War and Human Capital Accumulation: Evidence from the Second Sino-Japanese War

Jinyuan LIU*

May 31, 2026

Abstract

War losses are often geographically concentrated, but their long-run consequences depend on how states respond after conflict. We study this mechanism in the Sino-Japanese War by combining a new county-level measure of wartime loss, built from digitized individual martyr records in the national registry, with pooled microdata from the 1982, 1990, and 2000 Chinese population censuses. Our design is a cohort difference-in-differences that compares pre- and post-war birth cohorts across counties with different wartime-loss intensity. Post-war cohorts from higher-loss counties attain 0.09 to 0.16 additional years of schooling relative to comparable cohorts from lower-loss counties. The result is stable across treatment definitions, cutoff choices, historical county controls, and functional-form checks. The gains are concentrated at basic schooling margins and are larger in counties with weaker pre-war educational capacity. Mechanism evidence is consistent with a supply-side response through higher local education spending, alongside persistence in education-oriented values across generations.

Keywords: War exposure; Human capital; Education; Collective memory

JEL codes: I25; J24; N45; O15

*China Academy of Public Finance and Public Policy, Central University of Finance and Economics, Beijing, China. Email: jinyuanliu@email.cufe.edu.cn. I am especially grateful to Ruiming LIU, Tengyu ZHAO, and Chunbing XING for their guidance and detailed comments. I also thank Xingzhi YAO, Shihan WENG, and Fengsheng XU for helpful discussions. I thank participants at the 2026 Spring CETW (Renmin University of China), the YES × Chengdu: China Economic Theory and Policy Workshop at Southwest University of Finance and Economics (2026), and the 2026 Annual Workshop on Economics of Education in China (AWEEC) at Tsinghua University for valuable comments. All remaining errors are my own.

I Introduction

The Anti-Japanese War (1931–1945) destroyed lives and local state capacity on a massive scale. It killed an estimated 14 to 20 million Chinese soldiers and civilians (Lary, 2010; Mitter, 2013). Counties lost working-age men, schools were destroyed or requisitioned, teachers were displaced, and local public finance collapsed. The same war later became a core source of political legitimacy for the post-1949 state. As Johnson (1962) emphasized, wartime resistance forged durable ties between rural communities and the Communist Party.

This paper asks a simple question: did counties that bore heavier wartime losses receive a stronger post-war educational push? Existing work gives reasons for either sign. Conflict can reduce schooling for exposed cohorts (Ichino and Winter-Ebmer, 2004; de Walque, 2006; Lai and Thyne, 2007; Akresh and de Walque, 2008; Shemyakina, 2011; Chamarbagwala and Morán, 2011; León, 2012; Akbulut-Yuksel, 2014). It can also induce durable political and social responses, including collective action and value change (Bellows and Miguel, 2009; Blattman, 2009; Blattman and Miguel, 2010; Voors et al., 2012; Bauer et al., 2016; Becker et al., 2020). Which force dominates is an empirical question.

We bring new data to that question. We digitize individual martyr records from the *Zhonghua Yinglie Wang* (China Martyrs Website), the national registry maintained by the Ministry of Veterans Affairs (Ministry of Veterans Affairs of the People's Republic of China, 2026). The site is a memorial archive, with one searchable profile per martyr and no bulk download. A scripted crawler extracts profile-level JSON records. We retain Anti-Japanese War profiles, remove duplicates, map records to counties, and harmonize to current county boundaries using administrative-code matches and a historical crosswalk. This yields a new county-level measure of wartime loss: martyr count per 1953 county population.

We merge this county-level treatment to pooled microdata from the 1982, 1990, and 2000 population censuses. The 1982 and 1990 files are 0.1 percent random samples, and the 2000 file is from the 0.095 percent long-form questionnaire. After harmonizing geography and schooling measures, and restricting to birth years 1920–1956, the estimation sample contains about 9.1 million individuals in 2,296 counties.

Our identification strategy is a cohort difference-in-differences in the spirit of Duflo (2001).

We interact county-level wartime loss with a post-war cohort indicator and include county, birth-year, and census-wave fixed effects. The baseline defines post-war cohorts as those born in 1940 or later. A stricter design uses 1946, the first full birth year after the Japanese surrender.

The main result is positive and precisely estimated. Post-war cohorts from high-loss counties attain more schooling than comparable cohorts from lower-loss counties. Under the continuous treatment, the coefficient on $Post \times IHS(martyrs\ per\ 100{,}000)$ is 0.034. Under binary treatment definitions, the implied gain is 0.09 to 0.16 years of schooling. Event-study estimates are flat before 1940 and turn positive afterward. The result survives alternative cutoffs, functional-form choices, and controls for sent-down youth, famine intensity, Cultural Revolution victims, grain output, and urbanization.

Three additional patterns help interpret the estimate. First, gains are concentrated at basic margins: primary and junior-high completion rise, while upper-tail schooling moves little. Second, effects are larger where pre-war educational capacity was weaker. Third, effects are stronger in non-treaty-port prefectures and in counties without ancestral clans. The profile is consistent with post-war educational catch-up targeted to places with larger initial deficits and weaker private substitutes.

We then study mechanisms. On the supply side, counties with heavier wartime loss record higher fiscal education spending over 1964–1990, and higher spending aligns with improvements at basic completion margins. This pattern matches the priorities of the post-1949 literacy campaigns and the expansion of compulsory basic education (Pepper, 1996; Hannum, 1999). On the demand side, we document what we call *spiritual inheritance*: persistent value transmission shaped by wartime sacrifice and commemoration. In the 2008 China General Social Survey, respondents from high-loss counties place more weight on ambition and hard work. In the 1995 China Household Income Project, households in those counties spend more on schooling. At the prefecture level, wartime loss is positively correlated with a collectivism index.

This paper contributes to three literatures.

First, we add to work on war and human capital by showing that long-run educational effects can turn positive when wartime loss is followed by sustained, geographically targeted

reconstruction (Ichino and Winter-Ebmer, 2004; Akresh and de Walque, 2008; León, 2012; Akbulut-Yuksel, 2014). Most existing studies emphasize destruction effects on exposed cohorts. We document a different margin: within a single country and common institutional setting, counties that suffered heavier wartime losses later experienced larger gains in schooling for post-war cohorts. The result links conflict exposure to state-led catch-up, and it helps reconcile the destruction channel with the institutional-response channel emphasized in work on conflict and long-run social change (Bellows and Miguel, 2009; Voors et al., 2012; Gilligan et al., 2014; Bauer et al., 2016).

Second, we contribute to research on the long-run legacy of violence by identifying a concrete value channel linked to human-capital accumulation (Voigtländer and Voth, 2012; Cantoni et al., 2017; Bazzi et al., 2019; Fouka, 2020). Existing studies often focus on trust, civic participation, backlash, and intergroup attitudes. We instead study whether wartime experience reshapes beliefs and behavior around education investment. We show that individuals from higher-loss counties place more weight on ambition and hard work, and households from these counties spend more on schooling. This evidence provides a measurable demand-side mechanism that complements fiscal supply-side responses.

Third, we contribute to the empirical study of state-building in modern China by providing county-level evidence that the geography of wartime sacrifice shaped the geography of post-war educational expansion (Meng et al., 2015; Chen et al., 2020; Wang, 2008; Ouyang and Yuan, 2021; Chen et al., 2024). Methodologically, we build a new county-level treatment from digitized individual martyr records and link it to pooled census microdata at scale. Substantively, we show that the Anti-Japanese War, the central legitimacy event in the political narrative of the People's Republic, left a measurable developmental legacy in education. The findings clarify one channel through which wartime political capital translated into post-war resource allocation.

The rest of the paper is organized as follows. Section II provides historical background. Section III describes the data and empirical strategy. Section IV presents the main results, robustness checks, alternative interpretations, and heterogeneity. Section V investigates mechanisms. Section VI provides additional results. Section VII concludes.

II Historical Background

A The Anti-Japanese War and Local Mobilization

The Anti-Japanese War (1931–1945) was one of the most devastating episodes in modern Chinese history. Japan invaded Manchuria in September 1931 and established the puppet state of Manchukuo. Full-scale war followed the Marco Polo Bridge Incident in July 1937. By the end of 1938, Japanese forces had occupied most major coastal cities, including Shanghai, Nanjing, Guangzhou, and Wuhan, as well as the rail corridors that connected them (Eastman, 1984; van de Ven, 2017). Eight years of fighting left an estimated 14 to 20 million Chinese dead, counting both military and civilian losses from combat, famine, and disease (Lary, 2010; Mitter, 2013).

