorts with exposure to sent-

[1] I choose numbers of sent-down youth in 1969 for the following two reasons: 1. I only find centralized sent-down youth data for 1969, which covers 61 out of 63 rural counties, in Heilongjiang Archive. Data for other years are incomplete, typically comprising less than 40% of rural counties. I construct density with 1969 numbers to minimize the sample selection bias risk and maximize sample size. 2. The number of sent-down youth in 1969 is 91.6% positively correlated with the total amount in the period 1968-1978 based on 13 counties with complete annual sent-down youth data in county gazetteers—the county-level historical archives.

[2] The annual county-level population statistics was incomplete in 1969. The Bureau of Statistics malfunctioned during the Culture Revolution 1966-1976. Thus, we use the population data from the first official census after the Culture Revolution for normalization.

down youth during their school years. I then estimate a difference-in-difference specification to show that a one percent increase in sent-down youth density translates to an increase of 0.1-0.2 years of education in the 1990 Census. Using the same identification strategy, I find that the density of sent-down youths reduces the primary school dropout rate but increases the high school dropout rate.

Then, I estimate the labor market outcomes of urban youth exposure with the difference-in-difference approach. In 1990, exposure to sent-down youth led to lower non-agricultural employment. One possible explanation is that political campaigns during Mao's Mass Rustication Movement caused the significant drop in educational quality. People educated during the rustication era suffered from higher unemployment and lower non-agricultural employment conditional on the same educational attainment.

Conversely, in 2000, exposure to sent-down youth increased non-agricultural employment and reduced unemployment, even after controlling for the individual education level. More employment benefits from exposure to urban youth appeared after a decade of rapid urbanization[3] and technological progress. Furthermore, I also find significant heterogeneity across occupations. The employment improvement mainly was concentrated in productive jobs (e.g., factory workers, service workers, and specialists), but very modestly in bureaucratic employment (e.g., government officials and administrative staff).

Sent-down youth also changed the composition of future migration flows. Among the population who moved from 1985 to 1990, high-density counties attracted more young migrants who were exposed to sent-down youth in their schooling, both from other counties in Heilongjiang and from other provinces.

Sent-down urban youth also altered rural people's social values and beliefs. Those who were exposed to more sent-down youth during their school years, expressed stronger social trust, more scientific medical beliefs—evidenced by more confidence in Western medicine than traditional Chinese medicine—and higher levels of well-being. Measures of well-being include

[3] Appendix Figure 1 reports the annual percentage of the urban population from 1949 to 1990. The urbanization rate remained at 18% during the sent-down movement decade 1968-1978 and took off after 1980.

height, subjective health evaluation, self-reported happiness, happiness relative to peers, and feelings about the fairness of life.

Sent-down youth also shifted job attitudes from objective cognitive assessments to affective job satisfaction. People ascribed more importance to the spiritual amenities of jobs (e.g., opportunities to fully use their talents, satisfy their interests, gain more respect, and comfort themselves), but put less stress on the practical purposes of jobs (e.g., ability to make a living, and build connections). Also, urban youth undermined the rural dwellers' willingness to expend effort under unfavorable conditions. People with high exposure to urban youth were less likely to pay effort when they were in poor health conditions, when the task was undesirable, and when the task did not pay off until far in the future.

Inspired by historical narratives, I hypothesize that urban youth worked as rural teachers during the rustication period. To test for this mechanism, I collect a new county-level unbalanced panel of sent-down youth, students, teachers, and schools from Chinese county gazetteers—the county-level historical archives. On the extensive margin, I compare growth rates of educational variables from the pre-rustication decade, 1958-1967, to the rustication decade, 1968-1977. Counties with high sent-down youth densities experienced faster expansion in numbers of students, teachers, and schools; and slower growth in student-teacher ratios, which indicates that the increase in teacher size exceeded the increase in student size. On the intensive margin, I select counties without missing values in the rustication period, 1968-1978, and use the cumulative sent-down youth density to forecast the per-capita number of teachers, students, and schools with panel regressions. Higher cumulative sent-down youth density predicts more teachers, but not more students or schools. On both margins, the coefficients for teachers in secondary schools are twice as large as the coefficients for teachers in primary schools.

This paper contributes to the literature in the following dimensions. First, this paper broadly relates to the literature on human capital externality. The following table summarizes the literature by human capital flow direction: high to high, low to high, and high to low. The "high to high" literature, commonly referred to as an agglomeration economy, investigates benefits from people gathering together in cities and industrial clusters. A large body of "low to high" literature studies the outcomes of relocating disadvantaged individuals into high human capital areas (e.g., Moving to Opportunity (MTO) and resettling refugees). However, the "high to low" literature, evaluating impacts of migrants with high human capital in low human capital areas, is still quite limited.

Despite the lack of empirical investigation, many policy practices aim to instill well-educated population into low-income areas (e.g., Teach for America, the Barefoot College in India, and the University Student Village Official Program in China). Some historical events also share this feature (e.g., missionaries in Africa, Nunn et al. 2014). Thus, it is hugely policy relevant to understand the impacts of well-educated migrants in under-developed areas. This paper quantitatively estimates the effects on educational improvement, short-run versus long-run employment outcomes, and shifts in social values.

Migration		**To**	
		High Human Capital	**Low Human Capital**
From	**High Human Capital**	Rauch (1993), Glaeser and Gottlieb (2009), Moretti (2004a), Moretti (2004b)	Nunn et al. (2014), Hornung (2014), Wantchekon, Klašnja, and Novta (2014), Chen et al. (2018) **This Paper**
	Low Human Capital	Katz, Kling, and Liebman (2001); Gould, Lavy, and Paserman (2004); Kling, Ludwig, and Katz (2005); Kling, Liebman, and Katz(2007); Card and Lewis (2007); Gould, Lavy, and Paserman (2009); Gould, Lavy, and Paserman (2011); Beaman (2012); Abramitzky, Boustan, and Eriksson (2012); Dustmann, Frattini, and Pretson (2013)	Beyond the interests of economists

Second, this paper contributes to the literature on historical origins of trust. Nunn and Wantchekon (2011) document that the slave trade in Africa undermined interpersonal trust; Lowes and Montero (2018) show that the colonial medical campaign of 1921-1956 by the French government reduced peoples' trust in medicine. However, there is a lack of empirical evidence showing any plausible intervention that boosts trust. This paper adds to this strand of literature by suggesting that injection of human capital increases trust among people and trust in modern Western medicine.

Third, this paper provides suggestive evidence that an exogenous migration shock causes the changes in local cultures. Several theoretical and empirical studies explain the cultural persistence and its economic outcomes (Bisin and Verdier 2000, Bisin and Verdier 2001, Guiso, Sapienza, and Zingales 2006, Tabellini 2008, Tabellini 2010, Montgomery 2010, Bisin and Verdier 2011, Giavazzi, Petkov, and Schiantarelli 2014, Alesina and Giuliano 2015, You 2018). Giuliano and Nunn (2017) investigate the origins of cultural changes and argue that an unstable environment undermines the importance of maintaining tradition. This paper proposes migration as another critical factor—an urban-to-rural migration can shift job attitudes and willingness to exert effort in the receiving area.

Additionally, this paper supplements the understanding of Mao's Mass Rustication Movement and its critical impact on development in rural China. Most economic research evaluates the disadvantageous impacts on sent-down youth (Meng and Gregory 2007, Zhang, Liu and Yung 2007, Li, Rosenzweig, and Zhang 2010). This paper focuses on rural residents—the converse side of the issue—and explores the outcomes in educational attainment (Chen et al. 2018), employment, and social values. The quantitative results also complement the narratives documented by historians and political scientists (Burnstein 1977, Liu 2004, Ding and Liu 2009, MacFarquhar and Schoenhals 2009, Walder 2009, Bonnin and Horko 2013, Shi and Tang 2014, Honig and Zhao 2015, Jin and Jin 2015, Walder 2015, Zhang 2015).

The remainder of the paper is organized as follows. Section 2 introduces the background of the Cultural Revolution and Mao's Mass Rustication Movement. Section 3 describes the data and identification strategies. Section 4 documents the improvement in education. Section 5 studies the short-run and long-run employment outcomes and deconstructs effects by occupational category. Section 6 explores the impact of sent-down youth on trust, medical beliefs, subjective well-being, and job attitudes. Section 7 shows suggestive evidence that sent-down youth educated rural students as their teachers. Section 8 concludes.

2. *Background*

Cultural Revolution

The Cultural Revolution, ten years of turmoil from 1966 to 1976, is the critical political background for Mao's Mass Rustication Movement (Walder 2009, Walder 2015). The Cultural Revolution started in Beijing and spread rampantly throughout the entire country. Economic, educational, and political institutions suffered severe setbacks and lapsed into disorder (Bai and Wu 2017). Most schools and universities closed for years, and students participated in "Red Guard" organizations to fight against class enemies, including some teachers and professors. The educational system broke down in the cities. As revolution mania became uncontrollable, the government found it very difficult to persuade people to return to school or go back to work. Although no employment data is available, it is widely acknowledged that most graduates could not find any job position in the cities (Liu and Zhuang 2009). Mao's rustication campaign aims to alleviate the rocketing unemployment and uncontrollable Red Guard violence in urban China.

Mao's Mass Rustication Movement

Mao's Mass Rustication Movement[4] officially started on December 22, 1968.[5] *The People's Daily*, an authorized media outlet of the Chinese Communist Party, headlined Mao's Chairman Command and printed his famous quote, "*Educated Youth should go to the rural areas and be re-educated by peasants. It is necessary!*" The peak year of the rustication movement was 1969. Approximately 2.6 million urban youths, including most graduates in years 1966, 1967, and 1968,[6] were sent down to rural areas. The sent-down movement was designed as a permanent program ex ante. However, many social problems occurred when large crowd of urban youth unexpectedly migrated to rural areas and led to the termination of the movement in October

[4] Also translated as, "Down to the Countryside Movement," "Up to the Mountains and Down to the Villages Campaign," and "Sent-down Youth Movement."
[5] Before 1968, some urban youth voluntarily participated—rather than being sent down—in the rural labor force and helped with agricultural work. But the migration numbers were negligible. The history of urban-to-rural migration dates back to 1955.
[6] Students who graduated in these three years are referred to as the Three Oldest Classes (老三届, "Lao San Jie"). Due to the Cultural Revolution, the Three Oldest Classes did not have opportunities to continue their education or secure jobs. They remained mostly unemployed until Mao's Mass Rustication Movement sent them to rural areas in 1969.

1978. The government allowed urban youth to return home in early 1979. By the end of 1980, most of the sent-down population had returned to urban areas.

MacFarquhar and Schoenhals (2009) estimate that 16,470,000 urban youths were dispatched between 1968 and 1979, which accounts for almost ten percent of the entire urban population. The size of the rustication migration is stunning. Recall that 12.5 million African slaves were shipped to the New World during the entire history of the Trans-Atlantic Slave Trade between 1525 and 1866. Mao's Rustication Movement displaced thirty percent more people within only one decade compared to the slave trade that lasted more than three centuries.

There were three primary motivations for the Chinese government to initiate the rustication campaign. First, Mao discharged the uncontrollable "Red Guards" to rural areas to calm social unrest in the cities. Second, high unemployment posed threats to the command economy. Third, the government pursued communist ideology and aimed to eliminate the "three major inequalities": inequality between urban and rural areas, inequality between agricultural and non-agricultural workers, and inequality between blue-collar and white-collar workers. Sending educated urban residents to the rural areas served as a policy tool to reduce inequality.

Mao's government organized the rustication in a top-down manner: The central government commanded provincial governments to devise annual rustication plans for each county, and provincial governments dispatched urban youth to meet the pre-determined plans. Urban youth were not able to select their preferred destinations;[7] instead, they were directly assigned to particular counties. Each county received a tiny settlement fee associated with each sent-down youth to compensate for their living costs in rural areas. No historical narrative documents that the settlement fee was re-allocated to any other purpose. The settlement fee was also too little to generate any substantial impact on rural counties.

[7] Some migrants voluntarily moved to rural areas because they firmly believed they could make more of a contribution to China in this way. Furthermore, urban residents could find safety from the violence and murder (Walder 2009) occurring in urban areas by moving to rural areas. On the contrary, other migrants were reluctant to participate in the massive migration for reasons such as worse living conditions in rural areas, separation from their families, disruption to their education, communication barriers with local peasants, and tedious agricultural work.

3. *Data and Research Design*

Data

First: Heilongjiang Archival Data. Heilongjiang, a province on the northern border of China, was one of the paramount destinations for sent-down urban youth. I obtain administrative records about the urban youth settlement in 1969 in 61 of 63 rural counties from the Heilongjiang Archive. This paper uses the sent-down youth density—the number of sent-down youth in 1969 divided by the population size in the 1982 Census—as the primary source of variation.[8] In Figure 1, I graph the county-level sent-down youth density in the 1982 Census county map.

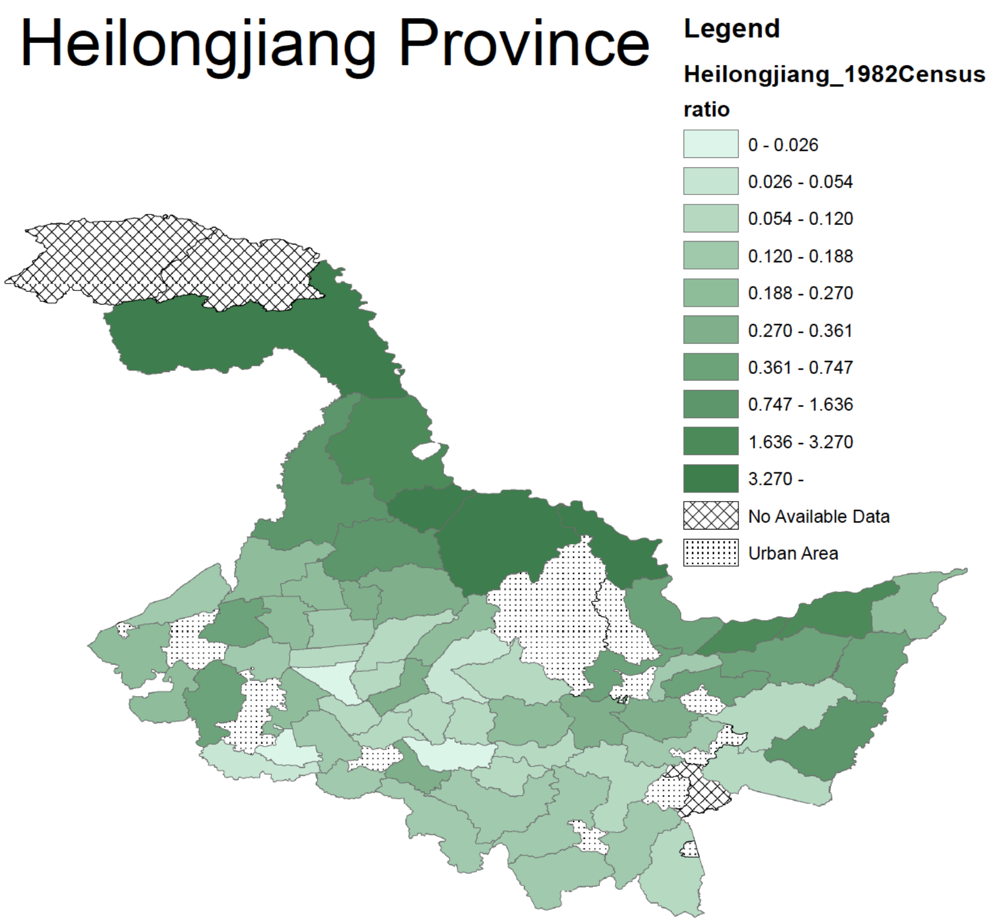


Fig 1. Sent-down youth density distribution in the Heilongjiang County Map. The sent-down youth density is defined as the number of sent-down youth in 1969 over the total county population in 1982.

[8] The measurement is consistent across time because county/city boundaries did not change from 1968 to 1980.

Figure 2 plots the correlation between the sent-down youth density and other county features. The sent-down youth density is 38% (t=3.18)[9] positively correlated with the log geographical distance, 17% (t=1.31) positively correlated with the pre-rustication[10] educational attainment, -50% (t=-4.41) negatively correlated with the county population in 1982, and 44% (t=3.72) positively correlated with the percentage of people working in non-agricultural employment in 1990.