The war's footprint was vast but uneven. The formal occupation zone covered much of eastern and northern China, but fighting intensity varied widely across counties. Some saw protracted guerrilla warfare and intense social mobilization. Others were bypassed by major operations. Military geography was central. Counties along strategic rail lines were fought over repeatedly; mountainous or remote counties were harder for Japanese forces to hold. Resistance bases mattered as well. Counties within or adjacent to Communist-led base areas (*genjudi*) in northern China, especially in Shanxi, Hebei, and Shandong, saw particularly intense conflict (Johnson, 1962). Counties in the unoccupied interior, behind the Nationalist lines, experienced the war mainly through economic dislocation and refugee inflows.

As Johnson (1962) argued, wartime mobilization reshaped the relationship between the rural population and the emerging Communist state. Guerrilla warfare required deep local mobilization. Peasants served as soldiers, porters, and intelligence sources. Local cadres organized production and mutual aid. The shared experience of collective resistance created lasting bonds between communities and the Communist movement. These wartime networks formed the backbone of the post-1949 local state apparatus in much of northern China.

The war also imposed severe short-run costs on human capital. School buildings were taken over by military forces or destroyed in fighting. Local tax revenues dried up, and educational financing collapsed. Teachers were displaced or conscripted. Students were forced

to flee or abandon their studies. The direct human toll was concentrated among working-age men, depleting local labor forces across the occupied regions.

The war was a dual shock. In the short run, it destroyed physical and human capital on a massive scale. In the long run, it created a new political order. Counties that experienced intense wartime mobilization developed stronger organizational capacity, deeper ties to the emerging Communist state, and stronger claims on the resources of the post-1949 government.

B Martyrdom and the Post-War State

We measure local wartime sacrifice using county-level martyr counts from the national martyr registry. In the Chinese political system, martyrs (*lieshi*) are individuals formally recognized by the state for having died in service to the revolution or the nation. The designation is not automatic. It requires a formal application, supporting documentation, and approval by local and provincial civil affairs authorities. The criteria have evolved over time, but throughout the post-1949 period, Anti-Japanese War combatants and civilian supporters who died during the war have been eligible for recognition (Ouyang and Yuan, 2021).

Martyr designation carries concrete consequences. Surviving family members receive a monthly pension and preferential access to social services. Names are inscribed on local memorials and recorded in county chronicles. Counties with more martyrs typically have larger martyr cemeteries, more prominent memorial halls, and more active commemorative ceremonies. These practices are not merely symbolic. They connect local communities to the national revolutionary narrative and anchor the patriotic-education campaigns carried out in schools and local organizations.

The martyr system serves two functions for our analysis. It records where the human costs of war were concentrated: counties with more martyrs experienced more intense fighting, more effective resistance mobilization, or both. It also links wartime sacrifice to post-war state attention. Counties with heavier sacrifice were folded more tightly into the revolutionary narrative of the new state. The link distinguishes our treatment from generic conflict-exposure measures that capture only the destructive side of war.

C Post-War Educational Expansion

After 1949, the People's Republic launched an ambitious program of educational reconstruction. The new state treated mass illiteracy as a central obstacle to socialist construction, political integration, and economic development. In 1952, the State Council formally launched the National Literacy Campaign (*quanguo saomang yundong*), a nationwide effort to teach basic reading and writing to adults while pushing toward universal primary enrollment for children (Pepper, 1996; Hannum, 1999). The campaign was not peripheral to early state-building. It sat alongside land reform, cadre construction, and local administrative consolidation as one of the core projects of the early People's Republic.

Implementation was decentralized. County governments built schools, hired and trained teachers, and organized enrollment drives. The center set targets and guidelines; the actual provision of education depended on the capacity and initiative of local cadres. Political priority mattered. Counties classified as revolutionary base areas (*geming laoqu*), or with strong ties to the wartime resistance, stood to receive more attention and more resources.

The logic parallels Chen et al. (2020) on the sent-down youth movement. There, counties that received more sent-down youth absorbed a positive shock to the local supply of educated adults and later saw higher educational attainment among local children. Here the channel is different but the structure is the same. Wartime sacrifice created localized political pressure for educational investment, and that investment flowed through the post-1949 state apparatus.

D Channels: From Wartime Sacrifice to Post-War Education

The baseline expectation in the conflict literature is that war destroys human capital (Ichino and Winter-Ebmer, 2004; Miguel and Roland, 2011; Akbulut-Yuksel, 2014). Our argument is narrower and more historically specific. In the Chinese case, the institutional and social responses to wartime sacrifice were strong enough to generate educational catch-up for post-war cohorts. We conceptualize that response through two complementary channels.

D.1 Channel 1: Government Supply

The first channel runs through the supply of educational resources. After 1949, the new government faced a war-ravaged country and limited fiscal resources. In a centralized system, reconstruction priorities reflected political judgments about which regions should be rebuilt first. Counties that had contributed most to the resistance occupied a privileged position in the revolutionary narrative. They were classified as revolutionary old base areas (*geming laoqu*) or liberated areas (*jiefangqu*), designations that carried both symbolic prestige and practical consequences for resource allocation.

Education was built into the decentralized implementation structure. County cadres in high-sacrifice counties had stronger incentives to meet central targets for school construction, teacher recruitment, literacy work, and enrollment. They could invoke the county's wartime contribution when requesting resources from higher levels. They also faced pressure from local communities, including martyr families, to rebuild what the war had destroyed. High-sacrifice counties should therefore have devoted more fiscal resources to education and made faster progress on the basic-schooling margins the early state prioritized.

This channel generates several testable predictions. First, fiscal education expenditures should be higher in high-sacrifice counties. Second, the gains in educational attainment should be concentrated at the basic level (primary and junior high), because the post-1949 state prioritized literacy and basic education rather than elite schooling. Third, the effect should be larger where pre-war educational capacity was weaker, because the marginal return to state investment is highest where the starting level is lowest. We test these predictions in Section V.

D.2 Channel 2: Spiritual Inheritance

The second channel runs through household demand. Schools and teachers matter only if families send their children and invest in their schooling. We hypothesize that the wartime experience durably reshaped household attitudes toward education in high-sacrifice counties. We call this channel spiritual inheritance (*jingshen chuancheng*): the intergenerational transmission of values and aspirations forged by the war.

Three forces could drive this demand response. First, shared suffering and resistance strengthened community solidarity and collective commitment to reconstruction, including investment in the next generation's schooling (Bellows and Miguel, 2009; Voors et al., 2012; Bauer et al., 2016). In the Chinese case, post-war institutions reinforced this mechanism. Mutual-aid groups, production teams, patriotic study sessions, and commemorative practices carried wartime solidarity into peacetime collective action.

Second, the destruction of physical assets may have shifted household beliefs about the relative returns to physical and human capital. Land, livestock, and tools can be destroyed or confiscated. Education cannot. Families that watched their physical assets disappear had reason to shift toward human capital as a more durable store of value (Nunn, 2014).

Third, the post-war commemorative system reinforced these values across generations. Patriotic education campaigns in schools featured local martyr stories. Memorial ceremonies connected children to the wartime generation's sacrifices. County chronicles and local histories celebrated the connection between wartime resistance and post-war achievement. These practices created a sustained cultural environment in which education was valued not only as a private investment but also as a form of collective repayment for the sacrifices of the previous generation.

This demand-side channel generates its own testable predictions. Individuals from high-sacrifice counties should hold more education-oriented attitudes. Households in these counties should spend more on education. Communities with stronger wartime legacies should display stronger collective-action values. We test these predictions using survey data in Section V.

D.3 Complementarity

The supply and demand channels are complements, not competing explanations. State investment is more effective when households value education enough to send their children. Household demand is more easily met when the state provides nearby schools and qualified teachers. The interaction generates a multiplier effect: state-provided schools raise the return to household investment, and household demand raises the return to state investment. If

wartime loss strengthened both channels at once, the combined effect on educational attainment exceeds what either channel generates alone.

A positive association between wartime loss and post-war schooling is not evidence that war is beneficial for development. The Anti-Japanese War caused enormous suffering. Our claim is narrower. In the institutional, political, and cultural setting of post-1945 China, the responses to wartime loss were strong enough to produce educational catch-up that offset, and eventually reversed, the initial destruction.

III Data and Empirical Strategy

A Martyr Data

The measure of wartime loss comes from the *Zhonghua Yinglie Wang* (China Martyrs Website), the national martyr registry maintained by the Ministry of Veterans Affairs (Ministry of Veterans Affairs of the People's Republic of China, 2026). The site hosts profiles for martyrs recognized in every major Chinese conflict since the founding of the Communist Party. Each profile records the martyr's name, the conflict, the native place, and the place of sacrifice. The registry is a public memorial archive rather than a research database. There is no bulk download, and each record sits behind a separate search-result page.

We build the county-level dataset in four steps. A scripted crawler extracts the structured JSON behind each record, yielding nearly 1.89 million profiles across all periods. Each profile reports the conflict, the native place (*jiguan*), and administrative codes for the native province, prefecture, and county. We retain profiles associated with the Anti-Japanese War (1931–1945) and drop duplicates. We assign each record to a contemporary geographic unit through a hierarchical rule on the administrative codes, prioritizing the county-level node and falling back to the prefecture when the county identifier is missing or invalid. We harmonize the records to a unified current-county geography by matching sequentially on the 12-digit and 6-digit administrative codes and then applying a historical-to-current crosswalk that reflects county boundary changes since 1949. The procedure resolves the vast majority of valid records to a unique current county.

The procedure yields county-level martyr counts for 2,296 counties.[1] We normalize each county's count by its 1953 population, the earliest reliable post-war census tabulation, which predates the demographic disruption of the Great Leap Forward. The preferred continuous treatment is the inverse hyperbolic sine (IHS) of the per-capita martyr rate:

$$X_c = \text{IHS}\left(\frac{Martyrs_c}{Population1953_c} \times 100{,}000 \right).$$

The IHS approximates the natural logarithm for large values but is defined at zero, so it avoids the arbitrary-constant problem that afflicts a plain log transformation (Bellemare and Wichman, 2020). The binary alternative is an indicator for an above-median martyr count. It compares high- and low-loss counties without imposing any functional form on the treatment.

The baseline assigns each martyr to the native place (*jiguan*). Native place is the most direct proxy for the community that bore the human cost of the war and that the post-war state recognized through revolutionary-base designations, veteran benefits, and commemorative infrastructure. Each profile also records the county of death. As a robustness check, we rebuild the treatment from the place-of-death (sacrifice-place) count and replicate the baseline, robustness, heterogeneity, and mechanism regressions. Appendix C reports the results.

B Census Data

We measure educational outcomes with individual-level microdata from the 1982, 1990, and 2000 Chinese population censuses. The 1982 and 1990 files are 0.1 percent random samples. The 2000 file is the 0.095 percent long-form questionnaire. All three waves report completed schooling, birth year, county of residence, and ethnicity.

We harmonize the three waves to a common format. Years of completed schooling follow the standard Chinese-census coding: 0 for no schooling, 3 for incomplete primary, 6 for primary, 9 for junior high, 12 for senior high, 15 for a two- or three-year college (*dazhuan*), and 16 for a four-year bachelor's degree or above. We winsorize at the first and ninety-ninth

[1] Appendix A describes the registry, the crawling workflow, and the geocoding procedure in greater detail. Appendix B reports the cross-county distribution of the treatment and the cohort composition of the pooled sample.

percentiles. County codes in each wave are mapped to the same current-county geography as the martyr data, so both sides of the merge refer to the same geographic units.

The estimation sample is individuals born between 1920 and 1956, and all baseline specifications use this birth-year range unless noted otherwise. The lower bound leaves a long pre-war window. The upper bound keeps the youngest individuals at least 26 years old at the 2000 census, by which point schooling is nearly always complete. After merging and dropping records with missing education or county, the pooled sample contains 9.1 million individuals: 3.4 million from 1982, 2.8 million from 1990, and 2.9 million from 2000.

C Historical Controls

To guard against confounding by correlated historical shocks, we assemble a set of county-level controls following Chen et al. (2020): (1) sent-down youth density, the per-capita number of urban youths relocated to rural areas during the Cultural Revolution, digitized from local gazetteers (*xianzhi*); (2) Great Leap Forward famine intensity (1959–1961), constructed with the cohort-size method of Meng et al. (2015) (see also Chen and Zhou 2007); (3) per-capita Cultural Revolution victims from the county-annals dataset of Walder (2017); (4) 1965 grain output, digitized from the gazetteers; and (5) the 1964 urbanization rate, digitized from the same gazetteer sources. The mechanism section also draws on gazetteer-based school counts, teacher counts, and fiscal education expenditures. Gazetteer coverage varies across regions, so the effective sample for each control is not identical. In Section B, we interact each control with the post-war cohort indicator and add it to the baseline specification. Table 1 reports summary statistics for the main variables.

D Empirical Strategy

We estimate the effect of wartime loss on educational attainment with a cohort difference-in-differences design in the style of Duflo (2001). The design exploits two sources of variation: cross-county differences in the intensity of wartime loss, and cross-cohort differences in ex-

posure to post-war schooling. The baseline equation is:

$$Y_{icw} = \beta\,(Post_i \times X_c) + \gamma Z_i + \lambda_c + \delta_b + \mu_w + \varepsilon_{icw}, \tag{1}$$

where Y_{icw} is years of completed education for individual i born in year b, residing in county c, observed in census wave w. X_c is county-level treatment intensity. Z_i contains individual controls (an indicator for ethnic minority status). λ_c, δ_b, and μ_w are county, birth-year, and census-wave fixed effects. Standard errors are clustered at the county level (Bertrand et al., 2004).

County fixed effects absorb all time-invariant determinants of schooling, including the level of wartime loss itself. Birth-year fixed effects absorb national trends, such as the aggregate expansion of basic education in the 1950s. The coefficient β identifies the differential change in schooling across cohorts between high- and low-loss counties. The identifying assumption is that, absent the war, schooling in high- and low-loss counties would have evolved in parallel across birth cohorts.

The post-war cohort indicator $Post_i$ uses two cutoffs. The primary specification sets $Post_i = 1$ for individuals born in 1940 or later and drops the 1939 boundary cohort. Children born around 1940 reached primary-school entry age in 1946–1947 and were the first cohort to receive their primary education entirely under post-war institutions. A stricter check sets $Post_i = 1$ for individuals born in 1946 or later and drops the 1945 cohort. These individuals were born after the Japanese surrender.

Table 2 reports the unconditional difference-in-differences. Post-war schooling rises by 2.93 years in above-median counties and 2.81 years in below-median counties. The raw difference is 0.12 years. Figure 2 plots the event-study coefficients, and Figures 3a and 3b plot the residualized cohort profiles. The gap between high- and low-loss counties is flat before 1940 and widens afterward.

IV Main Results

Post-war cohorts born in counties with heavier wartime losses attained more schooling than cohorts born elsewhere. We document this in three steps. We report baseline difference-in-differences estimates and event-study evidence on parallel trends. We then show robustness to specification choices and correlated historical shocks. We close with four heterogeneity tests that sharpen the interpretation.

A Baseline and Event Study

Post-war cohorts in high-loss counties attained more schooling than counterparts in low-loss counties. Unless stated otherwise, all baseline specifications use birth years 1920–1956. Table 3 reports the pooled difference-in-differences under the 1940 cutoff. Panel A uses the IHS of martyrs per 100,000 as the continuous treatment. Panel B uses a binary indicator for above-median martyr counts. Columns (1)–(3) estimate the specification wave by wave; column (4) pools all three waves and adds census-wave fixed effects.

The pooled IHS estimate in column (4) is 0.034 and is significant at the one-percent level. A one-unit increase in IHS martyr intensity implies 0.034 additional years of schooling for post-war cohorts relative to the within-county pre-war baseline. The interquartile range of treatment is 2.5 IHS units, so moving a county from the 25th to the 75th percentile of wartime loss raises post-war schooling by 0.085 years. The binary counterpart in Panel B is 0.157: post-war cohorts in above-median counties gain 0.16 more years of schooling than cohorts in below-median counties, against a sample mean of 4.5 years.

The wave-specific estimates are tightly clustered around the pooled coefficient: 0.030 in 1982, 0.037 in 1990, and 0.033 in 2000. Three independent random samples of the adult population, drawn eight and eighteen years apart, deliver the same answer.

Table 4 repeats the exercise with a stricter 1946 cutoff. Treatment is assigned only to cohorts born after the Japanese surrender, and the 1945 boundary cohort is dropped. These children entered primary school under the People's Republic and were not directly exposed to wartime disruption. The pooled IHS estimate is 0.030 and the binary estimate is 0.091. The

baseline is not driven by schooling interruptions around the 1940 cutoff.

Figure 3 gives a nonparametric view of the same pattern. Panel (a) plots mean residualized years of schooling by birth cohort separately for above- and below-median counties. Both series rise steadily and overlap before 1940, then diverge. Panel (b) traces the gap directly: flat before 1940, widening after.

Figure 2 formalizes the timing with an event study that interacts each birth-year indicator with the treatment, using 1939 as the reference cohort. Pre-1940 coefficients hover around zero under both treatments, and the 95 percent confidence intervals contain zero throughout. Pre-trends are absent. Starting with the 1940 cohort, the coefficients turn positive and rise with each successive cohort, as expected when later cohorts receive a larger share of their schooling under post-war institutions.