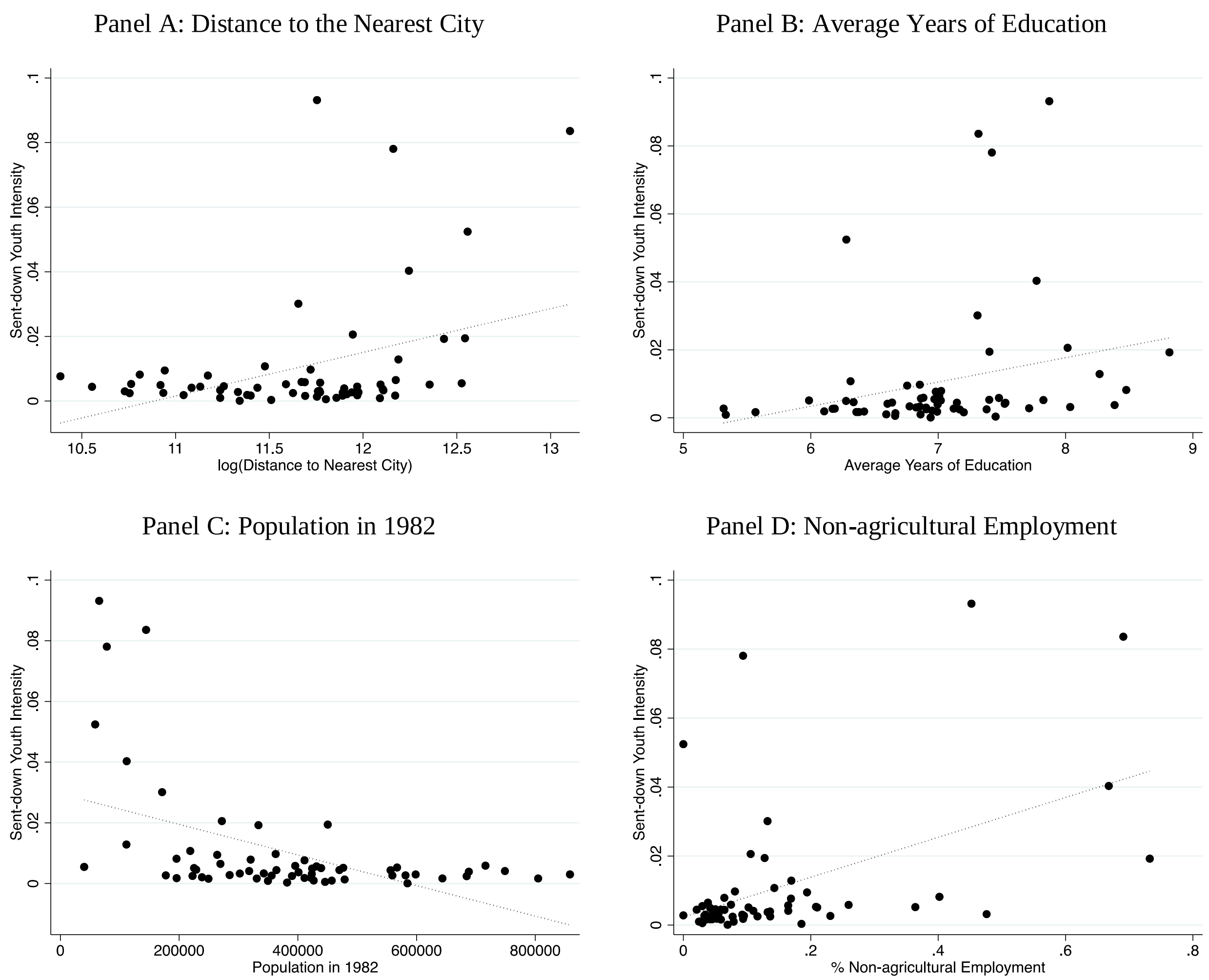


Fig 2. The sent-down youth density and county characteristics. log(Distance to Nearest City) is the log distance to the center of the nearest city. Average years of education and percentage of non-agricultural employment are computed with the population born between 1943 and 1950 — the pre-rustication birth cohorts — in the 1990 Census 1% micro-sample. The population size is from the 1982 Census.

[9] The t-statistics are taken from the coefficients by regressing county features on the sent-down youth density with 61 observations (the 61 rural counties in the sample).

[10] The Pre-rustication educational attainment—years of education—is calculated from the pre-rustication birth cohorts 1943 - 1950.

Second: Chinese Gazetteer Data. I collect county-level unbalanced panel data about the numbers of students, teachers, and schools from the education sections in Chinese county gazetteers.[11] Appendix Table 1 reports the checklist for gazetteer data collection from the 61 rural counties in my sample. Figure 3 reports the average number of students and teachers in primary and secondary schools in the counties with complete annual data from 1958 to 1980. The numbers of teachers in both primary and secondary schools accelerated after 1968, and the number of enrolled students soared correspondingly. Figure 3 visualizes the rapid educational expansion during Mao's Mass Rustication Movement.

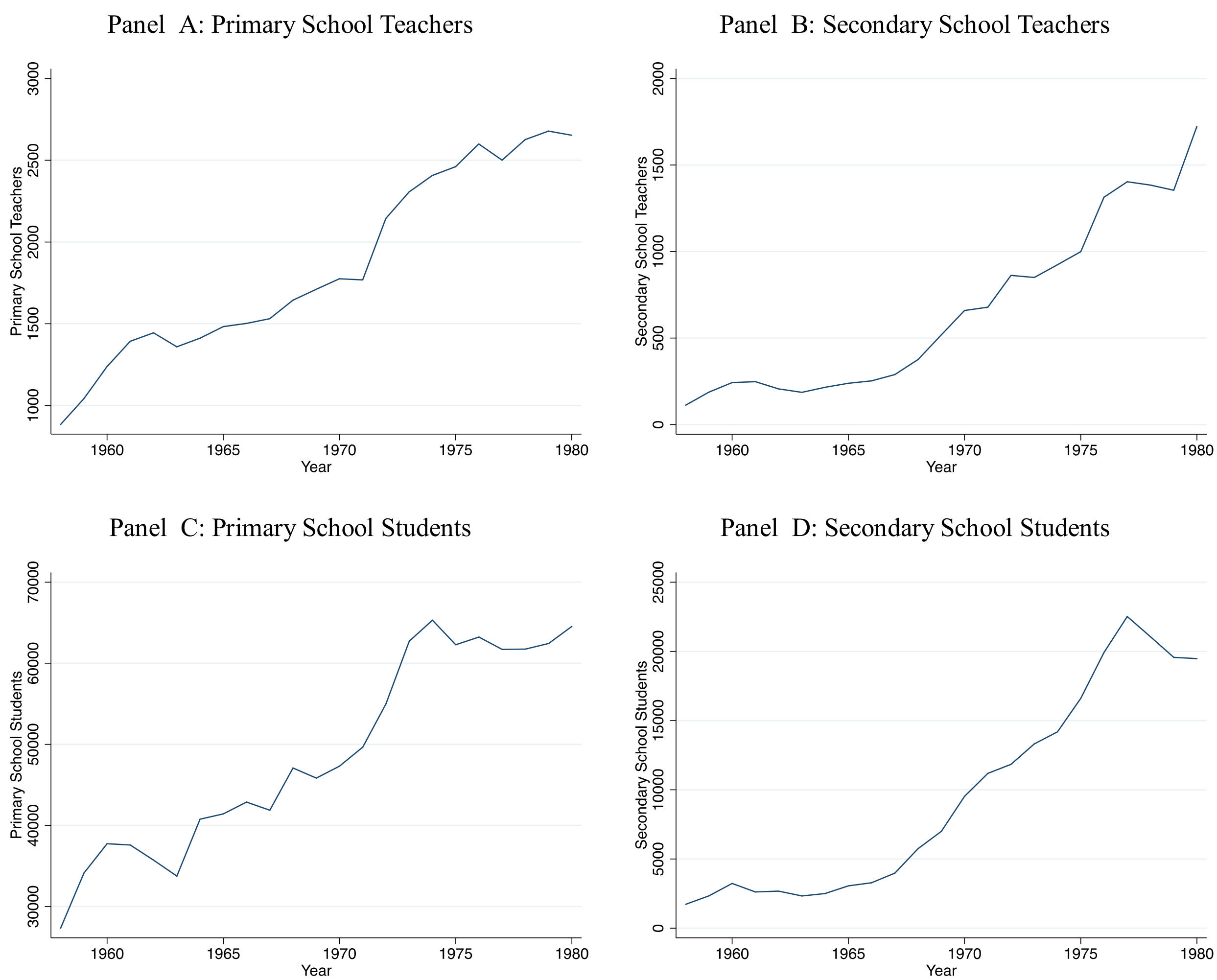


Fig 3. The numbers of students and teachers in 1958-1980. The samples only include Heilongjiang counties with complete annual data from 1958 to 1980. Panel A plots the average number of primary school teachers in 18 counties. Panel B plots the average number of secondary school teachers in 17 counties. Panel C plots the average number of primary school students in 18 counties. Panel D plots the average number of secondary school students in 18 counties.

[11] Chinese county-level historical statistical archives cover the period between 1949 and 1986.

Some gazetteers document the annual number of received sent-down youth. I compile another county-level unbalanced panel data of sent-down youth.[12] In Chinese gazetteers, 15 rural counties recorded the number of sent-down youth arriving in 1969. The numbers from county gazetteers are 95% correlated with the numbers from the Heilongjiang Archive. The correlation remains at 73% after normalization with the 1982 population size. The data consistency validates the reliability of the data sources.

Third: Census of Population Survey (CPS) and China Labor Dynamic Survey (CLDS). The 1990 Census 1% sample, 2000 Census 0.1% sample, and 2005 Population Survey 1% sample[13] provide demographic information, education level, employment status, and occupational choices. The years of education is imputed from the education level.[14] I mainly focus on non-agricultural employment, unemployment, and occupational choices as labor market outcomes. No credible information about income is available in CPS data. Using the Census map, I also compute the geographical distance from the 61 rural counties to the nearest city as an important control variable. Figure 4 shows the average years of education, primary school enrollment rate, middle school enrollment rate, and high school enrollment rate by birth cohort. Panels C and D show that middle and high school enrollments rose sharply for the cohorts born after 1952, peaked for the 1962 birth cohort, and declined after the 1962 birth cohort. The average years of education in Panel A confirms the pattern. These stylized facts are consistent with the findings in Figure 2.

[12] The gazetteer data does not necessarily match the statistics found in the Heilongjiang Archive since the county statistical system is separate from the provincial statistical system.
[13] The 2005 Population Survey randomly selects 1% of the population in China to participate in the survey. My micro sample includes 2,585,481 observations accounting for 0.2% of the Chinese population.
[14] I assume six years in a primary school, three years in a middle school, three years in a high school, and four years to earn a college degree.

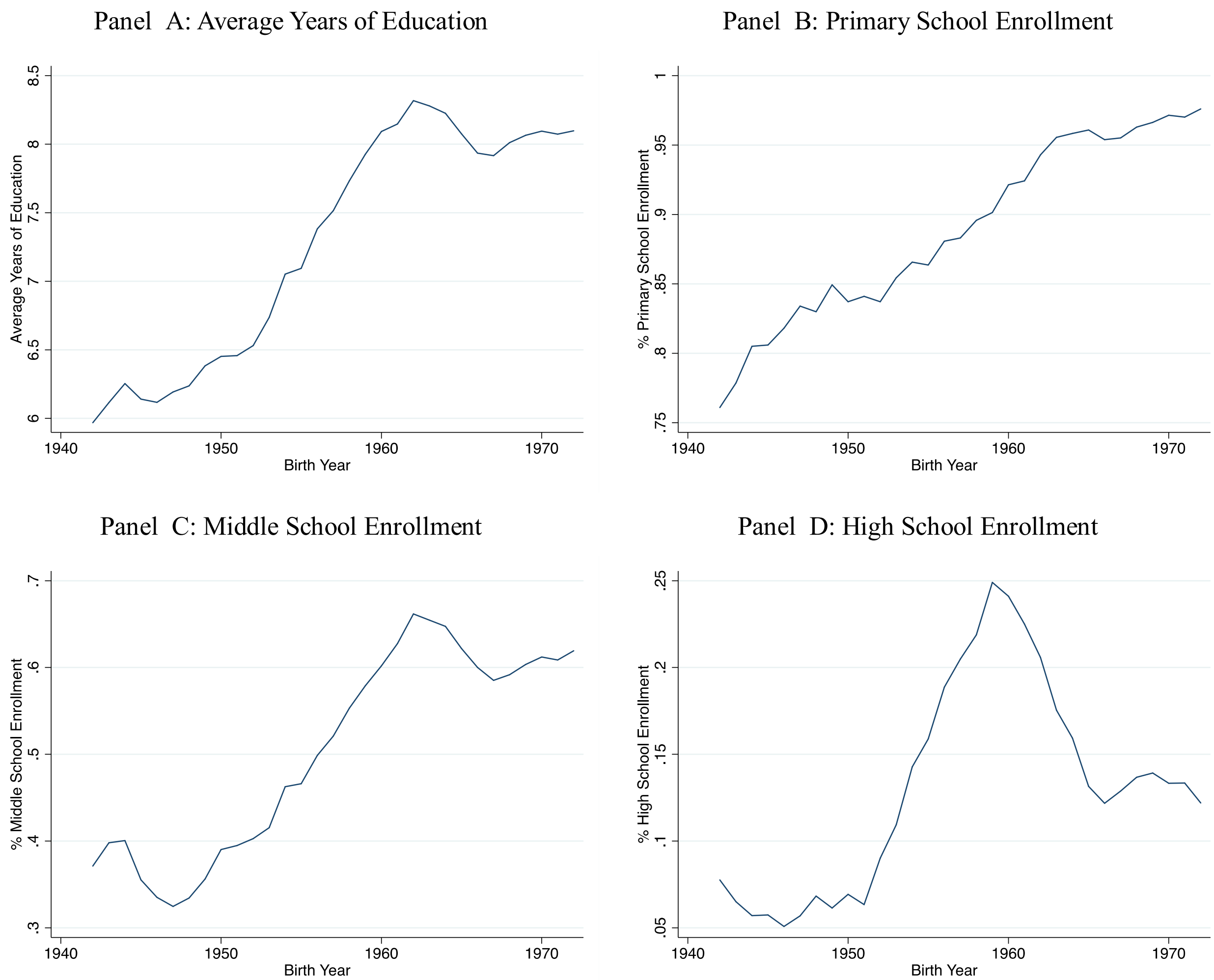


Fig 4. Educational outcomes by birth cohort 1942-1972. Panel A plots the average years of education by birth cohort. Panel B plots the average primary school enrollment rate. Panel C plots the average conditional middle school enrollment rate (defined as the probability of middle school entry conditional on primary school enrollment). Panel D plots average conditional middle school enrollment rate (defined as the probability of high school entry conditional on middle school enrollment).

CLDS provides survey questions about interpersonal trust, medical beliefs, self-reported well-being, job attitudes, and willingness to expend effort. The sample selection criteria are people born between 1942 and 1982 with a birthplace in Heilongjiang Province. CLDS only provides four-digit prefecture identifiers rather than six-digit county codes, thus I conduct my analysis at prefecture level with CLDS data. In total, the sample consists of 339 individuals from 12 prefectures in Heilongjiang province.

Research Design

The empirical analysis consists of four parts of analysis: education, labor market, social values, and channel tests. I first document the impact on educational attainment with both aggregate data and individual data. With the county-level panel data, I compare years of education in the six highest-density counties with the 20 lowest-density counties in the raw data and estimate the causal impact using the synthetic control approach. Then, I impose a linear structure to analyze individual-level samples from CPS by estimating the baseline specification (1):

$$y_{i,j,t} = \beta\ Density_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (1)$$

$y_{i,j,t}$ represents educational outcomes of individual i born in county j[15] and year t including years of education, school enrollment, and dropout rate. Our data sample only includes the people does not include the sent-down youth because most sent-down youths left rural areas by 1980 and all datasets used in this paper were collected from 1990 onward. $Density_j$ is the sent-down youth density defined as a ratio of the sent-down youth number in 1969 divided by the population size in 1982. $Expo_t$ is the dummy for exposure to sent-down youth during schooling. I define two cleavage birth years, 1956 and 1962, for the first watershed event—the onset of the massive sent-down movement in 1968. Since most rural residents only received primary school education, people born before 1956—12 years old in 1968—were unlikely to interact with sent-down youth as students. People born after 1962—6 years old in 1968—inevitably were exposed to sent-down youth in primary school. As a benchmark, $Expo_t$ equals one if born after 1956 (including 1956) and $Expo_t$ equals zero if born before 1956. Alternately, Appendix reports parallel results with $Expo_t$ equals one if born after 1962 (including 1962) and $Expo_t$ equals zero if born before 1956. α_j and δ_t are the county and birth cohort fixed effects. Standard errors are clustered at the county level.

To address the concern of the endogenous sent-down youth assignment, I construct two control variables: the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. For robustness, the specification (2) includes control variables and county-specific linear trends:

---

[15] The only exception is the definition of location in Census 1990. Census 1990 does not collect birth-county because the strict Hukou system bound people to their birthplace back in 1990. Thus we proxy the birth-county with the county they lived in.

$$y_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + Linear\ Trend_j + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (2)$$

I also generalize the specification (1) to the specification (3) to allow for birth-cohort-specific β_t, plot β_t as a function of the birth year, and examine how the function behaves before and after the cleavage birth cohort. The specification (3) addresses two concerns: The first concern is the p-hacking risk by twisting the assumption of birth cohort exposure. Another concern is that the pre-trend may contaminate the estimates. α_j is the county fixed effect, and δ_t is the birth cohort fixed effect.

$$y_{i,j,t} = \sum_t \beta_t Dummy_{i,j,t} * Density_j + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (3)$$

Then, I implement an additional test using the second watershed event—sent-down youth rehabilitation in 1980. The hypothesis is that counties with higher sent-down youth density should have experienced larger decrease in educational attainment when urban youth left rural areas. I substitute $Expo_t$ with a post-rustication dummy $Post_t$ in specifications (1) and (2). $Post_t$ equals one if born after 1975 (including 1975) and $Post_t$ equals zero if born before 1975. People born after 1975—6 years old in 1981—were unlikely to interact with sent-down youth during their school years

$$y_{i,j,t} = \beta Density_j \times Post_t + \gamma\ Controls_j \times Post_t + Linear\ Trend_j + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (4)$$

The second part analyzes the employment outcomes and migration. During the Cultural Revolution, educational quality sharply declined in rampantly expanding rural schools. I document the educational quality decrease with a naïve regression that compares employment outcomes conditional on the same education level. α_j is the county fixed effect.

$$Employment_{i,j,t} = \beta_{edu} Expo_t + Exp_{i,j,t} + Expsq_{i,j,t} + \alpha_j + \varepsilon_{i,j,t} \qquad (5)$$

Then, I evaluate the employment effects of sent-down youth exposure using specifications (1), (2), and (3) with employment variables as the dependent variables. To further separate out the effect driven by educational expansion, specification (6) includes a set of education-level dummies as controls. I estimate specifications with both the 1990 Census and 2000 Census, and show the short-run versus long-run employment effects.

$$Employment_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \xi\ Edu\ Dummies_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (6)$$

Sent-down youth also altered future migration patterns. $Migration_{i,j,t}$ is one if rural youth move in county j between 1985 and 1990; it is zero otherwise.[16]

$$Migration_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (7)$$

In the third part, I estimate the influence on trust, medical beliefs, and values about work. The 2012 CLDS elicits opinions regarding trust, belief in Western medicine and traditional Chinese medicine, and various components of job attitudes. Since the CLDS only provides four-digit prefecture codes,[17] I aggregate sent-down youth density at the prefecture level and estimate a pair of equations in the specification (8). $Controls_j$ represents the average years of education of the pre-rustication birth cohorts 1943-1950.[18] α_j is the prefecture fixed effect, and δ_t is the year fixed effect. Standard errors are clustered at the prefecture level.

$$y_{i,j,t} = \beta_{Expo_t} Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (8)$$

$$y_{i,j,t} = \beta_{Post_t} Density_j \times Post_t + \gamma\ Controls_j \times Post_t + \alpha_j + \delta_t + \varepsilon_{i,j,t} \qquad (8)$$

Consistent with the previous analysis, I estimate β_{Expo} with the exposure dummy $Expo_t$ and β_{Post} with the post-rustication dummy $Post_t$, respectively, and then compare these two coefficients. If urban youth spread some values that differed from local values, I hypothesize that β_{Expo} and β_{Post} should have opposite signs.