B Robustness and Alternative Interpretations

The baseline could be biased if wartime loss is correlated with later shocks that independently raised schooling in the same counties. We first test functional-form choices, then add major historical confounders one by one.

Functional form. Years of schooling is bounded below at zero with a mass point there, so a multiplicative treatment effect would bias the OLS estimate. Table D1 in Appendix D re-estimates the baseline with Poisson pseudo-maximum likelihood (Silva and Tenreyro, 2006). PPML coefficients are positive and significant under both treatments. The results also hold with the raw martyr rate in place of the IHS transformation and with quartile dummies in place of the above-median indicator.

Cohort size. Wartime loss reduced the size of post-war birth cohorts in heavily affected counties, and smaller cohorts mechanically receive more schooling inputs per capita. Adding log cohort size at the county-by-birth-year level leaves the baseline coefficient essentially unchanged, as does controlling for log county population in 1953. Resource dilution is not the driver.

Migration. Treatment is assigned at the county of residence at the time of the census, so internal migration could attenuate or contaminate the coefficient. Appendix E re-estimates the baseline on the subsample of non-migrants identified through the household-registration question in the 1990 and 2000 censuses. The coefficient is essentially unchanged. Migration is not the driver either.

Schooling margins. Years of schooling pools several distinct margins. The mechanism evidence in Section V predicts that the gains are concentrated at the basic-literacy margins. Appendix F decomposes the years-of-schooling effect into primary and junior-high completion using the categorical schooling variable in each census wave directly. The effect is concentrated at the primary margin, larger and more precisely estimated for males, and stable to the addition of historical county controls.

Historical county controls. No single twentieth-century shock absorbs the treatment effect. Table 5 adds five historical covariates to the baseline, each interacted with the post-war cohort indicator so its effect varies across cohorts in the same way as the treatment. Following Chen et al. (2020), the controls are: (1) sent-down youth density during the Cultural Revolution, which channeled urban human capital into rural areas; (2) Great Leap famine intensity (1959–1961), a catastrophic shock to schooling and health (Chen and Zhou, 2007; Meng et al., 2015); (3) per-capita Cultural Revolution victims, a proxy for political disruption; (4) per-capita grain output in 1965, capturing pre-Cultural-Revolution agricultural productivity; and (5) the 1964 urbanization rate, capturing early industrialization. Column (1) reports the baseline. Columns (2)–(6) add each control in turn. Column (7) includes all five simultaneously. The coefficient on Post × IHS(rate) is stable in magnitude and significant at the 1 percent level throughout. The education effect does not collapse into any of these later shocks.

Long March and Korean War falsification. Treatment could capture a general tradition of military mobilization rather than the specific Anti-Japanese War shock. If so, the state may have rewarded these areas regardless of conflict type. Table 6 addresses this by adding martyr intensity from two other conflicts as separate controls: the Long March (1934–1935)

and the Korean War (1950–1953). Their geographies differ sharply. Long March casualties were concentrated in southern and southwestern provinces, while Korean War casualties were concentrated near the northeastern border. Column (1) reports the baseline. Columns (2) and (3) add Long March and Korean War intensity, respectively, each interacted with the post-war cohort dummy. Column (4) adds both. The Anti-Japanese War coefficient remains positive, significant, and stable, while placebo coefficients are small and inconsistent. The effect is specific to this war.

C Heterogeneity

The baseline is an average across counties and individuals. If wartime sacrifice operated through post-war state reconstruction and catch-up in basic schooling, the effect should be larger along dimensions that correspond to the gaps the catch-up was meant to close. We test four such dimensions. Each is grounded in a distinct literature, and each generates a sharp prediction. All four predictions are borne out in the data.

C.1 Gender

Mid-century rural China featured large gender gaps in schooling, with households favoring sons when resources were scarce (Hannum, 1999; Qian, 2008). The early post-war literacy campaigns and primary-school drives therefore reached boys first. If the response to wartime loss ran through those same household and policy channels, the schooling gain should appear earlier and more strongly for males.

Table 7 estimates the baseline separately by sex. The effect is larger for males than for females under both treatments; the female estimate is positive but smaller. The pattern matches the prediction.

C.2 Pre-War Educational Capacity

The catch-up literature argues that the marginal return to public educational investment is highest where pre-existing capacity is lowest (Duflo, 2001; Chen et al., 2020). New classrooms, teachers, and enrollment drives yield their largest gains when they reach children who would

otherwise receive no schooling. Chen et al. (2020) find exactly this pattern for the sent-down youth program, with effects concentrated in less-educated counties. If the post-war state turned wartime loss into reconstruction investment, the same logic predicts that the effect should concentrate in counties with the weakest schooling base.

Table 8 splits the sample at the median of pre-1940 county mean years of schooling. The effect is concentrated in the low-base group. The IHS coefficient is 0.069 in low-base counties and 0.016 in high-base counties, with statistical significance only in the former. The binary contrast is sharper: 0.352 years versus −0.120. The gains came from counties with the most room to grow.

C.3 Ancestral-Clan Presence

Ancestral clans are lineage organizations that managed communal resources in rural China and, in many places, ran lineage schools (Chen et al., 2024). In clan-heavy counties these institutions supplied a non-state channel for local education long before 1949. In counties without clans the post-1949 state was effectively the sole provider. If the estimated effect runs through post-war state-led expansion, it should be larger where no private substitute was in place (Chen et al., 2024).

Table 9 splits the sample by whether the county has any recorded ancestral clan. The IHS coefficient is positive and significant in both groups but more than twice as large without clans (0.062) as with them (0.025). State investment mattered most where the private alternative was absent.

C.4 Treaty-Port Status

Treaty ports were the earliest sites of China's engagement with Western commerce, modern schooling, and institutional reform (Jia, 2014). By the early twentieth century, prefectures containing treaty ports had denser modern schools and larger urban populations. If the post-war reconstruction response drove catch-up from a low base, the effect should be muted in treaty-port prefectures, which were already on a faster modernization path, and pronounced in non-treaty-port prefectures, which relied more on the post-war state to close their gap.

The prediction parallels the pre-war-capacity argument but uses an institutional rather than an outcome-based proxy for prior modernization.

Table 10 confirms the prediction. The schooling response is concentrated in non-treaty-port prefectures, with an IHS coefficient of 0.036 and a binary coefficient of 0.096. In treaty-port prefectures both estimates are positive but small and statistically indistinguishable from zero. Pre-war modernization, whether measured by outcomes or by institutional exposure, crowds out the marginal effect of the post-war reconstruction channel.

C.5 Summary

The four dimensions are motivated by distinct literatures: household gender preferences, educational catch-up, lineage-based private schooling, and pre-war modernization. Each predicts where the effect should be larger. All four predictions are borne out. The effect is largest for males, in low-base-education counties, in counties without clans, and in non-treaty-port prefectures. The pattern supports the interpretation that the post-war state's reconstruction response took the form of basic-schooling catch-up, targeted at counties with the largest pre-war deficits.

V Mechanisms

The baseline results show that post-war cohorts in high-loss counties attained more schooling. Heterogeneity results place these gains where catch-up is most plausible: at the bottom of the schooling distribution, in counties with weaker pre-war educational capacity, and in counties where the post-1949 state faced fewer private substitutes. This section asks why. We study two channels. The first is supply-side: the state directed more educational resources toward counties that had borne heavier wartime sacrifice. The second is demand-side: wartime experience and post-war commemoration reshaped household beliefs and community norms in ways that favored education. These channels follow the broader conflict literature (Blattman and Miguel, 2010; Voors et al., 2012; Bauer et al., 2016; Becker et al., 2020), but in China they operate through a distinct institutional setting.

The auxiliary datasets in this section are smaller than the pooled census, so the evidence is supportive rather than stand-alone. The goal is not to load all interpretation onto a single regression. It is to test whether independent sources point to the same mechanism.

A Government Educational Investment

Counties designated "revolutionary old base areas" were promised preferential treatment in post-war reconstruction. These counties were also those most deeply marked by Anti-Japanese War sacrifice (Johnson, 1962; Mitter, 2013). If political recognition translated into material support, counties with more wartime martyrs should have received more public funds for education in subsequent decades. That is the direct test of the supply-side channel.

A.1 Fiscal Education Spending

We test this prediction with county-level fiscal education expenditures digitized from local gazetteers (*xianzhi*), following Chen et al. (2020). The series report county-year fiscal aggregates with consistent coverage from 1964 onward. Because treatment is time-invariant at the county level, we collapse the panel to a long-run county average. We compute the average log fiscal education expenditure over 1964–1990 and regress it on wartime-loss intensity with province fixed effects. Identification comes from within-province comparisons: among counties in a common provincial environment, did counties that sacrificed more later spend more on education?