In the last part, I suggest that urban youth influenced rural areas as rural teachers with two empirical approaches: the "Pre-Post" comparison (extensive margin) and the panel regression (intensive margin). As the "Pre-Post" comparison, I compute the growth rate of the average educational outcomes from the pre-rustication period, 1958-1967, to the average outcomes in the rustication period, 1968-1977.[19]

$$Growth_j = \frac{\overline{y_{1968-1977}} - \overline{y_{1958-1967}}}{\overline{y_{1958-1967}}}$$

[16] Sent-down youth moved back to cities in 1980. The rehabilitation does not overlap with the migration from 1985-1990.
[17] CLDS does provide six-digit county codes, however they are encoded and do not match with the official Guobiao county code. Thus, I cannot identify birthplace at the county level.
[18] Due to the geographical aggregation, I drop the distance to nearest city in control variables. The average years of education is calculated with Census 2000.

[19] The average outcomes are calculated from all available gazetteer observations.

The educational outcomes include the number of primary/secondary school teachers, number of primary/secondary school students, number of primary/secondary schools, and student-teacher ratio in primary/secondary schools. Then, I run uni-variate regressions of $growth_j$ on the sent-down youth density.

$$Growth_j = \beta\ Density_j + \gamma + \varepsilon_j \qquad (9)$$

Furthermore, I explore the time variation in the sent-down youth panel data. The data sample only consists of counties with more than five observations available during the rustication period, 1968-1978. The specification (10) forecasts educational outcomes with the lagged cumulative sent-down youth density.[20] $\%\ y_{j,t}$ is the per-capita number of teachers, students, and schools (normalized with the 1982 population size). α_j is the county fixed effect, and δ_t is the year fixed effect.

$$\%\ y_{j,t} = \beta Cumulative\ Density_{j,t-1} + \alpha_j + \delta_t + \varepsilon_{j,t} \qquad (10)$$

In addition to all statistical analysis described above, I extensively reviewed the historical narratives to validate these empirical findings. My first source was the education sections of Chinese County Gazetteers, the official statistical archives of China. The second source was memoirs and novels written by sent-down youth. Many memoirs discussed sent-down youths' experiences as rural teachers and their interactions with rural students. The historical narratives corroborate my findings in the data.

4. *Education*

Model-Free Approach

The empirical investigation starts with a model-free comparison using aggregate statistics.[21] Figure 5 plots the average years of education by birth cohort for the six counties with the highest sent-down youth densities and the 20 counties with the lowest densities. The average sent-down

[20] The cumulative sent-down youth density is the total number of sent-down youth arriving before year t divided by the total population in 1982.

[21] Given the unique historical background, I am concerned that the data may be inaccurate (e.g., county statistics do not match with provincial records). Thus, I start with a very conservative approach by ranking counties and comparing the high versus the low values.

youth density is 0.16% in the 20 low-density counties and 6.29% in the six high-density counties. Although the rustication movement occurred in every county, I treat the 20 low-density counties as the quasi-control group since the maximal sent-down youth density is only 0.27%, much smaller than 6.29%.

The high-density counties have slightly higher average educational attainment than the low-density counties in the pre-rustication cohorts, and the education gap significantly widens after birth cohort 1956, persists, and slightly shrinks after birth cohort 1964. The average difference in years of education is 0.34 for cohorts 1943-1955, 1.23 for cohorts 1956-1962, and 1.28 for cohorts 1963-1970.[22] Counties with high densities of sent-down youth experienced a much larger boost in educational attainment. Note that the data sample does not include sent-down youth because over 97% of them had moved back to cities by 1980 and the Census data was collected in 1990. This finding also rules out the possibility of the "catch-up" effect because the government did not intend to send urban youth to low-education counties.

Fig 5. Average years of education in six counties with highest density and 20 counties with lowest sent-down youth density. The solid line is the average of the six counties with the highest sent-down youth density, and the dashed line is the average of the 20 counties with the lowest sent-down youth density. Arrows flag birth cohorts if the difference in years of education is larger than 0.8 years. The left dashed vertical line denotes birth cohort 1956, and the right one denotes birth cohort 1962.

[22] In Appendix Figure 2, I restrict my data sample to residents with agricultural Hukou. It confirms the same pattern with Figure 5, but the education gap is much smaller. The average difference in years of education is 0.24 for cohorts 1943-1955, 0.94 for cohorts 1956-1962, and 0.98 for cohorts 1963-1970.

However, the high-density counties might not be comparable with the low-density counties in other dimensions. To address this concern, I apply the synthetic control method (Abadie, Diamond, and Hainmueller 2010, Abadie, Diamond, and Hainmueller 2015) to construct counterfactuals and compare "apples" with the counterfactual "apples." I match the six highest-density counties with the 30 lowest-density counties using seven variables: log geographical distance to the nearest city; percentage of the people at work; percentage of married population; gender ratio; and years of education for birth cohorts 1943, 1948, and 1953. The pre-trend matching period is 1943-1953 and the treatment year is 1956. Figure 6 shows the average years of education in the six high-density counties and the six synthetic control counties. In birth cohorts 1943-1955, the synthetic control counties match the pre-trend of high-density counties reasonably well. In birth cohorts 1956-1962, the average years of education in the six highest-density counties starts to exceed the synthetic control counties. After birth cohort 1962, the education gap further expands and persists to birth cohort 1970. Figure 6 confirms the stylized fact in the raw data.

Fig 6. The solid line is the average years of education in the six highest-density counties. The dashed line is the average years of education in the six counterfactual counties constructed with the 30 low-density counties using the synthetic control approach. The left dashed vertical line denotes birth cohort 1956 and the right one denotes birth cohort 1962.

The gap between the solid line (high-density counties) and the dashed line (synthetic control counties) is the "high minus low" effect, rather than the treatment effect since all the low-density counties are treated by sent-down youth, although the dose is small. Thus, the education gap in Figure 6 is the lower bound of the real treatment effect. To better understand its magnitude and economic significance, I compare the sent-down movement with the Indonesian Government School Construction Program. Duflo (2001) estimates the education improvement ranges from 0.12 to 0.19 years per new school built per 1000 children. The effect of sent-down youth in the six highest-density counties translates into 5 to 9 new schools constructed per 1000 children.[23]

Then I assess the statistical significance of the estimated effect. Appendix A proposes a theoretical framework for the generalized permutation test that extends the permutation test (Abadie, Diamond, and Hainmueller 2010) to the setting with multiple treatment units (Xu 2017) and access the joint statistical significance. The fundamental intuition is to evaluate the likelihood that the treated units are randomly selected.

The generalized permutation tests follow four steps. First, I choose a pool of control units and construct a synthetic unit for each county with the units in pool.[24] In my context, I pick the 30 low-density counties out of 61 counties as the pool of control units. Second, I calculate the estimated effect as the treated unit minus the synthetic control unit and compute the average estimated effect in the treated units. Third, I randomly select units from the full sample, calculate the average estimated effect, and bootstrap. The last step is to evaluate the average estimated effect of the treated units in the bootstrap distribution from random sampling. The "p-value" is the probability that a random draw has an estimated effect larger than that of the treated units. Appendix A includes more mathematical details.

---

[23] It is worth mentioning that people in rural China in 1982 received, on average, 6.91 years of education, which is much lower than the 8.02 years of education received in Indonesia. This fact undermines the validity of the calculation; however, it demonstrates the large size of the effect.

[24] For a unit in the control unit pool, I use other units in the pool to construct the synthetic control unit.

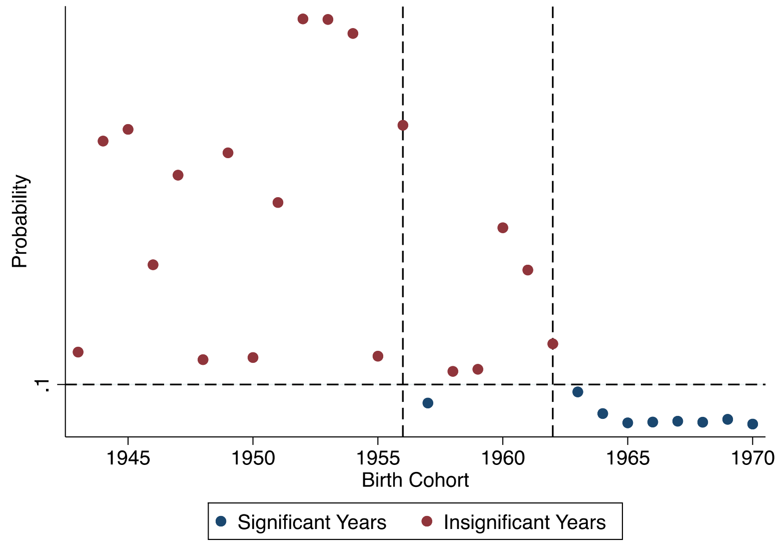


Fig 7. This figure reports the simulated probability of the event that the average estimated effect in the six highest-density counties is lower than the average estimated effect in six randomly chosen counties. The probability is calculated from 10,000 random draws. The left dashed vertical line denotes birth cohort 1956 and the right one denotes birth cohort 1962.

I evaluate the significance with the bootstrap of 10,000 draws and plot the "p-value" by birth cohort. In Figure 7, the blue dots represent the birth cohorts that are statistically significant at the 10% level, while the red dots represent the insignificant cohorts. The results show that all cohorts born after 1962 gain joint statistically significant higher educational attainment in the six highest-density counties.

Urban Youth Arrival in 1968

This section identifies the urban youth impact on education with the first watershed event**:** Sent-down youth arrived in rural areas in late 1968 and early 1969. I estimate a difference-in-difference specification using the variation of sent-down youth density, $Density_j$, the ratio of sent-down youth numbers in 1969 divided by the county population size reported in the 1982 Census. The exposure dummy is zero for cohorts born before 1956 and one for cohorts born after 1956 (including 1956). $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Table 1 reports the baseline specification (1) and the specification (2) with Census 1990 and Census 2000 data.

Table 1: Educational Expansion during Sent-down Movement

	Dependent Variable					
	Census 1990			Census 2000		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: $Edu_{i,j,t}$= Years of Education					
$Density_j \times Expo_t$	11.64**	14.63***	10.10*	5.08**	6.03***	12.96***
	(5.43)	(4.44)	(5.47)	(2.14)	(2.01)	(2.67)
Obs.	126,313	126,313	126,313	10,385	10,385	10,385
	Panel B: $Edu_{i,j,t}$ = Primary School Enrollment					
$Density_j \times Expo_t$	-0.46	0.43**	0.72**	-0.28**	-0.18	-0.10
	(0.33)	(0.20)	(0.31)	(0.13)	(0.13)	(0.18)
Obs.	126,313	126,313	126,313	10,385	10,385	10,385
	Panel C: $Edu_{i,j,t}$ = Middle School Enrollment					
$Density_j \times Expo_t$	2.22**	2.14**	1.42	0.90***	0.96***	1.72***
	(0.97)	(0.96)	(1.01)	(0.14)	(0.18)	(0.28)
Obs.	114,770	114,770	114,770	9,805	9,805	9,805
	Panel D: $Edu_{i,j,t}$ = High School Enrollment					
$Density_j \times Expo_t$	1.89***	1.33*	-1.12	0.47*	0.60*	0.91*
	(0.66)	(0.77)	(0.76)	(0.25)	(0.30)	(0.50)
Obs.	64,969	64,969	64,969	6,146	6,146	6,146
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes
$Linear\ Trend_j$	No	No	Yes	No	No	Yes

Notes: The left three columns report estimates with the 1% Census 1990 sample and the right three columns report estimates with the 0.1% Census 2000 sample. The data samples include birth cohorts 1942 to 1972. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Expo_t$. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) include the county-specific linear trends. $Expo_t$ is the exposure dummy: one if born after 1956 (including 1956); zero if born before 1956. $Density_j$ is the number of sent-down youth in 1969 over the population size in 1982. $Edu_{i,j,t}$ is years of education in Panel A, primary school enrollment in Panel B, conditional middle school enrollment in Panel C, and conditional high school enrollment in Panel D. All specifications include age and county dummies. Robust standard errors are clustered at the county level and reported in parentheses.

$$Edu_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + Linear\ Trend_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Table 1 Panel A shows that residents received more years of education if they were exposed to higher sent-down youth densities in school years. Adding controls and linear trends does not alter the point estimates and statistical significance much in the 1990 sample. In the 2000 sample, controlling for the linear trends strengthens the finding—the coefficient of interest jumps from 5.08 (t=2.37) to 12.96 (t=4.85). All six coefficients are statistically significant at the 10% level. These findings imply that a positive impact on educational attainment is robust to the choices of specification and data sample.

Then, I deconstruct the years of education into primary school enrollment, conditional[25] middle school enrollment, and conditional high school enrollment. As shown in Table 1 Panel B, the impact on primary school enrollment is sensitive to the specification. The coefficients β are both negative in the baseline specification (1), insignificant in the 1990, but significant in the 2000 sample. After including controls and linear trends, the coefficient moves to positive and statistically significant in the 1990 sample, but moves to zero in the 2000 sample. Panel C yields the most robust results. All specifications indicate that a higher sent-down youth density improves conditional middle school enrollment. In Panel D with Census 1990 data, sent-down youth yields sizable and significant positive effects on conditional high school enrollment in the baseline specification. The coefficient slightly decreases after adding controls and drops into the negative range after adding county-specific linear trends. In Panel D with Census 2000 data, the effect is positive but with limited statistical significance. In Appendix Table 2, I use alternative exposure dummy definition (zero for pre-1956 cohorts and one for post-1962 cohorts) to replicate Table 1 and find coefficients are similar to coefficients in Table 1. Our results are robust to the choice of exposure dummy.

I also examine the impact of sent-down youth on dropout rates in the 1990 Census with the same difference-in-difference specifications. Table 2 Panel A reports the results with the benchmark definition of exposure dummy (Before and After 1956), and Panel B reports the results with the alternative definition (Before 1956 and After 1962). Panel A reports the baseline results and Panel B excludes birth cohorts between 1950 and 1956 as the robustness check.

[25] "Conditional" means conditional on enrollment in previous level of education. Conditional middle school enrollment refers to middle school enrollment conditional on primary school enrollment.

Table 2: Dropout Rate during Sent-down Movement

	Dependent Variable							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A : $Expo = 1\ if\ born\ after$ 1956							
	Total Dropouts		Dropouts in Primary School		Dropouts in Middle School		Dropouts in High School	
$Density_j \times Expo_t$	-0.70	-0.62	-1.31**	-1.39**	-0.40	-0.05	1.63*	1.69
	(0.44)	(0.47)	(0.63)	(0.66)	(0.40)	(0.41)	(0.84)	(1.10)
Obs.	86,390	86,390	86,390	86,390	42,053	42,053	7,929	7,929
	Panel B : $Expo = 1\ if\ born\ after$ 1962							
	Total Dropouts		Dropouts in Primary School		Dropouts in Middle School		Dropouts in High School	
$Density_j \times Expo_t$	-0.83*	-0.78	-1.54**	-1.63**	-0.45	-0.06	1.48*	1.66
	(0.45)	(0.48)	(0.72)	(0.75)	(0.42)	(0.44)	(0.88)	(1.20)
Obs.	70,330	70,330	70,330	70,330	33,772	33,772	5,436	5,436
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Controls \times Post$	No	Yes	No	Yes	No	Yes	No	Yes

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: $y_{i,j,t}$ is total dropouts in Columns (1) and (2), dropouts in primary school conditional on primary education enrollment in Columns (3) and (4), dropouts in middle school conditional on middle school enrollment in Columns (5) and (6), dropouts in high school conditional on high school enrollment in Columns (7) and (8). Panel A adopts the definition that $Expo_t = 1$ if born after 1956 (including 1956) and Panel B adopts the definition that $Expo_t = 1$ if born after 1962 (including 1962). All specifications include age and county dummies. Robust standard errors are clustered at the county level and reported in parentheses.

$$y_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Sent-down youth reduce the overall dropout rate, but heterogeneity plays an important role. Sent-down youth substantially lessened the primary school dropout rate. I detect no significant impact on the middle school dropout rate and find suggestive evidence that sent-down youth led to higher high school dropout rates during the high school enrollment spikes as shown in Figure 4 Panel D. The coefficients in Panel B are slightly higher but still very similar to the coefficients reported in Panel A. Thus, the similarity between Panel A and Panel B implies that the results are not sensitive to the definition of the exposure dummy.

Fig 8. This figure plots the cohort-specific coefficients β_t in the specification (3) with years of education as the dependent variable for birth cohorts 1945-1972. The two dotted straight lines fit coefficients in pre-1956 and post-1956 birth cohorts respectively. The dashed lines plot the 90% confidence interval for β_t. The confidence interval is computed with the standard errors clustered at the county level. The left dashed vertical line denotes birth cohort 1956 and the right one denotes birth cohort 1962.

$$Eduyr_{i,j,t} = \sum_t \beta_t Dummy_{i,j,t} * Density_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Then, I estimate the specification (3) to study the dynamics. Figure 8 plots the cohort-specific β_t as a function of birth year. The two vertical dashed lines denote the cleavage birth cohorts 1956 and 1962. For the cohorts born before 1956, the β_ts conform to a dotted flat line in the vicinity of zero. For cohorts born after 1956, β_t jumps to 26 in the birth cohort of 1958 and mostly stays above 20.[26] Figure 8 rules out the concern that the pre-trend drives the identified effect. Notably, β_t spikes in birth cohort 1964 and gradually declines afterward. This pattern is

[26] Appendix Figure 3 replicates Figure 8 using population with agricultural Hukou. For the cohorts born before 1956, the β_ts conform to a dotted upward sloping line in the vicinity of zero. For cohorts born after 1956, β_t jumps to 20 in the birth cohort of 1958 and mostly stays above 10.

consistent with the shrinking education gap in the post-1964 cohorts, as shown in Figure 6. According to the visualization, a density of 1% more sent-down youth translates into 0.2 extra years of education for cohorts born after 1957. Recall that the average sent-down youth density in the top six counties is 6.29%. The baseline estimation predicts roughly 1.25 years of educational improvement in the cohorts who were exposed to sent-down youth in schooling years. The magnitude is also comparable with the education gap documented in Figure 5.

Rehabilitation in 1980

This section documents the reduction in education attainment after the second watershed event: Sent-down youth were rehabilitated back to cities in 1980. The logic is simple: Counties with higher inflows of urban youth experienced more substantial outflows at the end of the movement. Thus, a county with a high sent-down youth density should suffer more setbacks in education after sent-down youth left the county. The termination date of rustication is fuzzier than the start of the rustication campaign. "Returning to cities" was first officially allowed in 1979. However, some sent-down youths had returned to cities before 1979 due to disease, employment in urban factories, family reasons, etc.

I compare cohorts who received education after sent-down youth returned to cities with the more senior cohorts. The post-rustication dummy $Post_t$ equals one if born after 1975 (including 1975) and $Post_t$ equals zero if born before 1975. I estimate the specification (4) with the data of birth cohorts 1962-1982 in the 2000 Census, including eight post-rustication cohorts and 13 rustication cohorts.

Table 3: Educational Reduction in Post-Rustication

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Years of Education			Panel B: Primary School Enrollment		
$Density_j \times Post_t$	-3.10*	-2.09	-7.34***	0.0001	0.02	0.19
	(1.61)	(1.68)	(2.57)	(0.07)	(0.08)	(0.19)
Obs.	8,614	8,614	8,614	8,614	8,614	8,614
	Panel C: Middle School Enrollment			Panel D: High School Enrollment		
$Density_j \times Post_t$	-0.53**	-0.34	-0.57	0.11	0.03	-0.70
	(0.23)	(0.27)	(0.85)	(0.53)	(0.52)	(0.57)
Obs.	8,483	8,483	8,483	6,322	6,322	6,322
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes
$Linear\ Trend_j$	No	No	Yes	No	No	Yes

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: The data sample is the birth cohorts from 1962 to 1982 in Census 2000. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. $Post_t$ is the exposure dummy: one if born after 1975 (including 1975); zero if born before 1975. $Density_j$ is the number of sent-down youth in 1969 over the population size in 1982. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Post_t$. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) include the county-specific linear trends. All specifications include age and county dummies. Standard errors are clustered at the county level and reported in parentheses.

$$Edu_{i,j,t} = \beta Density_j \times Post_t + \gamma\ Controls_j \times Post_t + Linear\ Trend_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Table 3 Panel A, Panel C, and Panel D show that residents in counties with high sent-down youth density experienced lower years of education, conditional middle school enrollment, and conditional high school enrollment. An interesting pattern is that coefficient β decreases after controlling county-specific linear trends: from -2.09 to -7.34 in Panel A, from -0.34 to -0.57 in Panel C, and from -0.03 to -0.70 in Panel D. This implies the downward deviation from the linear trend is larger than the drop in level. These findings are consistent with the hypothesis that the arrival of urban youth increased educational attainment and "returning to cities" decreased educational attainment. The principal educational reduction shown in Panel A comes from the decrease in middle and high school enrollment. As shown in Panel B, the effect of the sent-down youth rehabilitation on primary school enrollment is insignificant. The coefficients are also quite small in magnitude (e.g., in columns (3) and (6), coefficient β is only 0.19 in Panel B, but -0.57 in Panel C, and -0.70 in Panel D).

Persistency

Another natural question concerns persistency: Did Mao's Mass Rustication Movement lead to a persistent divergence in educational attainment between counties with high and low sent-down youth densities, or just a temporary divergence with a convergence at the end of the rustication? The answer is: The effect on education remained robust and statistically significant until birth cohort 1970. However, the effect gradually diminished for birth cohorts after 1970, and almost disappeared until the birth cohort born in 1980. These results deliver an important policy implication that a temporary injection of human capital to under-developed areas may not be sufficient to improve education in the long run.

$$Edu2005_{j,t-2\rightarrow t+2} = \alpha_t\ Edu1990_j + \beta_t\ Density_j + \delta_t + \varepsilon_{j,t}$$

I run a sequence of county-level forecasting regressions by predicting future average years of education with the sent-down youth density and average years of education of the pre-rustication cohorts. Pre-rustication education, $Edu1990_j$, is defined as the average years of education of birth cohorts 1946 to 1956 in county *j* using the 1990 Census. The dependent variable $Edu2005_{j,t-2\rightarrow t+2}$ is computed with the 2005 1% Population Survey. To compensate for the small sample size of the 2005 Survey, I proxy the educational attainment of birth cohort *t* in county *j*, $Edu2005_{j,t-2\rightarrow t+2}$, using all observations of five consecutive birth cohorts from *t*-2 to *t*+2 in county *j*. Each regression pins down β_t for each birth cohort from 1944 to 1985.[27] Each β_t captures the predictive power of the sent-down youth density in the educational attainment of birth cohort *t*.

[27] My sample only includes people whose age was above 18. Thus, birth cohort 1985 is the youngest cohort in the study because $Edu2005$ intakes the educational attainment of birth cohort 1987 whose ages were 18 in 2005.

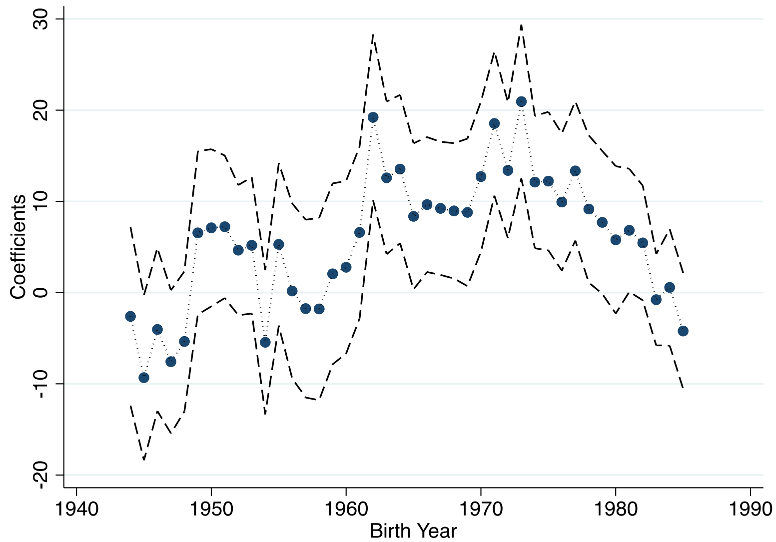


Fig 9. This figure plots coefficients β_t of the county-level forecast regressions as a function of birth year t. $Edu2005_{j,t-2\rightarrow t+2}$ is the average years of education of birth cohorts t-2 to t+2 (as the proxy for birth cohort t) in county j from Census 2005. $Edu1990_j$ is the average years of education of birth cohorts 1946 to 1956 in county j from Census 1990. $Density_j$ is the number of sent-down youth in 1969 over the population size in 1982. Each dot represents β_t estimated from each forecast regression. The dotted line shows the dynamics by connecting all coefficients. The dashed lines plot the 90% confidence interval.

$$Edu2005_{j,t-2\rightarrow t+2} = \alpha_t Edu1990_j + \beta_t Density_j + \delta_t + \varepsilon_{j,t}$$

Figure 9 plots the β_t as a function of the birth year t. Coefficients conform to a hump-shaped curve. The coefficients initially fluctuate around zero, jump up to the positive range (significant at the 90% confidence level) in birth year 1962, persist until birth year 1973, and then gradually decline over time. The predictability completely fades away after birth cohort 1982. The externality of sent-down youth in education continues in the birth cohorts of the 1970s but disappears in birth cohorts of the 1980s.

The main takeaway is that sent-down youth did not generate a permanent educational advantage among rural populations in the long run. There are many policy practices that advocate bringing human capital to low-income areas (e.g., Teach for America in the US and the University Student Village Official Program in China[28]). My findings cast doubts on these practices by showing that an injection of high human capital may generate only temporary improvements but fail to create persistent educational improvement.

[28] This program encourages well-educated university graduates to take non-governmental positions as village officials in rural areas. The program is entirely volunteer-based.

5. *Labor Market*

Analyzing the impact of urban youth on employment is more complex than assessing educational attainment. In addition to individual ability, employment also depends on the abilities of others in the same labor market and on labor-market demand. Time-varying market demand may lead to different employment outcomes. For example: In a technology-advanced economy, people have more occupational choices (e.g., to work in a factory) that may not be feasible in an agricultural economy. In this context, I study the employment outcomes in both 1990 and 2000—referred to as short-run and long-run effects—and find interesting dynamics.

Educational Quality

I commence with documenting weaker employment competence resulting from being educated during the rustication period. Almost all education chapters of county gazetteers in Heilongjiang Province document the severe drop in educational quality caused by the Cultural Revolution.[29] Classes in political indoctrination and agricultural practices replaced many basic courses (e.g., mathematics and Chinese). The Hailin County Gazetteer states: "In the Cultural Revolution years, secondary education quality was very low due to the rapid expansion in secondary education. The secondary school expansion also imposed an extra economic burden on rural peasants." This subsection confirms the decrease in educational quality manifested by poorer employment.

Table 4 presents OLS estimates of the specification (5) to compare pre-rustication and rustication cohorts in non-agricultural employment (Panel A) and unemployment (Panel B) in 1990, conditional on the same educational level. I add the post-graduation years, $Exp_{i,j,t}$, and the

[29] Almost every gazetteer in Heilongjiang mentions the sharp decline in educational quality. The Fangzheng County Gazetteer states: "In the early stage of the Cultural Revolution (1966-1968), students were immensely affected by the 'Gang of Four,' 'No School, Go Revolution,' 'Rebellion Justified, Revolution Legitimated.' Students did not go to school, but rebelled, criticized, and struggled with teachers and school heads as the 'Capitalist enemies.' In 1968, students returned to school. Propaganda teams and peasant representatives occupied schools and replaced basic curricula with the 'Great Struggle' activities. After returning to schools, students not only learned from teachers, but also from factory workers, peasants, and soldiers. Education only contained limited basic courses. Education quality tremendously declined, and the actual abilities of graduates were lower than before."

squared post-graduation years, $Expsq_{i,j,t}$, in the specification to control for the variation in employment driven by work experience.

Table 4: Educational Quality during Cultural Revolution

	Panel A: Non-Agricultural Employment					
	OLS			FE Logit		
$Expo_t \mid Primary$	0.01			0.05		
	(0.01)			(0.08)		
$Expo_t \mid Middle$		-0.05***			-0.04	
		(0.01)			(0.05)	
$Expo_t \mid High$			-0.02			-0.27**
			(0.02)			(0.081)
Obs.	36,924	39,502	13,203	36,924	39,502	13,203
	Panel B: Unemployment					
	OLS			FE Logit		
$Expo_t \mid Primary$	0.08***			0.58***		
	(0.01)			(0.09)		
$Expo_t \mid Middle$		0.04***			0.07	
		(0.01)			(0.11)	
$Expo_t \mid High$			-0.01			-0.71***
			(0.01)			(0.18)
Obs.	47,512	46,719	14,641	47,512	46,669	14,632

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: $y_{i,j,t}$ is the non-agricultural employment dummy in Panel A and unemployment dummy (defined as not working and not in school) in Panel B. The data samples include only birth cohorts 1942 - 1972. I split the full sample into three groups by education enrollment level: primary school enrollment, middle school enrollment, and high school enrollment. β_{edu} is the coefficient estimated for each education level. $Expo_t$ is the exposure dummy: one if born after 1956; zero if born before 1956. $Exp_{i,j,t}$ is the imputed post-education years, $Expsq_{i,j,t}$ is the squared $Exp_{i,j,t}$. All specifications include county fixed effects α_j. The left three columns report estimates with the OLS model and the right three columns report estimates with the fixed-effect Logit model. Robust standard errors are clustered at the county level and reported in parentheses.

$$y_{i,j,t} = \beta_{edu} Expo_t + Exp_{i,j,t} + Expsq_{i,j,t} + \alpha_j + \varepsilon_{i,j,t}$$

The specification is estimated within three subgroups by education level—people enrolled in primary school, middle school, and high school but without further education. In Table 4, the left three columns report estimates with an OLS model and the right three columns report estimates with a (county) fixed-effect Logit model. Panel A shows that acquiring middle school and high school education during the rustication period lowers the probability of non-agricultural employment by 4% and 4.5% respectively, and Panel B shows that rustication-period primary school and middle school education increases the unemployment probability in 1990 by 7.4% and 4.2% respectively. The magnitude of coefficients is sufficiently large to conclude that educational quality substantially regressed and led to poorer employment records in 1990.

Arrival of Urban Youth in 1968

This subsection studies the impact of sent-down youth on employment outcomes in 1990 and 2000. For each employment outcome, I estimate three specifications: the baseline specification (1), the specification (1) with control variables, and the specification (6). The main motivation of specification (6) is to tease out the effect of the educational expansion identified in Section 4. To explore heterogeneity, I also limit the data sample to residents with agricultural Hukou in Census 1990 and Census 2000 and report the results in Appendix Table 3 Panels A and B.[30]

[30] A caveat worth noting is that some people with an agricultural Hukou in 1990 may have transferred to the non-agricultural Hukou during the decade of rapid urbanization. In the 1990 Census sample, 77% of the population held the agricultural Hukou, and this number declined to 70% in the 2000 Census sample. For example, specialists with agricultural Hukou in 1990 could move to cities and hold non-agricultural Hukou in 2000. Due to data limitation, I am not able to track the change in Hukou type.

Table 5: Dynamic Employment Effects of Sent-down Youth Exposure

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Non-agricultural Employment					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-0.59*	-0.34	-1.36*	0.84***	1.15***	0.47**
	(0.32)	(0.43)	(0.72)	(0.14)	(0.23)	(0.22)
Obs.	99,356	99,356	99,356	8,447	8,447	8,447
	Panel B: Unemployment					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	0.22	-0.36	0.01	-0.78***	-0.69***	-0.68***
	(0.60)	(0.47)	(0.46)	(0.21)	(0.25)	(0.25)
Obs.	123,636	123,636	123,636	10,385	10,385	10,385
Age and County Dummies	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes
Education Dummies	No	No	Yes	No	No	Yes

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: The data samples include birth cohorts from 1942 to 1972 in Censuses 1990 and 2000. All specifications include age and county dummies. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Expo_t$. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) control for individual education-level dummies. The dependent variable is a dummy for non-agricultural employment in Panel A, and unemployment in Panel B. In each panel, the left three columns reports estimation with Census 1990 and the right three columns reports estimation with Census 2000. Robust standard errors are clustered at the county level and reported in parentheses.

$$Employment_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \xi\ Education\ Dummies_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Table 5 columns (1) and (4) report the baseline specification (1) with the non-agricultural employment dummy in Panel A and the unemployment dummy in Panel B as the dependent variables. Columns (2) and (5) add the pre-rustication education and geographical distance controls to the specification. Panel A shows that the aggregate effect in non-agricultural employment tends to be negative in 1990. A 1% increase in sent-down youth density reduced non-agricultural employment by 0.34% to 0.59% in the full sample, and 0.49% to 0.69% in the population with agricultural Hukou.[31] Figure 10 plots the β_t estimated from the specification (3) with non-agricultural employment as the dependent variable using the 1990 Census sample. The β_t for birth cohorts 1945-1955 form a slightly upward sloping line, ruling out the pre-trend concern. β_t remains steady in the birth cohorts 1956-1962, and then a declining trend appears in

[31] Appendix Table 3 Panels A and B report the parallel results in the population with agricultural Hukou.

cohorts born after 1962.[32] The cohort-specific β_t plot confirms the negative effect in Table 5 Panel A. Conversely, a decade later, the direction of the impact on employment flipped the sign. A 1% increase in sent-down youth density encouraged non-agricultural employment by 0.84% to 1.15% in the full sample, and 0.65% to 0.79% in population with agricultural Hukou.

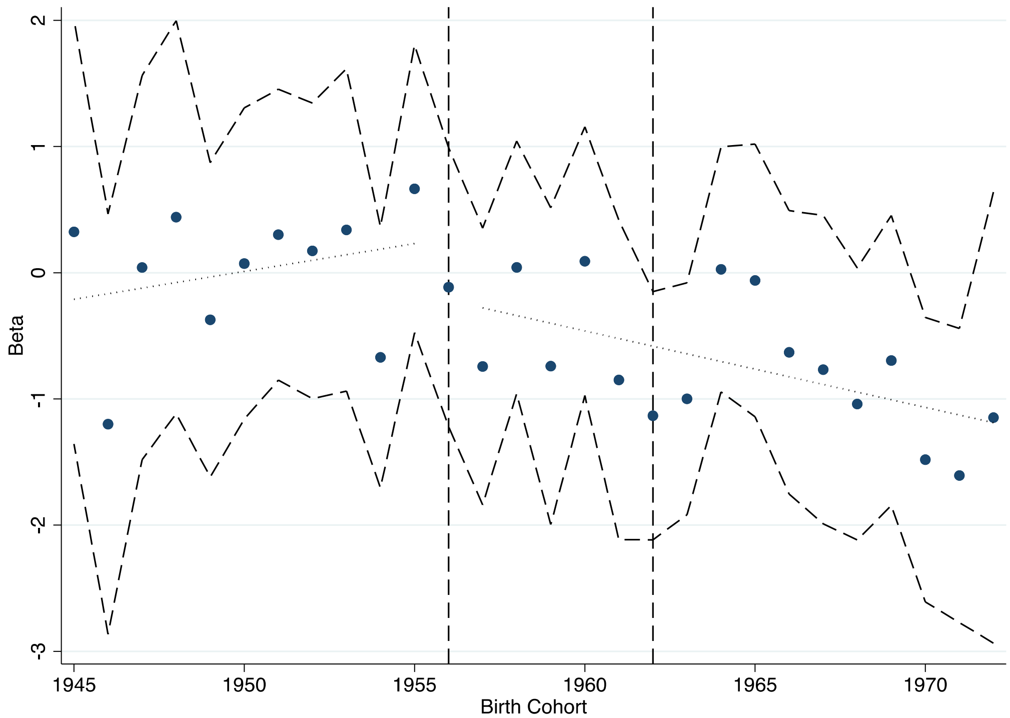


Fig 10. This figure plots the cohort-specific coefficients β_t with non-agricultural employment as the dependent variable for birth cohorts 1945-1972. The two dotted straight lines fit coefficients in pre-1956 and post-1956 birth cohorts respectively. The dashed lines plot the 90% confidence interval for β_t. The confidence interval is computed with the standard errors clustered at the county level. The left dashed vertical line denotes birth cohort 1956, and the right one denotes birth cohort 1962.

$$Industry_{i,j,t} = \sum_t \beta_t Dummy_{i,j,t} * Density_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Panel B shows that sent-down youth had no significant impact on unemployment in 1990. A decade later, the unemployment rate reduced by 0.69% - 0.78% if the sent-down youth density increased by 1%. Reduction of unemployment was concentrated mainly in the population with non-agricultural Hukou registration in 2000.