Table 11 reports the answer. The IHS coefficient is 0.034 and the binary coefficient is 0.327, both significant at the one-percent level. Above-median counties spent roughly 33 percent more on education on average over the long run. Wartime sacrifice translated into a durable fiscal advantage, not a brief symbolic gesture.

A.2 Basic-School Completion

If fiscal investment is the mechanism, the gains should appear where post-war policy was aimed. Early PRC education policy had two explicit priorities: reduce mass illiteracy and universalize primary schooling (Pepper, 1996; Hannum, 1999). Junior high expanded as the

next margin once primary enrollment rose. Senior high and college remained selective and were only loosely tied to county fiscal capacity. The prediction: treatment effects appear at the primary and junior-high margins and weaken sharply above them.

Table 12 tests the prediction in the pooled census with schooling-completion indicators as outcomes. The effect is concentrated exactly where the policy history would lead us to expect. Primary completion ($\geq$ 6 years) and junior-high completion ($\geq$ 9 years) rise. Under the binary treatment, post-war cohorts in high-loss counties are 2.0 to 2.9 percentage points more likely to reach those margins. Senior-high completion ($\geq$ 12 years) does not move in a consistent or statistically meaningful way. The distribution is central to the interpretation. The state did not use wartime sacrifice to create an elite-education advantage. It used sacrifice to push basic schooling further and faster in the places that had suffered most.

B Spiritual Inheritance

Supply alone is unlikely to explain the full pattern. The conflict literature shows that violence can reshape preferences long after fighting ends, often by strengthening trust, cooperation, civic identification, or forward-looking investment (Bellows and Miguel, 2009; Voors et al., 2012; Bauer et al., 2016; Becker et al., 2020). In China, wartime sacrifice was repeatedly narrated and commemorated. Local martyr halls, patriotic education, county chronicles, and revolutionary-base rhetoric linked educational advancement to a language of sacrifice and repayment. We call this demand-side channel *spiritual inheritance*. If this channel matters, individuals from high-loss counties should hold more effort-oriented beliefs, their households should spend more on education, and their communities should display stronger collectivist values.

B.1 Attitudes toward Determinants of Success

The first prediction is attitudinal. Individuals from high-loss counties should place more weight on ambition, hard work, and effort as determinants of success. The prediction follows from Voors et al. (2012) and related work showing that severe adversity can shift beliefs about effort, risk, and forward-looking behavior. If wartime sacrifice became part of the local

moral environment, the legacy should still be visible in survey attitudes decades later.

Table 13 uses the 2008 China General Social Survey (CGSS), which records the respondent's county of birth and asks respondents to rate the importance of different determinants of success in life. We regress responses on the interaction of county-of-origin wartime loss with a post-war cohort indicator, controlling for demographics and province and cohort fixed effects. Under the IHS specification, the coefficients on ambition (0.247) and hard work (0.199) are positive and statistically significant, as is the coefficient on personal connections (0.190). We place the greatest weight on the first two outcomes. Together they indicate that individuals from high-loss counties attach greater importance to effort-based determinants of success, the attitudinal shift the spiritual-inheritance channel would predict.

B.2 Household Education Spending

Attitudes alone are not enough. The second prediction is behavioral. If households in high-loss counties truly value education more, they should allocate more resources to it. Household education spending is the cleanest observable counterpart to the attitudinal channel (Kremer et al., 2019).

Table 14 uses the 1995 China Household Income Project (CHIP), which reports detailed household expenditures, including education spending. Under the IHS specification, the coefficient on log total education spending is 0.097, implying roughly 10 percent higher education expenditure in high-loss counties. The estimate is marginally significant. Statistical precision is unsurprisingly weaker here than in the census, given the smaller sample of roughly 7,000 households. Sign, size, and interpretation all align with the attitudinal evidence. Families from high-loss counties do not only say that effort and ambition matter more. They also devote more household resources to schooling.

B.3 Collectivism

The third prediction moves from individuals and households to the local social environment. Educational investment is partly a collective-action problem: schools have to be staffed, maintained, and socially supported, and enrollment has to be sustained over time (Miguel and

Gugerty, 2005; Nannicini et al., 2013). If wartime sacrifice strengthened norms of cooperation through shared mobilization and later commemoration, high-loss communities should display stronger collectivist values today. Those values do not replace fiscal investment. They amplify it.

Table 15 tests the prediction at the prefecture level using the CollectiveSim collectivism index (Wei et al., 2026), which measures the prevalence of collectivist values from text analysis of historical and contemporary sources. In a prefecture-level regression with province fixed effects, wartime loss and the collectivism index are positively and significantly correlated. High-loss communities carry the cooperative norms that make coordinated educational investment feasible.

C Summary

The mechanism evidence points to two complementary channels. On the supply side, counties with greater wartime loss received more fiscal investment in education in the post-war decades, and the gains sit at the basic-schooling margins the early PRC targeted. On the demand side, individuals from high-loss counties express more effort-oriented beliefs, their households spend more on schooling, and their communities display stronger collectivist norms. Neither channel alone captures the main pattern. Together they do. The state invested more where sacrifice was politically salient, and the investment was absorbed more effectively where communities had come to value education as part of the moral legacy of war.

VI Additional Analysis

The main results and mechanism evidence focus on years of schooling and the channels behind the gains. This section examines downstream outcomes. Career choice tests whether added schooling translated into occupational upgrading, as true human-capital accumulation would imply. Political and military engagement tests whether wartime loss also reshaped civic identification, as suggested in the literature on war and long-run preferences (Bellows

and Miguel, 2009; Blattman, 2009; Voors et al., 2012; Bauer et al., 2016). In the Chinese context, this question also connects to the long-noted link between wartime mobilization and post-war political integration (Johnson, 1962; Chen et al., 2020).

A Career Choice

The human-capital literature predicts that additional schooling shifts workers out of manual and agricultural occupations and into professional, technical, and managerial roles (Duflo, 2001; Glewwe and Kremer, 2006). If schooling gains in high-loss counties reflect real human-capital accumulation, these gains should appear in the occupational distribution. If they reflect only credential inflation, they should not. Occupational upgrading is therefore a validity check on the main education result.

Table 16 uses the 1990 census, which reports detailed occupational codes for employed respondents. We collapse occupations into four categories: white-collar (professional, technical, and managerial), service, agricultural, and manual. The specification follows the baseline cohort DID, interacting the county-level treatment with the post-war cohort indicator. Under the IHS treatment, the white-collar share rises by 0.003 and the manual share falls by 0.003, both significant at conventional levels. The binary contrasts are larger: white-collar employment rises by 1.6 percentage points and manual employment falls by 2.6 percentage points, both significant at the one-percent level. The magnitudes are modest, consistent with the size of the underlying schooling gains (0.1 to 0.2 years) and with the limited scope for occupational mobility in the early reform period. The direction confirms the schooling result: additional education moved workers out of manual occupations and into white-collar jobs.

B Party Membership and Military Service

The conflict literature predicts that wartime loss also reshapes civic identification. If shared wartime experience and state-sponsored commemoration shape willingness to serve the state, individuals from high-loss counties should be more likely to join the Communist Party and to serve in the military decades later (Bellows and Miguel, 2009; Blattman, 2009; Voors et al., 2012; Bauer et al., 2016; Becker et al., 2020). Party membership and military service are a hard

test of the spiritual-inheritance hypothesis. Both are costly forms of commitment to the post-war state, and neither can be generated by fiscal transfers alone. For China, the prediction connects directly to Johnson (1962), who argued that wartime resistance forged lasting bonds between rural communities and the Communist Party, bonds later reinforced through the commemoration of revolutionary old base areas.

We test these predictions with three waves of the China General Social Survey (CGSS). CGSS 2003 and 2006 ask about current Party membership and about whether the respondent has ever served in the military. CGSS 2008 asks about Party membership but not military service, which explains the asymmetric structure of the table. The public-use files for CGSS 2003 and 2006 release province but not county identifiers, so we aggregate the treatment to the province level for those waves. CGSS 2008 releases county identifiers, so we use the county-level treatment there. All specifications follow the baseline cohort DID with demographic controls, region fixed effects, and birth-cohort fixed effects. Standard errors are clustered at the region of treatment aggregation.

Table 17 reports the estimates. Columns (1)–(3) use Party membership as the dependent variable; columns (4)–(5) use military service. The Party-membership coefficients are small and statistically indistinguishable from zero in all three waves under both the IHS and the binary treatments. Post-war enrollment into the Party was a nationwide movement driven by centralized institutional incentives rather than by local wartime legacies, and the CGSS samples are too small to rule out modest effects.