[32] Appendix Figure 4 replicates the same practice using the subsample with agricultural Hukou. β_t fluctuates around zero and starts to decline after birth cohort 1956. The negative impact on employment also holds in the population with agricultural Hukou.

Table 5 columns (3) and (6) report the estimates of the specification (6) including education-level dummies. From 1990 to 2000, the coefficients jump from -1.36 (*t*=-1.89) to 0.47 (*t*=2.08) in Panel A, and drop from 0.01 (*t*=0.02) to -0.68 (*t*=-2.70) in Panel B. It implies that higher exposure to sent-down youth during school years generated more employment benefits—more non-agricultural employment and less unemployment—in 2000 than in 1990 after controlling for education-level dummies. In the short term, exposure to sent-down youth reduced non-agricultural employment in 1990. However, exposure to sent-down youth started to create a positive premium in non-agricultural employment when the Chinese economy was more industrialized and urbanized in 2000.

In Table 6, I divide non-agricultural employment into five occupational categories: factory workers, service workers, specialists, government officials, and administrative staff. Focusing on columns (1), (2), (4) and (5), the non-agricultural reduction in the 1990 sample is mainly driven by employment of factory workers, service workers, and government officials. In the baseline specification (1), a density of 1% more sent-down youth corresponds to 1.12% (*t*=-2.50) fewer factory workers, 0.97% (*t*=-2.48) fewer service workers, and 2.65% (*t*=-2.58) fewer government officials. In the 2000 Census, sent-down youth increased employment in all five non-agricultural occupation categories: 1% significance for factory workers, service workers, and specialists; and 5% significance for administrative staff.[33]

[33] Appendix In the 2000 Census rural sample, the coefficients are much smaller and less statistically significant. One plausible explanation is that rural residents with non-agricultural jobs moved to urban areas and earned the non-agricultural Hukou. Thus, I obtain a weaker effect if I limit the sample to the rural sample of the 2000 Census.

Table 6: Sent-down Youth Exposure and Occupation Choice

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Factory Workers					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-1.12**	-0.98**	-1.76***	1.08***	1.03***	0.74**
	(0.45)	(0.48)	(0.53)	(0.29)	(0.37)	(0.32)
Obs.	83,192	83,192	83,192	6,716	6,716	6,716
	Panel B: Service Workers					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-0.97**	-0.89**	-1.55***	1.01***	1.19***	0.68***
	(0.39)	(0.37)	(0.34)	(0.14)	(0.22)	(0.22)
Obs.	76,133	76,133	76,133	6,526	6,526	6,526
	Panel C: Specialists					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-0.11	0.13	-0.79**	1.44***	1.51***	0.48***
	(0.29)	(0.26)	(0.31)	(0.13)	(0.17)	(0.13)
Obs.	75,946	75,946	75,946	6,298	6,298	6,298
	Panel D: Government Officials					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-2.65**	-2.06**	-2.22***	0.15	0.22	0.09
	(1.03)	(0.79)	(0.67)	(0.15)	(0.18)	(0.09)
Obs.	71,623	71,623	71,623	6,026	6,026	6,026
	Panel E: Administrative Staff					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-0.10	0.05	-0.29	0.57**	0.47**	0.07
	(0.21)	(0.23)	(0.26)	(0.22)	(0.22)	(0.13)
Obs.	71,482	71,482	71,482	6,093	6,093	6,093
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes
$Education\ Dummies$	No	No	Yes	No	No	Yes

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: The data samples include birth cohorts from 1942 to 1972 in Censuses 1990 and 2000. All specifications include age and county dummies. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Expo_t$. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) control for individual education-level dummies. The dependent variable is dummy for factory workers in Panel A, service workers in Panel B, specialists in Panel C, government officials in Panel D, and administrative staff in Panel E. In each panel, the left three columns reports estimation with Census 1990 and the right three columns reports estimation with Census 2000. Robust standard errors are clustered at the county level and reported in parentheses.

$$Occupation_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \xi\ Education\ Dummies_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Column (6) yields another interesting finding by comparing across occupations. After controlling for educational expansion, a 1% sent-down density translates into increases of 0.74% (t=2.33) in factory workers, 0.68% (t=3.06) in service workers, 0.48% (t=3.57) in specialists, 0.09% (t=1.02) in government officials, and 0.07% (t=0.56) in administrative staff. In 2000, sent-down youth exposure was critical for productive occupations (factory workers, service workers, and specialists), but less important for bureaucratic occupations (government officials and administrative staff) in both coefficient magnitude and statistical significance.

The dynamic effect implies that human capital can generate different impacts on employment in the short and long horizons. In 2000, exposure to urban youth empowered rural people to benefit more from economic development; in particular, they were more likely to fill productive positions in society. However, in 1990, the benefits of urban knowledge did not yet exist. Thus, the evaluation of educational intervention in the short run may potentially understate total economic benefits.

Rehabilitation in 1980

This subsection documents that people in high-density counties suffer from poorer labor market performance in 2000. Table 7 corresponds to the right three columns in Tables 5 and 6 with the 2000 Census. Panels A-G report regression results with dependent variables as non-agricultural employment, unemployment, factory workers, service workers, specialists, government officials, and administrative staff. Panel A and Panel B show that birth cohorts after 1975 in high-density counties were less likely to gain non-agricultural employment and more likely to be unemployed. The unemployment increase is statistically significant: After sent-down youth returned to cities, a density of 1% more sent-down youth corresponds to 0.96% (t=2.62) more unemployment in the baseline specification, and the number increases to 1.18% (t=4.38) after adding controls and educational-level dummies.

Table 7: Employment in Post-Rustication Period

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Non-agricultural Employment			Panel B: Unemployment		
$Density_j \times Post_t$	-1.01	-0.98	-0.38	0.96**	1.18***	1.18***
	(0.84)	(0.83)	(0.58)	(0.37)	(0.32)	(0.27)
Obs.	6,909	6,909	6,909	8,614	8,614	8,614
	Panel C: Factory Workers			Panel D: Service Workers		
$Density_j \times Post_t$	-0.50	-0.57	-0.27	-0.75	-0.62	-0.30
	(0.65)	(0.66)	(0.64)	(0.55)	(0.58)	(0.60)
Obs.	5,685	5,685	5,685	5,502	6,526	6,526
	Panel E: Specialists			Panel F: Government Officials		
$Density_j \times Post_t$	-2.00***	-2.05***	-0.87***	0.17	0.24	0.24
	(0.33)	(0.35)	(0.30)	(0.35)	(0.33)	(0.44)
Obs.	5,233	5,233	5,233	4,962	4,962	4,962
	Panel G: Administrative Staff					
$Density_j \times Post_t$	-0.27	-0.24	0.06			
	(0.44)	(0.42)	(0.57)			
Obs.	5,087	5,087	5,087			
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes
$Education\ Dummies$	No	No	Yes	No	No	Yes

Notes: The data samples include birth cohorts from 1962 to 1982 in Census 2000. $Post_t$ is the post-rustication dummy: one if born after 1975 (including 1975); zero if born before 1975. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Post_t$. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) control for individual education-level dummies. The dependent variable $Employment_{i,j,t}$ is a dummy for non-agricultural employment in Panel A, unemployment in Panel B, factory workers in Panel C, service workers in Panel D, specialists in Panel E, government officials in Panel F, and administrative staff in Panel G. Robust standard errors are clustered at the county level and reported in parentheses.

$$Employment_{i,j,t} = \beta Density_j \times Post_t + \gamma\ Controls_j \times Post_t + \xi\ Education\ Dummies_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Panels C through G report the breakdown into the five occupation categories. Only coefficients for the specialist group are negative with statistical significance. Coefficients for factory workers and service workers are also negative but with limited statistical power. Recall that people with urban exposure did significantly better in these three productive job categories as shown in the previous subsection. The negative coefficients in Panels C-E confirm that urban youth rehabilitation reduced the employment advantages in these three occupation categories. In Panels F and G, I find no significant effect in bureaucratic jobs. This finding echoes the insignificant results in Table 6 Panel D and Panel E.

Migration Flow Composition

This subsection documents how sent-down youth affected future migration flows after the rustication campaign. In the period 1985-1990, counties with higher densities of sent-down youth tended to attract more young and educated migrants—both within provinces and from other provinces—who had previous exposure to sent-down youth. This fact implies that urban youth also empowered the county to attract more young people from other places. Table 8 reports the coefficients of interest in the specification (7) estimated with both OLS and Logit models.

Table 8: Future Migration Inflow

	Dependent Variable			
	Panel A: All Migration			
	OLS		FE Logit	
$Density_j \times Expo_t$	2.27**	1.79**	10.28	6.81
	(0.86)	(0.76)	(7.25)	(6.46)
Obs.	130,179	130,179	128,659	128,659
	Panel B: Migration from Other Provinces			
	OLS		FE Logit	
$Density_j \times Expo_t$	0.88**	0.67**	14.82*	10.83
	(0.38)	(0.33)	(8.44)	(6.91)
Obs.	127,379	127,379	113,268	113,268
	Panel C: Migration from Other Heilongjiang Counties			
	OLS		FE Logit	
$Density_j \times Expo_t$	1.82**	1.50**	9.37	5.76
	(0.70)	(0.65)	(7.36)	(7.13)
Obs.	129,113	129,113	127,593	127,593
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	No	Yes

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: $Migration_{i,j,t}$ is a dummy for migrants who arrived in county j between 1985 and 1990. The dependent variable indicates all migrants born outside county j in Panel A, migrants from other provinces in Panel B, and migrants from other Heilongjiang counties in Panel C. $Controls$ includes the log geographical distance to the nearest city, GDP, employment percentage, and agricultural population percentage in 1982. $Expo_t$ is the exposure dummy: one if born after 1956; one if born before 1956. All specifications include age and county dummies. The left two columns report OLS results and the right two columns report results estimated with a fixed-effect Logit model. Robust standard errors are clustered at the county level and reported in parentheses.

$$Migration_{i,j,t} = \beta\ Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

The dependent variable, $Migration_{i,j,t}$, indicates all types of migrants who moved to county j during 1985-1990 in Panel A, migrants from other provinces in Panel B, and migrants from other Heilongjiang counties in Panel C. In the OLS linear probability model, all coefficients are significant at the 5% level. A 1% increase in sent-down youth density corresponds to approximately 2% more young migrants in the county. In that 2% of young migrants, 30% of them moved to the county from another province, and 70% came from other Heilongjiang counties. Estimates with the county fixed-effect Logit model confirm the positive effect on migration inflow; however, the statistical power is limited.

The data is not sufficient to identify why sent-down youth affected migration choices, but I propose two plausible channels. First, the sent-down youth brought opportunities and values[34] that were more attractive to young and educated workers. Another possibility is that sent-down youth revealed more information about their sent-down counties and built more networks with other areas (Kinnan, Wang, and Wang 2018).

6. *Trust, Well-being, and Job Attitudes*

Sent-down youth also could have brought urban mindsets to the rural areas and even altered cultures there. Interaction with urban youth might have nurtured new cultural attitudes among rural people that persist in the long run. This section evaluates the impact of exposure to urban youth during school years on social values: trust, well-being, and job attitudes. The China Labor Dynamics Survey 2012 (CLDS) provides various relevant survey questions. In the specification (8), consistent with sample selection criteria in previous sections, I estimate β_{Expo} with birth cohorts 1942-1972 (183 respondents) and β_{Post} with birth cohorts 1962-1982 (161 respondents). Since CLDS only provides the four-digit prefecture identifier, I aggregate the sent-down youth density at the prefecture level. The primary focus of this section is coefficients β_{Expo} and β_{Post}.

Trust

Nunn and Wantchekon (2011) and Lowes and Montero (2018) suggest that important historical events may alter social trust in the long run. Table 9 Panel A reports the results of the impact of urban youth on interpersonal trust. Three questions in the CLDS capture three dimensions of interpersonal trust: general trust towards the general public (scaled from one to four), trust towards neighbors who are familiar to the respondent (scaled from one to five), and trust towards an interviewer who meets the respondent for the first time (scaled from one to four). β_{Expo} is 2.31 (*t*=2.85), 22.53 (*t*=28.16), and 0.46 (*t*=0.26) correspondingly. Thus, the arrival of sent-down youth drove up social trust, particularly in-group trust (trust among familiar people). β_{Post} is -

[34] Sent-down youth participated in almost all aspects of the rural society (e.g., as teachers, barefoot doctors, agricultural workers on collective farms, infrastructure construction workers, and even as “entrepreneurs” in state-owned enterprises).

5.96 (t=-3.05), -11.01 (t=-2.36), and -6.84 (t=-5.03) correspondingly. The rehabilitation of sent-down youth weakened all three types of trust.

Well-educated migrants may spread knowledge and build more trust in modern technology.[35] Specifically, I study medical beliefs: trust in traditional Chinese herbal therapy and trust in modern Western medicine. The medicine of the West, which was foreign to rural Chinese people, represents new technology developed through scientific research. Chinese herbal therapies feature mixes of a wide variety of herbs used to strengthen the entire immune system rather than fight specific diseases. Thus, it is very difficult to prove or disprove that herbal therapy is effective. Most traditional Chinese herbal treatments are based on medical practices rather than scientific experiments. In China, these herbal therapies still enjoy widespread popularity. Especially in underdeveloped rural areas, peasants continue to hold many superstitions about traditional herbal therapies and miss the opportunity to cure their diseases.[36]

Table 9 Panel B shows that exposure to sent-down youth discredits the traditional Chinese therapies in both specifications: β_{Expo} is -10.22 (t=-3.05) and β_{Post} is 15.07 (t=3.13). People who were more exposed to urban youth gained more confidence in Western medicine: β_{Expo} is 11.95 (t=12.99) but no evidence shows trust in Western medicine declined in post-rustication cohorts. These findings indicate that sent-down youth spread modern medical beliefs and weakened existing medical beliefs by shifting people's trust from traditional Chinese herbal therapy to modern Western medicine.

[35] Hornung (2014) shows that skilled Huguenot migrants spread technology in Prussia and improved the productivity of textile factories. Our evidence indicates trust in modern technology may serve as a channel to explain the technology adoption.

[36] Some rural peasants refuse to take antibiotics. Instead, they insist on using herbal therapies and die unnecessary.

Table 9: Trust, Medical Beliefs, and Well-being

	Panel A: Trusts						Panel B: Medical Beliefs			
	General Public (1-4: Distrust-Trust)		Neighbors (1-5: Low Trust-High Trust)		Survey Interviewers (1-4: Distrust-Trust)		Traditional Chinese Medicine (1-5: Low Trust-High Trust)		Western Medicine (1-5: Low Trust-High Trust)	
$Density_j \times Expo_t$	2.31**		22.53***		0.46		-10.22**		11.95***	
	(0.81)		(0.80)		(1.76)		(3.35)		(0.92)	
$Density_j \times Post_t$		-5.96**		-11.01**		-6.84***		15.07**		-0.50
		(1.96)		(4.67)		(1.36)		(4.81)		(2.52)
Obs.	183	161	183	161	183	161	183	161	183	161

	Panel C: Well-being Measures									
	Height (cm)		Self-reported Health (1-5: Unhealthy-Healthy)		Happiness (1-6: Unhappy-Happy)		Happiness relative to peers (1-6: Unhappy-Happy)		Fairness of life (1-5: Very Unfair-Completely Fair)	
$Density_j \times Expo_t$	134.86***		13.72***		21.01***		24.76***		6.35*	
	(34.83)		(1.74)		(0.82)		(2.01)		(3.13)	
$Density_j \times Post_t$		-103.48**		-4.79		-3.30		-6.51		-9.49
		(41.15)		(3.89)		(8.21)		(10.53)		(8.05)
Obs.	183	161	183	161	183	161	183	161	183	161

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: $Controls_j$ only includes the average years of education of the pre-rustication birth cohorts 1943-1950 in Census 2000. $Expo_t$ is the exposure dummy: one if born after 1956; zero if born before 1956. The data samples include birth cohorts from 1962 to 1982. $Post_t$ is the post-rustication dummy: one if born after 1975 (including 1975); zero if born before 1975. The data samples include birth cohorts from 1942 to 1972. Panel A reports social trust, Panel B reports medical beliefs, and Panel C reports well-being measures. β_{Expo} is the coefficient on the left and β_{Post} is the coefficient on the right for each survey question. All specifications include age and prefecture dummies. Robust standard errors are clustered at the prefecture level and reported in parentheses.

$$y_{i,j,t} = \beta_{Expo} Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

$$y_{i,j,t} = \beta_{Post} Density_j \times Post_t + \gamma\ Controls_j \times Post_t + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Well-being

Table 9 Panel C shows that exposure to sent-down youth has a positive impact on people's height, a comprehensive well-being measure that reflects nutrition, health, and environments. Even more interesting, β_{Expo} and β_{Post} are quite comparable in magnitude: 134.07 (*t*=3.81) versus -103.48 (*t*=-2.51). Thus, I conclude that the rustication cohorts are taller than the birth cohorts educated both before and after the rustication.