The military-service estimates are mixed. Under the binary specification, the CGSS 2003 coefficient is -0.083 and significant at the one-percent level; the CGSS 2006 coefficient is small, positive, and only marginally significant. The mixed sign rules out a clean conclusion. Contemporary PRC military service is allocated through nationwide recruitment quotas that do not map neatly onto historical county-level loss.

Taken together, the Party and military results do not support a broad long-run imprint of wartime loss on individual political or military engagement. This null pattern is informative. The schooling, fiscal, attitudinal, and household-expenditure results point in one direction: wartime loss shaped post-war *educational* trajectories, not all downstream outcomes.

Party membership and military service were governed by centralized allocation systems only weakly tied to local fiscal resources or local cultural transmission. The effect appears most clearly in outcomes where both supply-side and demand-side channels operate locally.

VII Conclusion

This paper asks whether the Anti-Japanese War left a persistent imprint on human capital accumulation in China, and whether the post-war state's differential response to war-affected communities transmitted that shock into long-run educational outcomes. We address this question by constructing a new county-level measure of wartime loss from the national martyr registry and merging it with pooled microdata from the 1982, 1990, and 2000 population censuses. In a cohort difference-in-differences design, post-war cohorts born in counties that bore heavier wartime losses attained more schooling than comparable cohorts born elsewhere.

The result holds under alternative treatment definitions, cohort cutoffs, functional-form checks, and a rich set of historical county controls. Falsification exercises with Long March and Korean War martyrs show that the effect is specific to the Anti-Japanese War, rather than a generic consequence of participation in Party-affiliated conflict.

Three features of the evidence pin down the interpretation. The gains sit at the bottom of the schooling distribution: primary and junior-high completion rise, with little movement at the upper tail. The effect is largest where pre-war educational capacity was weakest and where ancestral clans were absent. High-loss counties received higher fiscal education spending over the post-war decades. The pattern matches a post-war state that concentrated educational catch-up on the counties that had borne the heaviest losses, with basic schooling as the margin of adjustment.

On the demand side, individuals from high-loss counties hold more education-oriented attitudes, their households spend more on schooling, and their communities display stronger collectivist values. The attitudinal and behavioral differences are consistent with spiritual inheritance: the intergenerational transmission of values and aspirations forged by the wartime experience and reinforced by post-war commemoration.

These findings do not imply that war is beneficial for development. The Anti-Japanese War produced immense suffering and loss of life (Lary, 2010; Mitter, 2013). Our claim is narrower. In this setting, counties that bore heavier wartime losses became the focus of a stronger post-war reconstruction response, and that response appeared over subsequent decades as higher educational attainment. The mechanism is not any virtue of destruction. It is the institutional and social response that destruction set in motion. More broadly, the long-run consequences of historical shocks depend not only on what they destroy, but also on how states and communities respond (Blattman and Miguel, 2010; Nunn, 2014; Bauer et al., 2016).

Two implications extend beyond the Chinese context. First, post-conflict educational reconstruction is most effective when political commitment to rebuild is geographically targeted to the communities that bore the heaviest losses. Centralized fiscal transfers alone are unlikely to produce the pattern we observe. The effect here depends on a shared narrative that linked wartime loss to claims on public resources and on local cadres with the standing and motivation to act on those claims. Second, household values and community norms are complements to state investment. Where families already view education as a way to honor past loss, public schools reach more children and the return to each unit of state spending is higher. Policy designs that ignore the demand side and assume that building schools is enough understate the importance of the local cultural environment that sustains enrollment and effort over time. These lessons are relevant to any setting in which post-conflict states rebuild human capital under fiscal constraints: the effectiveness of reconstruction depends jointly on how the state targets its resources and on the social meaning attached to education in the affected communities.

Tables

Table 1: Summary Statistics for the Pooled Census Sample

	1982		1990		2000		Pooled	
	mean	sd	mean	sd	mean	sd	mean	sd
Years of education	4.566	(4.110)	5.173	(4.179)	5.835	(3.977)	5.137	(4.126)
Minority	0.058	(0.235)	0.074	(0.261)	0.071	(0.257)	0.067	(0.250)
Post-war cohort	0.592	(0.492)	0.615	(0.487)	0.664	(0.472)	0.621	(0.485)
IHS(martyrs/100k)	3.189	(1.947)	3.172	(2.051)	3.319	(2.203)	3.223	(2.060)
High martyr county	0.756	(0.429)	0.676	(0.468)	0.639	(0.480)	0.696	(0.460)
Observations	3,540,053		2,838,456		2,748,223		9,126,732	

Notes: Columns (1)–(3) report summary statistics by census wave; column (4) reports the pooled sample. Birth cohorts 1920–1956. Years of education are winsorised at the 1st and 99th percentiles. The baseline treatment interaction in the main regressions combines the county-level wartime exposure measure, IHS(martyrs per 100,000), with the post-war cohort indicator reported in this table. Standard deviations are reported in parentheses.

Table 2: Suggestive Evidence: Mean Years of Education (Cutoff 1940)

	Pre-war (born before 1940)	Post-war (born 1940–1956)	Δ (Post − Pre)
Control (count ≤ 12)	3.501	6.311	2.810
	[1,024,765]	[1,751,112]	
Treatment (count > 12)	3.272	6.202	2.929
	[2,437,567]	[3,913,288]	
Difference (T − C)	-0.229	-0.109	**0.120**
Observations	3,462,332	5,664,400	9,126,732
Counties	2,296		

Notes: Pooled individual-level data from the 1982, 1990, and 2000 Chinese population censuses. Birth cohorts 1920–1956. Years of education winsorised at the 1st and 99th percentiles. Treatment is a binary indicator equal to one if the county's total number of officially recognised war martyrs exceeds the cross-county median (12). The last column reports $\Delta_{\text{Treatment}} - \Delta_{\text{Control}}$, the raw difference-in-differences. Sample sizes in brackets.

Table 3: Pooled Baseline: Alternative Treatment Measures (Cutoff 1940)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Binary (above median)			Inverse hyperbolic sine			Log
Post × High count	0.0900**	0.1570***					
	(0.0422)	(0.0314)					
Post × High rate			0.1468***				
			(0.0260)				
Post × IHS(count)				0.0159	0.0336***		
				(0.0097)	(0.0063)		
Post × IHS(rate)						0.0342***	
						(0.0068)	
Post × ln(per 100k)							0.0346***
							(0.0075)
Cohort size control		Yes			Yes		
Minority control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment start	1940	1940	1940	1940	1940	1940	1940
Observations	8,912,983	8,912,983	8,779,147	8,912,983	8,912,983	8,779,147	8,779,147
R-squared	0.272	0.272	0.271	0.272	0.272	0.271	0.271
Counties	2,296	2,296	2,296	2,296	2,296	2,296	2,296

Notes: Individual-level regressions pooling the 1982, 1990, and 2000 population censuses. The dependent variable is years of education (winsorised at the 1st and 99th percentiles). Birth cohorts 1920–1956; year 1939 is dropped. Columns (1)–(3) use binary treatment: "High count" and "High rate" are dummies for counties above the median martyr count (12) or martyrs per 100,000 (6.3), respectively. Columns (4)–(6) use inverse hyperbolic sine (IHS) transformations. Column (7) reports the log specification for reference. Cohort size control is ln(county × birth year × wave cell count). All specifications include county fixed effects, birth-cohort fixed effects, wave fixed effects, and a minority-status control. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Robustness: Alternative Treatment Measures (Cutoff 1946)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Binary (above median)			Inverse hyperbolic sine			Log
Post × High count	0.0701**	0.1208***					
	(0.0350)	(0.0278)					
Post × High rate			0.1208***				
			(0.0237)				
Post × IHS(count)				0.0167**	0.0315***		
				(0.0083)	(0.0054)		
Post × IHS(rate)						0.0296***	
						(0.0060)	
Post × ln(per 100k)							0.0309***
							(0.0066)
Cohort size control		Yes			Yes		
Minority control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment start	1946	1946	1946	1946	1946	1946	1946
Observations	8,868,666	8,868,666	8,735,698	8,868,666	8,868,666	8,735,698	8,735,698
R-squared	0.271	0.271	0.270	0.271	0.271	0.270	0.270
Counties	2,296	2,296	2,296	2,296	2,296	2,296	2,296

Notes: Same specification as Table 3 but with post-war cohort defined as born in 1946 or later; year 1945 is dropped. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Historical County Characteristics Added One by One (Pooled Censuses)

	(1) Baseline	(2) +SDY	(3) +Famine	(4) +CR victims	(5) +Grain	(6) +Urban	(7) +All five
Panel A: IHS(rate) treatment							
Post × IHS(rate)	0.0342***	0.0342***	0.0403***	0.0382***	0.0368***	0.0346***	0.0442***
	(0.0068)	(0.0066)	(0.0075)	(0.0069)	(0.0070)	(0.0067)	(0.0067)
Panel B: Binary (above median) treatment							
Post × High count	0.0941**	0.1127***	0.0866	0.0832	0.0921*	0.0939**	0.1090**
	(0.0469)	(0.0396)	(0.0639)	(0.0519)	(0.0488)	(0.0458)	(0.0548)
Historical control	None	SDY density	Famine	CR victims	Grain output	Urban ratio	All five
Cohort size control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minority control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs (IHS)	8,779,147	8,779,147	8,779,147	8,779,147	8,779,147	8,779,147	8,779,147
Obs (Dummy)	8,912,983	8,912,983	8,912,983	8,912,983	8,912,983	8,912,983	8,912,983

Notes: This table adds the five historical county characteristics one at a time to the pooled individual-level *eduy* specification using the 1940 cutoff. Column (1) is the baseline with no historical controls. Columns (2)–(6) add each control interacted with the post-war cohort indicator: SDY density (from local gazetteers, scaled by 1964 county population), famine intensity (excess death rate during 1959–1961), ln(victims_cr + 1) (from the China Political Events Dataset, scaled by 1964 county population), ln(grain_output + 1) (grain output in 1965 from county statistical yearbooks), and the 1964 urbanization ratio (share of urban population from the 1964 population census). Column (7) includes all five controls simultaneously. Panel A reports the IHS(rate) specification. Panel B reports the binary above-median specification with cohort-size control. All specifications pool the 1982, 1990, and 2000 censuses and include the minority control, county fixed effects, birth-year fixed effects, and wave fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Additional Revolutionary-history Controls (Pooled Censuses)

	(1) Baseline	(2) +Long March	(3) +Korean War	(4) +Both
Panel A: IHS(rate) treatment				
Post × IHS(rate)	0.0342***	0.0392***	0.0309***	0.0369***
	(0.0068)	(0.0070)	(0.0066)	(0.0067)
Panel B: Binary (above median) treatment				
Post × High count	0.0941**	0.0998**	0.1043***	0.1140***
	(0.0469)	(0.0489)	(0.0385)	(0.0395)
Long March control (IHS)		Yes		Yes
Korean War control (IHS)			Yes	Yes
Cohort size control	Yes	Yes	Yes	Yes
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Obs (IHS)	8,779,147	8,779,147	8,779,147	8,779,147
Obs (Dummy)	8,912,983	8,912,983	8,912,983	8,912,983

Notes: This table starts from the pooled individual-level *eduy* baseline using the 1940 cutoff. Column (1) is the baseline. Column (2) adds Long March martyr intensity (IHS-transformed per 100,000 population, consistent with the main treatment) interacted with the post-war indicator. Column (3) adds Korean War martyr intensity (IHS-transformed). Column (4) adds both simultaneously. Panel A reports the IHS(rate) specification. Panel B reports the binary above-median specification with cohort-size control. All specifications pool the 1982, 1990, and 2000 censuses and include the minority control, county fixed effects, birth-year fixed effects, and wave fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Heterogeneity by Gender (Pooled Censuses)

	Female		Male	
	Dummy	IHS(rate)	Dummy	IHS(rate)
Post × Treatment	0.0082	-0.0023	0.2023***	0.0752***
	(0.0481)	(0.0098)	(0.0492)	(0.0086)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Treatment start	1940	1940	1940	1940
Observations	2,652,530	2,606,622	2,805,854	2,758,250

Notes: This table reports the pooled individual-level education regression separately by gender, using the 1940 cutoff. Columns (1)–(2) report results for females; columns (3)–(4) for males. Within each group, the first column uses the binary above-median count indicator and the second uses IHS(martyrs per 100k). All specifications include the minority control, county fixed effects, birth-year fixed effects, and wave fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 8: Heterogeneity by Pre-war Education Level (Pooled Censuses)

	Low base education		High base education	
	Dummy	IHS(rate)	Dummy	IHS(rate)
Post × Treatment	0.3518***	0.0688***	-0.1196	0.0155
	(0.0457)	(0.0088)	(0.0731)	(0.0125)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Treatment start	1940	1940	1940	1940
Observations	3,806,520	3,735,595	5,106,463	5,043,552

Notes: This table splits the pooled sample by pre-war county-level educational infrastructure using the 1940 cutoff. Pre-war education is measured as the mean years of education among pre-1940 birth cohorts in each county. Counties are split at the median. Columns (1)–(2) report results for counties with below-median pre-war education; columns (3)–(4) for above-median counties. Within each group, the first column uses the binary above-median count indicator and the second uses IHS(martyrs per 100k). All specifications pool the 1982, 1990, and 2000 censuses and include the minority control, county fixed effects, birth-year fixed effects, and wave fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 9: Heterogeneity by Ancestral-Clan Density (Pooled Censuses)

	No ancestral clans		Has ancestral clans	
	Dummy	IHS(rate)	Dummy	IHS(rate)
Post × Treatment	0.0663	0.0620***	0.1124***	0.0245***
	(0.0767)	(0.0127)	(0.0400)	(0.0081)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Treatment start	1940	1940	1940	1940
Observations	4,252,704	4,211,059	4,660,279	4,568,088

Notes: This table splits the pooled sample by whether the county has any recorded ancestral clans, using the 1940 cutoff. Columns (1)–(2) report results for counties with no ancestral clans; columns (3)–(4) for counties with ancestral clans. Within each group, the first column uses the binary above-median count indicator and the second uses IHS(martyrs per 100k). All specifications pool the 1982, 1990, and 2000 censuses and include the minority control, county fixed effects, birth-year fixed effects, and wave fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 10: Heterogeneity by Treaty-port Status (Pooled Censuses)

	Non-treaty-port		Treaty-port	
	Dummy	IHS(rate)	Dummy	IHS(rate)
Post × Treatment	0.0961**	0.0355***	0.1031	0.0243
	(0.0425)	(0.0071)	(0.0645)	(0.0168)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Treatment start	1940	1940	1940	1940
Observations	7,412,828	7,296,957	1,500,155	1,482,190

Notes: This table splits the pooled sample by whether the prefecture historically contained a treaty port, using the 1940 cutoff. Columns (1)–(2) report results for non-treaty-port prefectures; columns (3)–(4) for treaty-port prefectures. Within each group, the first column uses the binary above-median count indicator and the second uses IHS(martyrs per 100k). All specifications pool the 1982, 1990, and 2000 censuses and include the minority control, county fixed effects, birth-year fixed effects, and wave fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 11: Fiscal Education Spending

	(1) IHS(rate)	(2) High martyr (dummy)
Treatment	0.0339***	0.3272***
	(0.0109)	(0.0557)
Province FE	Yes	Yes
Observations	637	656

Notes: Cross-sectional OLS regressions of county-average log fiscal education expenditure, averaged over 1964–1990, on county-level wartime sacrifice. Column (1) uses IHS(martyrs per 100,000) as the treatment variable. Column (2) uses a binary indicator for above-median martyr count. Both specifications include province fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. We use the 1964–1990 window because it is the longest period with consistent county-level fiscal education data digitized from local gazetteers (*xianzhi*), following the source convention of Chen et al. (2020). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 12: Basic-schooling Completion Rates (Pooled Censuses)

	Cutoff 1940			Cutoff 1946		
	Primary	Junior high	Combined	Primary	Junior high	Combined
Panel A: IHS(rate) treatment						
Post × IHS(rate)	0.0056***	0.0012	0.0034***	0.0050***	-0.0001	0.0025***
	(0.0009)	(0.0012)	(0.0007)	(0.0008)	(0.0011)	(0.0006)
Panel B: Binary (above median) treatment						
Post × High count	0.0142***	0.0059	0.0101**	0.0126***	0.0017	0.0072**
	(0.0048)	(0.0063)	(0.0045)	(0.0042)	(0.0052)	(0.0035)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	County × cohort × wave population					
Obs (IHS)	197,482			197,482		
Obs (Dummy)	201,005			201,007		

Notes: This table reports pooled county×cohort×wave completion rates for three outcomes: primary completion (eduy $\geq$ 6 years), junior-high completion (eduy $\geq$ 9 years), and a combined basic-schooling measure equal to the average of the two rates. In every case, the denominator is the full county×cohort×wave population. Columns (1)–(3) use the 1940 cutoff; columns (4)–(6) use the 1946 cutoff. Panel A uses IHS(martyrs per 100k) and Panel B uses the above-median martyr-count indicator. All specifications include county fixed effects, birth-year fixed effects, and wave fixed effects, and are weighted by county×cohort×wave population. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 13: Attitudes and Educational Values (CGSS 2008)