Did sent-down youth improve rural people's subjective feelings about the world? To answer this question, I study four measures of subjective well-being: self-reported health condition, level of happiness rating, level of happiness relative to peers, and subjective feeling of fairness. The arrival of sent-down youth (β_{Expo}) induced higher subjective health evaluation significant at the 1% level, absolute and relative happiness significant at the 1% level, and sense of fairness significant at the 10% level. The rehabilitation of urban youth generated a modest decline in subjective well-being. All four coefficients, β_{Post}, are negative, but the statistical power is limited. Rural people who were exposed to urban youth during their school years developed more harmonic mindsets: they reported feeling healthier, happier about their lives, and that the world fairly compensated their efforts.

Job Attitudes

Job attitudes can potentially affect job choices and employment outcomes. The CLDS elicits answers to six questions about the importance of the features of an ideal job, and three questions about one's willingness to expend effort for work under disadvantageous conditions. I use these nine questions to measure job attitudes. Overall job attitudes can be broadly conceptualized into two categories: objective cognitive assessments and affective job satisfaction. Economists typically focus on objective cognitive assessments, such us pay, promotion opportunities, and other benefits. However, affective job satisfaction—which reflects one's global feelings about a job—is equally important but rarely studied by economists.

Table 10: Job Attitudes

	Panel A: Importance of Job Features											
	Make a living		Build connections		Comfort myself		Gain more respect		Satisfy my interests		Exploit my talents	
$Density_j \times Expo_t$	-6.62*		-11.29*		2.22		1.28		5.32***		11.26***	
	(3.30)		(5.15)		(1.56)		(4.20)		(0.68)		(2.93)	
$Density_j \times Post_t$		-2.73		1.26		-0.12		-2.54		-5.19		-16.30*
		(2.93)		(3.08)		(3.95)		(3.59)		(6.82)		(7.48)
Obs.	170	151	167	150	170	151	169	151	169	150	169	151

	Panel B: Willingness to Pay Effort					
	In Poor Health Conditions		Undesirable Tasks		Tasks only pay off after a long period	
$Density_j \times Expo_t$	-4.66***		-3.25		0.67	
	(0.92)		(5.05)		(1.34)	
$Density_j \times Post_t$		-0.51		6.09*		8.54***
		(1.84)		(3.19)		(2.26)
Obs.	183	160	183	160	183	160

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: $Controls_j$ only includes the average years of education of the pre-rustication birth cohorts 1943-1950 in Census 2000. $Expo_t$ is the exposure dummy: one if born after 1956; zero if born before 1956. The data samples include birth cohorts from 1942 to 1972. $Post_t$ is the post-rustication dummy: one if born after 1975 (including 1975); zero if born before 1975. The data samples include birth cohorts from 1962 to 1982. Panel A reports interpersonal trust, Panel B reports medical beliefs, and Panel C reports well-being measures. β_{Expo} is the coefficient on the left and β_{Post} is the coefficient on the right for each survey question. All specifications include age and prefecture dummies. Robust standard errors are clustered at the prefecture level and reported in parentheses.

$$y_{i,j,t} = \beta_{Expo} Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

$$y_{i,j,t} = \beta_{Post} Density_j \times Post_t + \gamma\ Controls_j \times Post_t + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Among the six questions about the importance of job features, I classify "make a living" and "build connections" as objective cognitive assessments since they may have direct impact on financial benefits. The other four features—"comfort myself," "gain more respect," "satisfy personal interests," and "exploit talents"—constitute affective job satisfaction. In Table 10 Panel A, I find the arrival of sent-down youth reduced people's incentives to pursue "make a living" (β_{Expo}=-6.62, *t*=-2.01; β_{Post}=-2.73, *t*=-0.93) and "build connections" (β_{Expo}=-11.29, *t*=-2.19; β_{Post}=1.26, *t*=0.41). Among the affective job satisfaction categories, exposure to urban youth increased people's motives to "satisfy personal interests" (β_{Expo}=5.32, *t*=7.82; β_{Post}=-5.19, *t*=-0.76) and "exploit talents" (β_{Expo}=11.26, *t*=3.84; β_{Post}=-16.30, *t*=-2.18), but had no significant impact on the importance of "comfort myself" and "gain more respect." These results indicate that sent-down youth broadly shifted job attitudes from objective cognitive assessments to affective job satisfaction. People valued intrinsic values of jobs more and paid less attention to the financial benefits.

In Table 10 Panel B, I evaluate the willingness to expend effort under disadvantageous circumstances and find suggestive evidence that urban exposure may have reduced effort levels. People who were born after the arrivals of sent-down youth were less willing to pay effort amid poor health conditions (β_{Expo}=-4.66, *t*=-5.07), expend effort on undesirable tasks (β_{Post}=6.09, *t*=1.91), and spend time on tasks that only paid off after an extended period (β_{Post}=8.54, *t*=3.78). This finding is consistent with the shift toward affective job satisfaction—people were less willing to sacrifice personal amenities for work.

A potential concern is that employment or income level may be responsible for these changes in social values. To address this concern, I include two individual-level control variables: total annual income and the dummy of being employed in 2011 in specification (8), replicate Tables 9 and 10, and report results in Appendix Tables 4 and 5. The coefficients β_{Expo} and β_{Post} remain of similar magnitude and statistical significance. Income and employment status cannot explain the shifts in trust, well-being, and job attitudes.

7. Channels

As opposed to the initial goal of being "re-educated by peasants," sent-down youth volunteered to educate peasants as rural teachers. This section documents the teacher supply channel—sent-down youth became rural teachers and provided more accessible education—which is likely to explain the identified effects.[37] I present two pieces of evidence with two newly collected datasets: an unbalanced panel of county-level education variables and an unbalanced panel of sent-down youth from Chinese county gazetteers.[38] I test the teacher supply channel on both the extensive margin (before and after the rustication) and the intensive margin (during the rustication years).

Extensive Margin Test

Did higher numbers of sent-down youth induce faster educational expansion? I collect eight educational outcomes: the number of primary/secondary teachers, number of primary/secondary students, number of primary/secondary schools, and the student-teacher ratio in primary/secondary schools. $Growth_j$ represents growth rate from the average outcomes in the pre-rustication period, 1958-1967, to the average outcomes in the rustication period, 1968-1977, for each county j. Then, I use the specification (9), a sequence of univariate regressions, to show the correlation between the density of sent-down youth ($Density_j$) and educational expansion ($Growth_j$)

[37] The Zhaoyuan County Gazetteer documents, "In 1968, each commune started to organize its secondary school." The Anda County Gazetteer documents, "Since 1970, rural commune secondary schools pervasively set up high schools."

[38] I do not use the county sent-down youth panel as the primary variation because the data is sparse. Thus, I choose the statistics in 1969 from Heilongjiang Archive as the primary sent-down youth density measure.

Table 11: Pre-Post Education Expansion

	Panel A: $Obs \geq 5$			Panel B: $Obs \geq 10$		
	(1)	(2)	(3)	(4)	(5)	(6)
$growth_j$	Coefficient	t-stat	# County	Coefficient	t-stat	# County
# Secondary Teachers	70.04**	[2.23]	33	40.84*	[1.77]	24
	(31.44)			(23.07)		
# Secondary Schools	33.30	[1.66]	34	21.55	[1.41]	24
	(20.052)			(15.32)		
# Secondary Students	**48.01*** **	[3.26]	36	50.38**	[2.23]	26
	(14.74)			(22.62)		
Student-Teacher Ratio in Secondary Schools	**-4.99*** **	[-4.66]	32	**-4.02*** **	[-4.55]	23
	(1.07)			**(0.88)**		
# Primary Teachers	10.54**	[2.57]	29	**13.41*** **	[4.65]	25
	(4.10)			**(2.88)**		
# Primary School	6.09*	[1.96]	33	**8.53*** **	[3.62]	28
	(3.13)			**(2.36)**		
# Primary Students	5.12*	[1.88]	34	**7.71*** **	[4.74]	29
	(2.72)			**(1.63)**		
Student-Teacher Ratio in Primary Schools	**-1.40*** **	[-3.76]	27	-1.17**	[-2.18]	24
	(0.37)			(0.54)		

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: Panel A includes counties with more than five observations and Panel B includes counties with more than ten observations in period 1958-1980. $Growth_j$ is defined as the growth rate from the average outcomes in the pre-rustication period 1958-1967 to the average outcomes in the rustication period 1968-1977. The average outcomes are calculated from all available data.

$$Growth_j = \frac{\overline{y_{1968-1977}} - \overline{y_{1958-1967}}}{\overline{y_{1958-1967}}}$$

$Growth_j$ represents growth rate of eight variables: number of primary/secondary teachers, number of primary/secondary students, number of primary/secondary schools, and student-teacher ratio in primary/secondary schools. This table reports the coefficients of univariate regressions of $growth_j$ on sent-down youth density. Columns (1) and (4) report the coefficients β and robust standard errors in the parentheses. Columns (2) and (5) report the t-statistics in brackets. Columns (3) and (6) report the number of counties in each regression. The coefficients in boldface are significant at the 1% level.

$$Growth_j = \beta\ Density_j + \gamma + \varepsilon_j$$

Table 11 reports the tests with the "extensive margin" specification (9). Since data completeness varies across counties, I estimate the specification with two samples of counties: 1) counties with no less than 5 observations in Panel A; and 2) counties with no less than 10 observations in Panel B. Panel A and Panel B yield similar qualitative results and statistical significance although the sample size in Panel B is smaller than the sample size in Panel A. The density of sent-down youth positively correlates with the "Pre-Post" growth in numbers of teachers, students, and schools; it negatively correlates with the student-teacher ratio. This finding indicates that higher densities of sent-down youth correlates with faster educational expansion overall, both in the amount of teachers and students. Lower growth in the student-teacher ratio implies that teacher size expansion, normalized with the growth of student size, is still larger in the high-density counties. This evidence implies that sent-down youth equipped counties with a relatively sufficient supply of teachers in the "Pre-Post" sense.

Intensive Margin Test

Did more sent-down youth in the county predict higher numbers of teachers during the rustication years? I use panel regression to forecast education variables with the lagged cumulative sent-down youth density. I compile an unbalanced panel of numbers of sent-down youth from *Chinese Gazetteer Sent-down Youth Historical Archival Collection* published in 2014 by two Chinese historians, Guangyao Jin and Dalu Jin.[39] This panel data supplements the snapshot of sent-down youth density and introduces the time variation; therefore, I can test the intensive margin during the rustication period. My sample contains counties with more than 5 observations out of the 11 rustication years 1968-1978. The cumulative sent-down youth density is the total number of sent-down youth who arrived before year *t* divided by the population size in 1982.

$$Cumulative\ Density_{j,t} = \sum_{k=1968}^{k=t} \#\ urban\ youth_k \Big/ population\ 1982$$

[39] They gather material related to sent-down youth from Chinese county gazetteers. Jin and Jin (2015) explain their archival work and discuss some of their insights about the Mao's Mass Rustication Movement.

Table 12: Time-varying Sent-down Density and Educational Expansion

	Dependent Variable					
	% Primary Teachers	% Secondary Teachers	% Primary Students	% Secondary Students	% Primary Schools	% Secondary Schools
	Panel A: Equal-Weighted Regressions					
Cumulative Numbers Sent-down Youth	0.10*	0.22***	0.03	-0.07	0.002	0.002
	(0.06)	(0.05)	(0.23)	(0.10)	(0.002)	(0.001)
	Panel B: Observation-Weighted Regressions					
Cumulative Numbers Sent-down Youth	0.10*	0.25***	0.05	-0.10	0.002	0.001
	(0.06)	(0.05)	(0.20)	(0.12)	(0.002)	(0.001)
Obs.	89	70	124	100	124	90

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: The table reports the panel regressions forecasting gazetteer education variables with the lagged cumulative sent-down youth density. The $Cumulative\ Density_{j,t}$ is defined as the total number of sent-down youth until year t (including year t) divided by the 1982 population. % $y_{j,t}$ is the per-capita number of teachers, students, and schools (normalized with the population size in 1982). α_j is the county fixed effect, and δ_t is the year fixed effect. The data sample includes counties with more than 5 observations available over the sent-down youth movement episode 1968-1978. Panel A reports equal-weighted regression results. Panel B reports "$\frac{11}{\#\ obs.}$" –weighted regression results. Standard errors are reported in parentheses.

$$\%\ y_{j,t} = \beta\ Cumulative\ Density_{j,t-1} + \alpha_j + \delta_t + \varepsilon_{j,t}$$

The dependent variables, % $y_{j,t}$, are the ***per-capita*** number of primary teachers, secondary teachers, primary students, secondary students, primary schools, and secondary schools. Table 12 reports the estimates of the specification (10).

Panel A reports equal-weighted OLS estimates of β. One concern is that counties have differing amounts of missing values in the gazetteer data. Thus, equal-weight regressions assign more weights on counties with more complete data. To address this concern, I re-weight the data by "$\frac{11}{\# obs.}$" so that each county gains the same weight in Panel B. Panel A and Panel B report similar results. The higher cumulative density of sent-down youth significantly predicts more primary teachers by 0.10 (t=1.75) in Panel A and 0.09 (t=1.74) in Panel B; also, it predicts more secondary teachers by 0.22 (t=4.39) in Panel A and 0.25 (t=4.57) in Panel B. The coefficients are quite sizable: a 1% increase in the cumulative density of sent-down youth led to an increase of 0.09-0.1% primary teachers per-capita and a 0.22-0.25% increase in secondary teachers per-capita. One in ten sent-down youths taught in primary schools and 2.2 in ten sent-down youths taught in secondary schools. Please note that the sample size is tiny: only 9 counties in the primary teacher regression and 8 counties in the secondary teacher regression. Thus, the numbers might not be applicable to the entire Heilongjiang Province nor to China as a whole. However, the coefficients are sufficiently large to illustrate the critical role of sent-down youth teachers in Chinese rural educational development. A higher cumulative density of sent-down youth also forecasts increased numbers of schools, but the statistical power is weak. No evidence shows that increased numbers of sent-down youth directly predicts more students in the one lag.

Historical narratives confirm the supply channel I emphasize in this section. Chairman Zedong Mao initiated the rustication by proposing, "Educated youths should be re-educated by rural peasants" in the *People's Daily* in December, 1968. One half-century later, the *Xinhua Daily Telegraph* and People.cn, two authoritative party-owned media outlets, posted the commentary, "Sent-down youth teachers and the unexpected luck for rural kids" in November, 2015.[40] The article quotes Shuxin Liu, president of the Sent-down Youth Museum, located in Heihe City, Heilongjiang Province, "From 1968 to 1979, sent-down youth teachers taught around 10 million rural students in the entire country." Memoirs and interviews also provide anecdotal

[40] Xinhua Daily Telegraph: http://www.xinhuanet.com/mrdx/2015-11/27/c_134860573.htm. People.cn: http://politics.people.com.cn/n/2015/1129/c70731-27867709.html.

evidence for the channel of sent-down youth teachers. Two chapters in Shi and Tang (2014) comprise reflections by two sent-down youth teachers. Zhang (2015) also mentions that sent-down youth filled vacant teacher positions in rural areas. Evidence also comes from many novels describing the rural experiences of sent-down youth.[41] For example, the novel, *A Golden Sun in Beijing*, written by Rui Li, is a tragic story about a sent-down youth teacher, Zhongyin Zhang, who attempted to spread knowledge in rural villages but was jailed during the Cultural Revolution.

8. Conclusion

This paper investigates the largest urban-to-rural migration in history and documents it impacts on rural education, employment, and social values. Mao's Mass Rustication Movement caused a substantial increase in educational attainment of rural Chinese. An increase of 1% in the sent-down youth density incurred 0.1-0.2 additional years of education on average. Using two unique China gazetteer datasets, I reveal the channel that sent-down youth educated rural students as teachers in rural counties. The teacher expansion was particularly large in secondary schools.

Sent-down youth initially reduced non-agricultural employment in 1990, but increased non-agricultural employment and led to lower unemployment rates in 2000. The dynamic indicates that the benefit of being educated by sent-down youth was much more substantial in 2000 when the economy was more urbanized and industrialized. The employment benefit is stronger among factory workers, service workers, and specialists; but tiny among government officials and administrative staff. It indicates the human capital instilled by sent-down youth was particularly useful in searching for productive jobs, but not for bureaucratic employment. Moreover, sent-down youth empowered counties to attract more young and educated laborers.

This paper additionally provides a new direction for contemplating the role of human capital shocks in cultural transitions. Exposure to sent-down youth during their school years may have shaped rural children's attitudes toward their lives, other people, new technology, and their

[41] "Sent-down Youth Literature" is an important strand of Chinese literature. This type of novel describes the special rural experiences of sent-down urban youth during the Cultural Revolution. This literature reflects people's political and social backgrounds from the individual viewpoint.

jobs. Sent-down youth spread more positive subjective evaluations of individual's welfare, built more trust (both in-group and out-group), and shifted confidence from traditional herbal Chinese therapy to modern Western medicine. These results imply that sent-down youth also generated positive non-cognitive benefits in the areas where they were stationed. Moreover, this paper also documents a systematic shift in job attitudes evidenced by the rural population placing more value on affective job satisfaction and less value on objective cognitive assessments in their employment choices. People also reported less willingness to expend effort under unfavorable conditions in the workplace. These empirical findings suggest that the massive external migration generated profound changes in the country's culture and social values today.

REFERENCES

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, 493-505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, 2015, Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59, 495-510.

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012): Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, 102, 1832-1856.

Alesina, Alberto, and Paola Giuliano, 2015, Culture and Institutions. *Journal of Economic Literature* 53, no. 4: 898-944.

Bai, Liang, and Wu Ling, 2017, Economic Legacies of Cultural Revolution *Working Paper*

Bisin, Alberto, and Thierry Verdier, 2000, "Beyond the melting pot": Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits. *The Quarterly Journal of Economics* 115, no. 3: 955-988.

Bisin, Alberto, and Thierry Verdier, 2001, The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory* 97, no. 2: 298-319.

Bisin, Alberto, and Thierry Verdier, 2010, The economics of cultural transmission and socialization. *Handbook of Social Economics*, vol. 1, pp. 339-416. North-Holland, 2011.

Beaman, Lori. A., 2012: Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the US. *The Review of Economic Studies*, 79, 128-161.

Bonnin, Michel, and Krystyna Horko, 2013, *The Lost Generation: The Rustication of China's Educated Youth (1968-1980)*. Chinese University Press.

Card, David, and Ethan G. Lewis, 2007, The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts. *Mexican Immigration to the United States*. Pp. 193-228, University of Chicago Press, 2007.

Chen, Yi, Ziying Fan, Xiaomin Gu, and Li-An Zhou, 2018, Arrival of Young Talents: Send-Down Movement and Rural Education in China. *SSRN Working Paper*

Ding, Yizhuang, and Xiaomeng Liu, 2009, *History of Chinese Sent-down Youth* **[In Chinese]**. Contemporary China Publishing House.

Duflo, Esther, 2001, Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, 91, 795-813.

Dustmann, Christian, Tommaso Frattini, and Ian P. Preston, 2013, The Effect of Immigration along the Distribution of Wages. *The Review of Economic Studies*, 80, 145-173.

Honig, Emily, and Xiaojian Zhao, 2015, Sent-down Youth and Rural Economic Development in Maoist China. *The China Quarterly*, 1-23.

Giavazzi, Francesco, Ivan Petkov, and Fabio Schiantarelli, 2014, Culture: Persistence and Evolution. No. w20174 *NBER Working Paper*.

Giuliano, Paola, and Nathan Nunn, 2017, Understanding Cultural Persistence and Change. No. w23617 *NBER Working Paper.*

Glaeser, Edward L., and Joshua D. Gottlieb, 2009, The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. *Journal of Economic Literature* 47, no. 4: 983-1028.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman, 2004, Immigrating to Opportunity: Estimating the Effects of School Quality Using a Natural Experiment on Ethiopians in Israel, *The Quarterly Journal of Economics*, 119, 489-526.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman, 2009, Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence. *The Economic Journal*, 119, 1243-1269.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman, 2011, Sixty Years after the Magic Carpet Ride: The Long-run Effect of the Early Childhood Environment on Social and Economic Outcomes. *The Review of Economic Studies*, 78, 938-973.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2006, Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives* 20, no. 2: 23-48.

Hornung, Erik, 2014, Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia. *American Economic Review* 104, no. 1: 84-122.

Jin, Guangyao, and Dalu Jin, 2014, *Chinese Gazetteer Sent-down Youth Historical Archival Collection* [In Chinese]. Shanghai Bookstore Publishing House.

Jin, Guangyao, and Dalu Jin, 2015, Understand the Sent-down Youth Movement from the Chinese Gazetteers [In Chinese]. *Modern Chinese History Study* no. 3: 112-122.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, 2001, Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *The Quarterly Journal of Economics*, 116: 607-654.

Kinnan, Cynthia, Shing-Yi Wang, and Yongxiang Wang, Relaxing Migration Constraints for Rural Households, *American Journal of Economics: Applied Economics, Forthcoming*

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz, 2005, Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. *The Quarterly Journal of Economics* 120, no. 1: 87-130.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, 2007, Experimental Analysis of Neighborhood Effects. *Econometrica* 75, no. 1: 83-119.

Li, Hongbin, Mark Rosenzweig, and Junsen Zhang, 2010, Altruism, Favoritism, and Guilt in the Allocation of Family Resources: Sophie's Choice in Mao's Mass Send-down Movement. *Journal of Political Economy* 118, no. 1: 1-38.

Liu, Xiaomeng, 2004, *Dictation History of Chinese Sent-down Youth* [In Chinese]. China Social Science Press.

Lowes, Sara, and Eduardo Montero, 2018, The Legacy of Colonial Medicine in Central Africa, *Working paper, Harvard University*.

MacFarquhar, Roderick, and Michael Schoenhals, 2009, *Mao's last revolution*. Harvard University Press.

Meng, Xin, and Robert Gregory, 2007, Exploring the Impact of Interrupted Education on Earnings: The Educational Cost of the Chinese Cultural Revolution. *IZA Discussion Paper*

Montgomery, James D, 2010, Intergenerational Cultural Transmission as an Evolutionary Game. *American Economic Journal: Microeconomics* 2, no. 4: 115-36.

Moretti, Enrico, 2004a, Workers' Education, Spillovers, and productivity: evidence from plant-level production functions. *American Economic Review* 94, no. 3: 656-690.

Moretti, Enrico, 2004b, Human capital externalities in cities. *Handbook of Regional and Urban Economics*, vol. 4, pp. 2243-2291.

Nunn, Nathan, Emmanuel Akyeampong, Robert Bates, and James A. Robinson, 2014, Gender and Missionary Influence in Colonial Africa. *African Development in Historical Perspective*

Nunn, Nathan, and Leonard Wantchekon, 2011, The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review* 101, no. 7: 3221-52.

Shi, Xiaoyu and Xi Tang, 2014, *Sent-down Memory: To our Immortal Youth*. [In Chinese] The Straits Literature and Art Press

Tabellini, Guido, 2008, The Scope of Cooperation: Values and Incentives. *The Quarterly Journal of Economics* 123, no. 3: 905-950.

Tabellini, Guido, 2010, Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association* 8, no. 4 (2010): 677-716.

Rauch, James E, 1993, Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. *Journal of Urban Economics* 34, no. 3: 380-400.

Walder, Andrew G, 2009, *Fractured Rebellion*. Harvard University Press.

Walder, Andrew G, 2015, *China under Mao: A Revolution Derailed*. Harvard University Press.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta, 2014, Education and Human Capital Externalities: Evidence from Colonial Benin. *The Quarterly Journal of Economics* 130, no. 2: 703-757.

Xu, Yiqing, 2017, Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis* 25, no. 1: 57-76.

You, Yang, 2018, Language Unification, Labor and Ideology. *Working Paper*

Zhang, Junsen, Pak-Wai Liu, and Linda Yung, 2007, The Cultural Revolution and Returns to Schooling in China: Estimates based on Twins. *Journal of Development Economics* 84, no. 2: 631-639.

Zhang Ling, 2015, *Collective Memories of 48 Sent-down Youths in LiaoXi-ShenYang*. American Academic Press.

Appendix A: Generalized Permutation Test

Yang You

Last Updated: July, 2018

This section extends the synthetic control method to the setting of multiple treated units and proposes semi-parametric and non-parametric generalized permutation tests to evaluate the statistical significance of the policy impact used in the *Model-Free Approach* part of section 4. The seminal paper by Abadie, Diamond, and Hainmueller (2010) proposes the synthetic control method and the standard permutation test when only one unit is treated. Xu (2017) extends the synthetic control method to multiple treated units. In a standard permutation test, researchers apply the synthetic control method to every unit in the sample and assess whether the estimated effect in the treated unit is large enough relative to the effect estimated for a unit chosen at random. In the same spirit as the permutation inference, the generalized permutation test gauges the probability of the estimated effect in randomly selected units is larger than the estimated effect in treated units.

The generalized permutation tests follow four steps. First, researchers choose a pool of control units and apply the synthetic control method to all units using units in the pool. Second, researchers calculate the estimated effect as the treated unit minus the synthetic control unit and compute the average estimated effect in the treated units. Third, researchers randomly select units from the full sample, calculate the average estimated effect, and bootstrap. Fourth, researchers evaluate the average estimated effect of the treated units in the bootstrap distribution from random sampling.

Consider n units X_i (i $\in \{1,2,3,...,n\}$), m units X_j (j $\in \{1,2,3,...,m\}$ and $m < n$) are exposed to the intervention. Each unit X_i has the outcome variable $Y_{i,t}$ and a vector of observed covariates $a_{i,t}$. (e.g., $Y_{i,t}$ is per-capita cigarette sales in Abadie, Diamond, and Hainmueller (2010). Denote $P_{i,t}$ is the estimated effect for the unit X_i in year t, where $P_{i,t} = Y_{i,t} - Y_{i,t}^{synthetic}$ [1] In year t, the average estimated effect of the intervention is $\bar{P}_t = \frac{\sum_{i=1}^{m} P_{i,t}}{m}$. Denote $\bar{P}_t^S$ as the average estimated effect of m random units.

Non-parametric Approach

To assess the statistical significance of $\bar{P}_t$, researchers construct the bootstrap distribution of the average estimated effect $\bar{P}_t^S$. In each resampling, I randomly draw m units

[1] The $Y_{i,t}^{synthetic}$ depends on the covariates a and Y as pre-specified in the synthetic control method.

$(X_{K_1}, X_{K_2}, ..., X_{K_m})$ from n units and compute $\bar{P_t}^{S1} = \frac{\sum_{i=1}^{m} P_{K_i,t}}{m}$. After N iterations of resampling, the sequence of $\bar{P_t}^{Sj}$ where $j \in \{1, 2, 3, ...N\}$ conform the bootstrap distribution.[2]

Without loss of generality, we hypothesize the intervention has positive effect on the outcome variable $Y_{i,t}$. Define the "p-value"

$$\alpha^{np} = Prob(\bar{P_t}^{Sj} > \bar{P_t}) = \frac{Card\{j|\bar{P_t}^{Sj} > \bar{P_t}\}}{N}$$

α measures the how likely randomly m units have larger average estimated effect than the m treated units. Lower α indicates higher statistical significance for $\bar{P_t}$

Parametric Approach

When m is large, we can approximate the bootstrap distribution with a normal distribution. Define $\mu = \frac{\sum_{i=1}^{n} P_{i,t}}{n}$ and $\sigma^2 = Var(P_{i,t}) < \infty$. μ and σ are bounded because $P_{i,t}$ can only take n values.

Lemma 1: $E(\bar{P_t}^{S}) = \mu$, $Var(\bar{P_t}^{S}) = \frac{\sigma^2}{m}$

Proof:

$$E(\bar{P_t}^{S}) = E(\frac{\sum_{i=1}^{m} P_{K_i,t}}{m}) = \frac{\sum_{i=1}^{m} E(P_{K_i,t})}{m} = \mu$$

$$Var(\bar{P_t}^{Sj}) = Var(\frac{\sum_{i=1}^{m} P_{K_i,t}}{m}) = \frac{\sum_{i=1}^{m} Var(P(K_i,t))}{m^2} = \frac{\sigma^2}{m}$$

Lemma 2: $\bar{P_t}^{Sj} \xrightarrow{d} N(\mu, \frac{\sigma^2}{m})$ as m approaches infinity.

Proof: The classical Lindeberg-Levy Central Limit Theorem implies the Lemma 2.

When m is large, the normal distribution $N(\mu, \frac{\sigma^2}{m})$ can approximate the distribution of $\bar{P_t}^{S}$. We can define the Z-score

$$z = \frac{\bar{P_t} - \mu}{\frac{\sigma^2}{m}}$$

Similarly to the non-parametric approach, we define the "p-value"

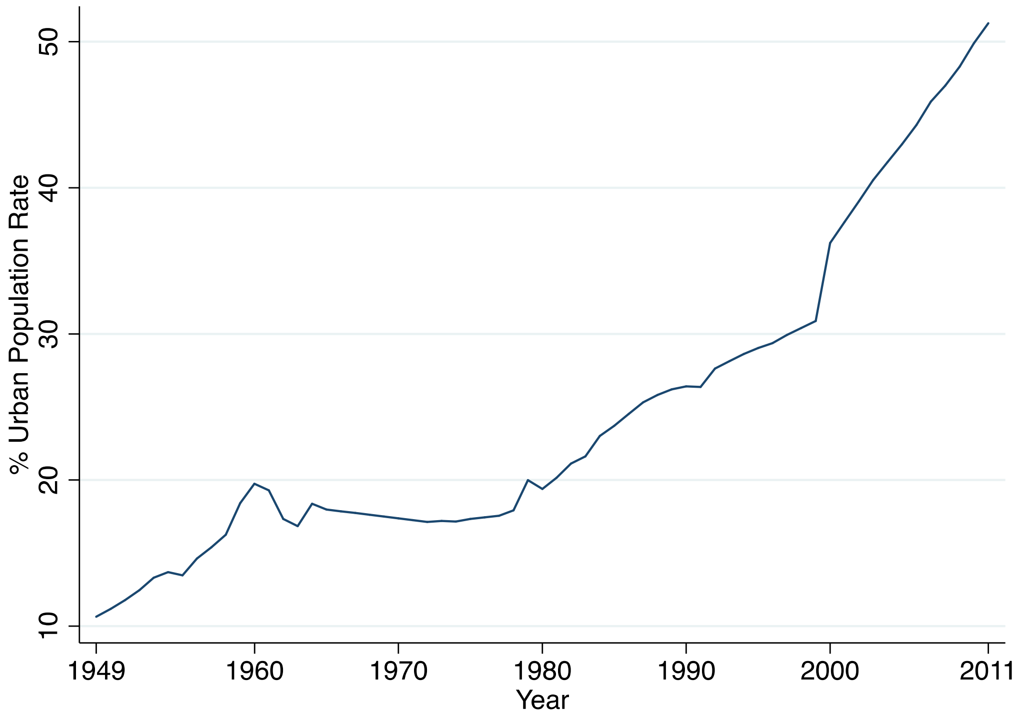
$$\alpha^{p} = 1 - \Phi(\frac{\bar{P_t} - \mu}{\frac{\sigma^2}{m}})$$

to assess the significance level.

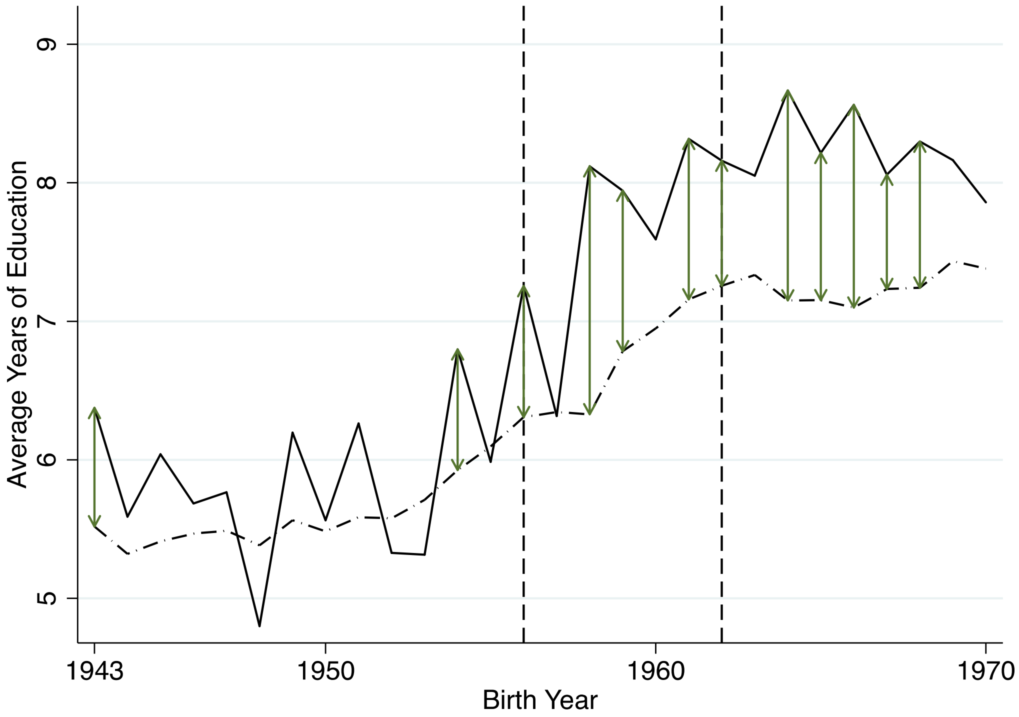
[2]For $t_1 \neq t_2$ and any given j, $P_{t_1}^{\bar{S}j}$ and $P_{t_2}^{\bar{S}j}$ are perfectly correlated. The relation is predetermined by the result of synthetic control. For $j_1 \neq j_2$ and any given t, $P_t^{\bar{S}j_1}$ and $P_t^{\bar{S}j_2}$ are i.i.d due to independent resampling process.

The advantage of the parametric approach is the computational efficiency. Researchers do not need to bootstrap for the distribution of ${\bar{P}_t}^S$. The disadvantage is that this approach can be applied only if m is reasonably large so that we can use a normal distribution to approximate the distribution of ${\bar{P}_t}^S$.

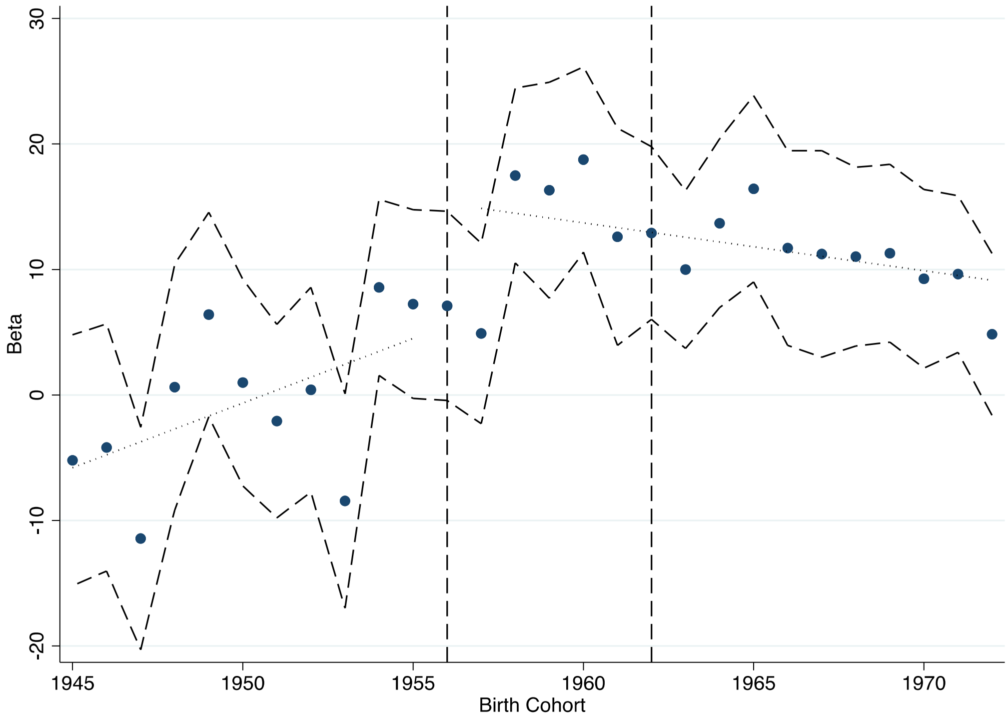
Appendix B: Figures and Tables



Appendix Fig 1. This figure plots the percentage of the urban population from 1949 to 2011. The data source is the Census Bureau of China.

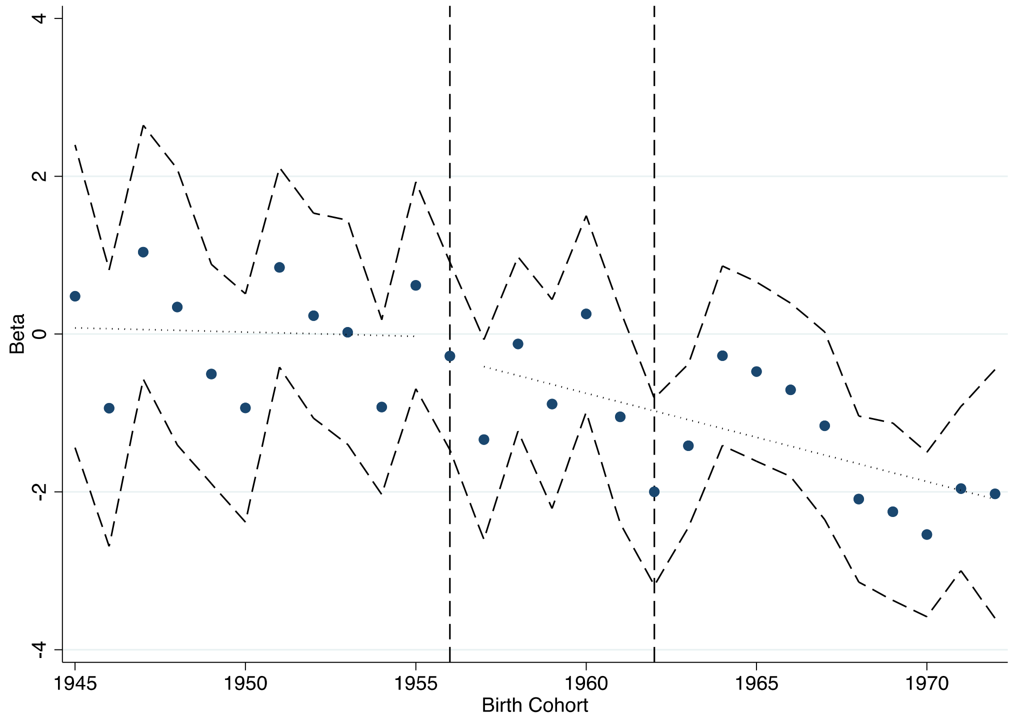


Appendix Fig 2. Average years of education in six counties with highest and 20 counties with lowest sent-down youth density. The data sample only includes the population with agricultural Hukou. The solid line is the average of the six counties with the highest sent-down youth density, and the dashed line is the average of the 20 counties with the lowest sent-down density. Arrows flag birth cohorts if the difference in education is larger than 0.8 years. The left dashed vertical line denotes birth cohort 1956, and the right one denotes birth cohort 1962.



Appendix Fig 3. This figure plots the cohort-specific coefficients β_t in the specification (3) with years of education as the dependent variable for birth cohorts 1945-1972. The data sample only includes the population with agricultural Hukou in Census 1990. The two dotted straight lines fit coefficients in pre-1956 and post-1956 birth cohorts respectively. The dashed lines plot the 90% confidence interval. The confidence interval is computed with the standard errors clustered at the county level. The left dashed vertical line denotes birth cohort 1956, and the right one denotes birth cohort 1962.

$$Eduyr_{i,j,t} = \sum\nolimits_t \beta_t Dummy_{i,j,t} * Density_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$



Appendix Fig 4. This figure plots the cohort-specific coefficients β_t with non-agricultural employment as the dependent variable for birth cohorts 1945-1972. The data sample only includes the population with agricultural Hukou in Census 1990. The two dotted straight lines fit coefficients in pre-1956 and post-1956 birth cohorts respectively. The dashed lines plot the 90% confidence interval for β_t. The confidence interval is computed with the standard errors clustered at the county level. The left dashed vertical line denotes birth cohort 1956, and the right one denotes birth cohort 1962.

$$Industry_{i,j,t} = \sum_t \beta_t Dummy_{i,j,t} * Density_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Appendix Table 1: China Gazetteer Data Collection Process.

County Name	Sent-down Youth Density	Availability	County Name	Sent-down Youth Density	Availability
Acheng	0.529%	Yes	Luobei	0.818%	Yes
Anda	0.568%	Yes	Mingshui	0.166%	Yes
Baiquan	0.272%	Yes	Mishan	0.249%	Yes
Baoqing	0.190%	No	Mulan	0.160%	Yes
Bayan	0.169%	No	Nehe	0.392%	Yes
Bei'an	0.510%	Yes	Nenjiang	1.943%	Yes
Binxian	0.009%	No	Ning'an	0.374%	No
Boli	0.336%	Yes	Qing'an	0.089%	Yes
Dedu	1.924%	Yes	Qinggang	0.056%	Yes
Dongning	0.270%	Yes	Raohe	1.287%	No
Dumeng	1.074%	Yes	Shangzhi	0.268%	Yes
Fangzheng	0.179%	Yes	Shuangcheng	0.244%	No
Fujin	0.584%	Yes	Suibin	3.012%	Yes
Fuyu	0.947%	Yes	Suihua	0.589%	Yes
Fuyuan	0.550%	Yes	Suileng	0.331%	Yes
Gannan	0.445%	Yes	Sunwu	5.243%	Yes
Hailin	0.316%	Yes	Tailai	0.412%	Yes
Hailun	0.169%	Yes	Tangyuan	0.789%	Yes
Huachuan	0.252%	Yes	Tieli	0.264%	Yes
Huanan	0.495%	Yes	Tonghe	0.511%	Yes
Hulan	0.301%	Yes	Tongjiang	4.031%	Yes
Hulin	2.059%	Yes	Wangkui	0.136%	Yes
Huma	8.357%	No	Wuchang	0.303%	Yes
Jiayin	9.314%	Yes	Xunke	7.805%	Yes
Jixian	0.766%	Yes	Yanshou	0.210%	Yes
Kedong	0.649%	Yes	Yi'an	0.521%	No
Keshan	0.447%	Yes	Yilan	0.974%	No
Lanxi	0.098%	Yes	Zhaodong	0.413%	Yes
Lindian	0.459%	Yes	Zhaoyuan	0.101%	Yes
Linkou	0.185%	No	Zhaozhou	0.034%	Yes
Longjiang	0.441%	Yes			

Notes**:** This table presents the data completeness of the 61 Heilongjiang counties. “Yes” in the “Availability” column indicates the county has annual data on students, teachers or schools. Otherwise, “No” fills in the “Availability” column.

Appendix Table 2. Educational Expansion in Population with Agricultural Hukou

	Dependent Variable					
	Census 1990			Census 2000		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: $Edu_{i,j,t}$= Years of Education					
$Density_j \times Expo_t$	11.09*	15.52***	14.03	1.96	2.79***	0.38
	(5.96)	(4.56)	(8.42)	(2.22)	(2.04)	(3.81)
Obs.	102,030	102,030	102,030	8,457	8,457	8,457
	Panel B: $Edu_{i,j,t} =$ Primary School Enrollment					
$Density_j \times Expo_t$	-0.60	0.40*	1.21*	-0.33***	-0.24*	-0.41*
	(0.40)	(0.21)	(0.71)	(0.13)	(0.13)	(0.18)
Obs.	102,030	102,030	102,030	8,457	8,457	8,457
	Panel C: $Edu_{i,j,t} =$ Middle School Enrollment					
$Density_j \times Expo_t$	2.26**	2.15**	0.44	0.66***	0.69***	0.52
	(1.00)	(1.02)	(1.52)	(0.14)	(0.20)	(0.75)
Obs.	92,858	92,858	92,858	7,959	7,959	7,959
	Panel D: $Edu_{i,j,t} =$ High School Enrollment					
$Density_j \times Expo_t$	2.24***	1.93*	0.65	0.18	0.34	0.02
	(0.77)	(0.81)	(1.13)	(0.34)	(0.38)	(0.87)
Obs.	51,659	51,659	51,659	4,957	4,957	4,957
Age and County Dummies	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes
$Linear\ Trend_j$	No	No	Yes	No	No	Yes

Notes: The data samples only include the population with agricultural Hukou born from 1942 to 1972 in Censuses 1990 and 2000. The left three columns report estimates with the 1% Census 1990 sample and the right three columns report estimates with the 0.1% Census 2000 sample. $Controls_j$ include the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Expo_t$. $Controls_j$ includes the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) include the county-specific linear trends. $Expo_t$ is the exposure dummy: 1 if born after 1962 (including 1962); 0 if born before 1956. $Density_j$ is the number of sent-down youth in 1969 over the total county population in 1982. $Edu_{i,j,t}$ is years of education in Panel A, primary school enrollment in Panel B, conditional middle school enrollment in Panel C, and conditional high school enrollment in Panel D. All specifications include age and county dummies. Robust standard errors are clustered at the county level and reported in parentheses.

$$Edu_{i,j,t} = \beta Density_j \times Expo_t + \gamma\ Controls_j \times Expo_t + Linear\ Trend_j + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Appendix Table 3. Dynamic Employment Effects of Sent-down Youth Exposure

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Non-agricultural Employment					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-0.69	-0.49	-1.67**	0.65**	0.79**	0.53*
	(0.53)	(0.54)	(0.68)	(0.31)	(0.31)	(0.31)
Obs.	76,847	76,847	76,847	5,920	5,900	5,900
	Panel B: Unemployment					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-0.65	-0.76	-0.17	0.40	0.30	0.35
	(0.79)	(0.85)	(0.89)	(0.37)	(0.40)	(0.40)
Obs.	95,516	95,516	95,516	6,813	6,789	6,789
	Panel C: Factory Workers					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-1.06*	-1.08*	-1.37**	0.25	0.22	0.18
	(0.63)	(0.64)	(0.66)	(0.32)	(0.36)	(0.35)
Obs.	72,041	72,041	72,041	5,631	5,611	5,611
	Panel D: Service Workers					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	0.12	0.12	-0.07	0.34**	0.50***	0.44***
	(0.14)	(0.14)	(0.15)	(0.14)	(0.12)	(0.11)
Obs.	70,089	70,089	70,089	5,560	5,545	5,545
	Panel E: Specialists					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	0.36**	0.52***	-0.37	0.07	0.10	-0.03
	(0.16)	(0.15)	(0.31)	(0.10)	(0.11)	(0.09)
Obs.	70,787	70,787	70,787	5,449	5,434	5,434
	Panel F: Government Officials					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	-1.43**	-1.20**	-1.31**	0.01	0.04	0.01
	(0.70)	(0.65)	(0.63)	(0.04)	(0.40)	(0.04)
Obs.	69,069	69,069	69,069	5,407	5,392	5,392
	Panel G: Administrative Staff					
	Census 1990			Census 2000		
$Density_j \times Expo_t$	0.01	0.04	-0.05	0.06	0.04	-0.003
	(0.11)	(0.10)	(0.10)	(0.05)	(0.06)	(0.04)
Obs.	68,769	68,769	68,769	5,400	5,385	5,385
$Age\ and\ County\ Dummies$	Yes	Yes	Yes	Yes	Yes	Yes
$Controls_j \times Expo_t$	No	Yes	Yes	No	Yes	Yes

Education Dummies	No	No	Yes	No	No	Yes

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes**:** The data samples only include the population with agricultural Hukou born from 1942 to 1972 in Censuses 1990 and 2000. All specifications include age and county dummies. Columns (1) and (4) report the baseline results. Columns (2) and (5) add the control variable interaction terms $Controls_j \times Expo_t$. $Controls_j$ include the log geographical distance to the nearest city and the average years of education of the pre-rustication birth cohorts 1943-1950. Columns (3) and (6) control for individual education-level dummies. The dependent variable is a dummy for non-agricultural employment in Panel A, and unemployment in Panel B, factory workers in Panel C, service workers in Panel D, specialists in Panel E, government officials in Panel F, and administrative staff in Panel G. In each panel, the left three columns reports estimation with Census 1990 and the right three columns reports estimation with Census 2000. Robust standard errors are clustered at the county level and reported in parentheses.

$$Employment_{i,j,t} = \beta Density_j \times Expo_t + \gamma \ Controls_j \times Expo_t + \xi \ Education \ \ Dummies_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Appendix Table 4: Trust, Medical Beliefs, and Well-being (with income controls)

	Panel A: Trusts						Panel B: Medical Beliefs			
	General Public (1-4: Distrust-Trust)		Neighbors (1-5: Low Trust-High Trust)		Survey Interviewers (1-4: Distrust-Trust)		Traditional Chinese Medicine (1-5: Low Trust-High Trust)		Western Medicine (1-5: Low Trust-High Trust)	
$Density_j \times Expo_t$	2.71***		24.78***		1.25		-11.69**		10.39***	
	(0.68)		(0.36)		(1.90)		(3.14)		(0.80)	
$Density_j \times Post_t$		-6.25**		-12.54**		-7.22***		15.32**		0.16
		(1.97)		(4.37)		(1.46)		(5.02)		(1.57)
Obs.	183	161	183	161	178	157	184	161	184	161
	Panel C: Well-being Measures									
	Height (cm)		Self-reported Health (1-5: Unhealthy-Healthy)		Happiness (1-6: Unhappy-Happy)		Happiness relative to peers (1-6: Unhappy-Happy)		Fairness of life (1-5: Very Unfair-Completely Fair)	
$Density_j \times Expo_t$	120.25***		12.34***		20.38***		24.22***		5.26	
	(29.87)		(2.29)		(0.85)		(2.39)		(3.57)	
$Density_j \times Post_t$		-96.83**		-3.85		-5.12		-8.37		-9.17
		(42.32)		(4.08)		(7.50)		(9.76)		(8.38)
Obs.	183	161	183	161	183	161	183	161	183	161

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes**:** $Controls_j$ only includes the average years of education of the pre-rustication birth cohorts 1943-1950 in Census 2000. $Expo_t$ is the exposure dummy: one if born after 1956; zero if born before 1956. The data samples include birth cohorts from 1962 to 1982. $Post_t$ is the post-rustication dummy: one if born after 1975 (including 1975); zero if born before 1975. The data samples include birth cohorts from 1942 to 1972. $Emp_{i,j,t}$ is the employment dummy: one if being employed in 2011; zero if not being employed in 2011. $Inc_{i,j,t}$ is the total annual income in 2011. Panel A reports social trust, Panel B reports medical beliefs, and Panel C reports well-being measures. β_{Expo} is the coefficient on the left and β_{Post} is the coefficient on the right for each survey question. All specifications include age and prefecture dummies. Robust standard errors are clustered at the prefecture level and reported in parentheses.

$$y_{i,j,t} = \beta_{Expo} Density_j \times Expo_t + \gamma \ Controls_j \times Expo_t + Emp_{i,j,t} + Inc_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

$$y_{i,j,t} = \beta_{Post} Density_j \times Post_t + \gamma \ Controls_j \times Post_t + Emp_{i,j,t} + Inc_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

Appendix Table 5: Job Attitudes (with income controls)

	Panel A: Importance of Job Features											
	Make a living		Build connections		Comfort myself		Gain more respect		Satisfy my interests		Exploit my talents	
$Density_j \times Expo_t$	-5.61		-11.24*		2.58		0.76		4.98***		11.70***	
	(3.74)		(5.22)		(1.67)		(3.95)		(1.12)		(2.78)	
$Density_j \times Post_t$		-3.15		1.05		-0.26		-1.91		-4.76		-15.61*
		(3.42)		(3.83)		(5.06)		(4.81)		(8.34)		(8.30)
Obs.	165	147	162	146	165	147	164	147	164	146	164	147

	Panel B: Willingness to Pay Effort					
	In Poor Health Conditions		Undesirable Tasks		Tasks only pay off after a long period	
$Density_j \times Expo_t$	-5.56***		-3.97		0.20	
	(1.12)		(4.94)		(1.95)	
$Density_j \times Post_t$		0.06		5.85**		7.79***
		(1.16)		(2.36)		(1.19)
Obs.	178	156	178	156	178	156

Level of Significance: * p<0.1 **p<0.05 ***p<0.01

Notes: $Controls_j$ only includes the average years of education of the pre-rustication birth cohorts 1943-1950 in Census 2000. $Expo_t$ is the exposure dummy: one if born after 1956; zero if born before 1956. The data samples include birth cohorts from 1942 to 1972. $Post_t$ is the post-rustication dummy: one if born after 1975 (including 1975); zero if born before 1975. The data samples include birth cohorts from 1962 to 1982. $Emp_{i,j,t}$ is the employment dummy: one if being employed in 2011; zero if not being employed in 2011. $Inc_{i,j,t}$ is the total annual income in 2011. Panel A reports interpersonal trust, Panel B reports medical beliefs, and Panel C reports well-being measures. β_{Expo} is the coefficient on the left and β_{Post} is the coefficient on the right for each survey question All specifications include age and prefecture dummies. Robust standard errors are clustered at the prefecture level and reported in parentheses.

$$y_{i,j,t} = \beta_{Expo} Density_j \times Expo_t + \gamma \ Controls_j \times Expo_t + Emp_{i,j,t} + Inc_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$

$$y_{i,j,t} = \beta_{Post} Density_j \times Post_t + \gamma \ Controls_j \times Post_t + Emp_{i,j,t} + Inc_{i,j,t} + \alpha_j + \delta_t + \varepsilon_{i,j,t}$$