	(1) Ambition	(2) Parents' edu	(3) Rich only	(4) Hard work	(5) Hometown	(6) Connections	(7) Religion
Panel A: IHS(rate) treatment							
Post × IHS(rate)	0.2468***	0.1495***	0.2117***	0.1987**	0.1372**	0.1903**	0.1788*
	(0.0642)	(0.0431)	(0.0747)	(0.0769)	(0.0533)	(0.0821)	(0.0911)
Panel B: Binary (above median) treatment							
Post × High count	0.5136**	0.2145	0.3811	0.5616**	0.2761**	0.4815**	0.1468
	(0.2018)	(0.1713)	(0.2733)	(0.2159)	(0.1377)	(0.2204)	(0.3187)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs (IHS)	2,368	2,387	2,377	2,386	2,394	2,381	2,314
Obs (Dummy)	2,396	2,415	2,405	2,414	2,422	2,409	2,342

Notes: This table reports the effect of wartime martyr exposure on individual attitudes from the CGSS 2008. The dependent variables measure agreement with whether each factor matters for success, coded on a standardized scale. Panel A uses IHS(martyrs per 100,000) as the treatment variable; Panel B uses a binary indicator for above-median county martyr count. All specifications include individual controls (female, minority, urban hukou, local hukou), county fixed effects, and birth-cohort fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 14: Household Education Spending (CHIP 1995)

	(1) ln(1+Edu total)	(2) ln(1+School fee)	(3) ln(1+Training)
Panel A: IHS(rate) treatment			
Post × IHS(rate)	0.0969*	0.0777	-0.0018
	(0.0545)	(0.0833)	(0.0289)
Panel B: Binary (above median) treatment			
Post × High count	0.3704*	0.2115	0.0445
	(0.2207)	(0.3170)	(0.1038)
Household controls	Yes	Yes	Yes
Prefecture FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Obs (IHS)	6,940	6,940	6,940
Obs (Dummy)	6,940	6,940	6,940

Notes: This table uses pooled urban and rural CHIP 1995 households, with county-level treatment merged when available and prefecture-level fallback otherwise. The dependent variables are $\ln(1 + y)$ transforms of total education spending, schooling fees, and training expenditures. Panel A uses IHS(martyrs per 100,000); Panel B uses a binary indicator for above-median martyr count. Household controls include household size, head gender, head working status, head years of education, and an urban-rural indicator. Standard errors, clustered at the matched prefecture level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 15: Collectivism as a Long-run Cultural Outcome

Panel A: Province Level (IHS + Dummy)

	(1) 1982	(2) 1990	(3) 2000
Panel A: IHS(rate) treatment			
IHS(rate)	0.1871	0.1204	0.3502*
	(0.3331)	(0.2232)	(0.1897)
Panel B: Binary (above median) treatment			
High count	0.5896	0.4790	1.4838**
	(1.0457)	(0.8057)	(0.6650)
Cross-sectional OLS		Yes	
Robust SE		Yes	
Observations	28	29	29

Panel B: Prefecture Level (2000, ln treatment)

	2000
Log martyr exposure per 100,000	0.187***
	(0.059)
Province FE	Yes
Observations	309
R-squared	0.033
Adj. R-squared	0.030

Notes: Panel A reports province-level cross-sectional regressions of the CollectiveSim collectivism index on province-level wartime martyr intensity for three census years. Panel B reports the prefecture-level regression for 2000, including province fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 16: Career Choice and Occupational Sorting (1990 Census)

	(1) White-collar	(2) Agricultural	(3) Manual
Panel A: IHS(rate) treatment			
Post × IHS(rate)	0.0026**	0.0007	-0.0032**
	(0.0010)	(0.0006)	(0.0014)
Panel B: Binary (above median) treatment			
Post × High count	0.0160***	0.0095***	-0.0255***
	(0.0049)	(0.0031)	(0.0071)
Individual controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Obs (IHS)	3,900,471	3,900,471	3,900,471
Obs (Dummy)	3,972,237	3,972,237	3,972,237

Notes: This table examines whether wartime exposure shifted occupational sorting in the 1990 census. Column (1) uses a white-collar indicator; column (2) an agricultural-occupation indicator; column (3) a manual-occupation indicator. Panel A uses IHS(martyrs per 100,000) as the treatment variable; Panel B uses a binary indicator for above-median county martyr count. All specifications include individual controls, county fixed effects, and birth-year fixed effects. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

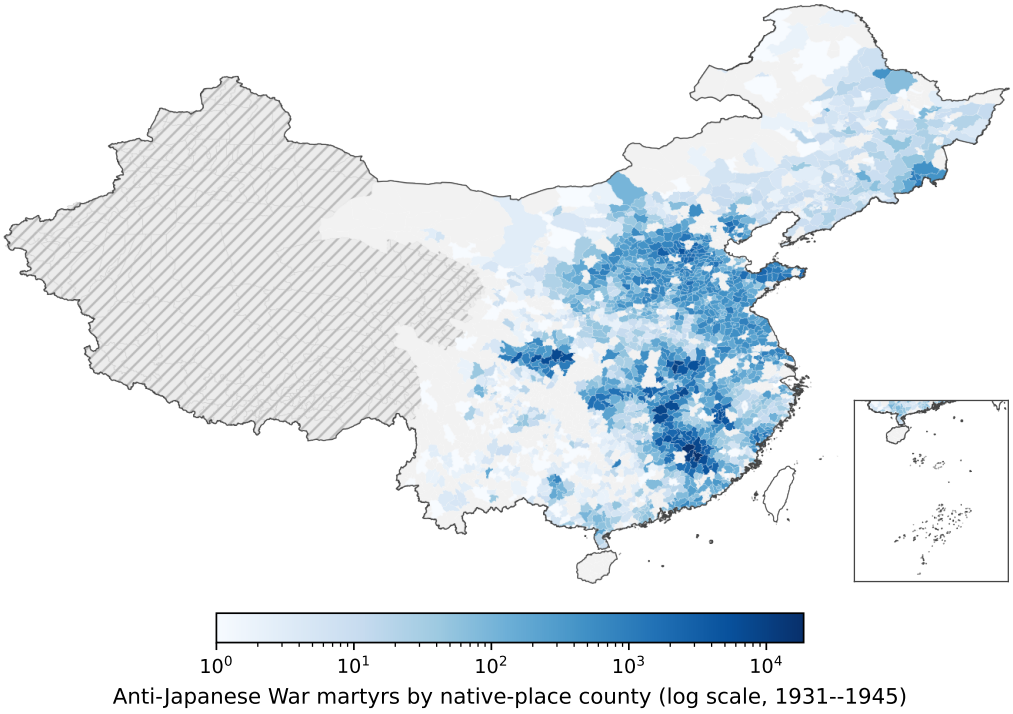
Table 17: Party Membership and Military Service

	CCP Membership			Military Service	
	(1) CGSS 2003	(2) CGSS 2006	(3) CGSS 2008	(4) CGSS 2003	(5) CGSS 2006
Panel A: IHS(rate) treatment					
Post × IHS(rate)	0.0090	-0.0160	0.0147	-0.0197*	0.0001
	(0.0144)	(0.0184)	(0.0230)	(0.0101)	(0.0100)
Panel B: Binary (above median) treatment					
Post × High	0.0507	-0.0162	0.0396	-0.0831***	0.0469*
	(0.0460)	(0.0484)	(0.0753)	(0.0239)	(0.0256)
Individual controls	Yes	Yes	Yes	Yes	Yes
Region FE	Prov.	Prov.	County	Prov.	Prov.
Cohort FE	Yes	Yes	Yes	Yes	Yes
Observations	5,627	9,969	4,771	5,720	9,946

Notes: This table examines whether wartime exposure is associated with differential probabilities of Communist Party membership and military service. Columns (1)–(3) use party membership as the outcome; columns (4)–(5) use military service. Data come from the 2003, 2006, and 2008 waves of the China General Social Survey (CGSS). The 2008 microdata do not record military service, so columns (4)–(5) include only the 2003 and 2006 waves. The 2003 and 2006 public-use files do not release county identifiers, so wartime exposure is aggregated to the province of birth; the 2008 microdata permit county-level treatment. Panel A uses IHS(martyrs per 100,000) as the treatment variable; Panel B uses a binary indicator for above-median intensity. All specifications include individual demographic controls, region (province or county) fixed effects, and birth-cohort fixed effects. Standard errors clustered at the region of treatment aggregation are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

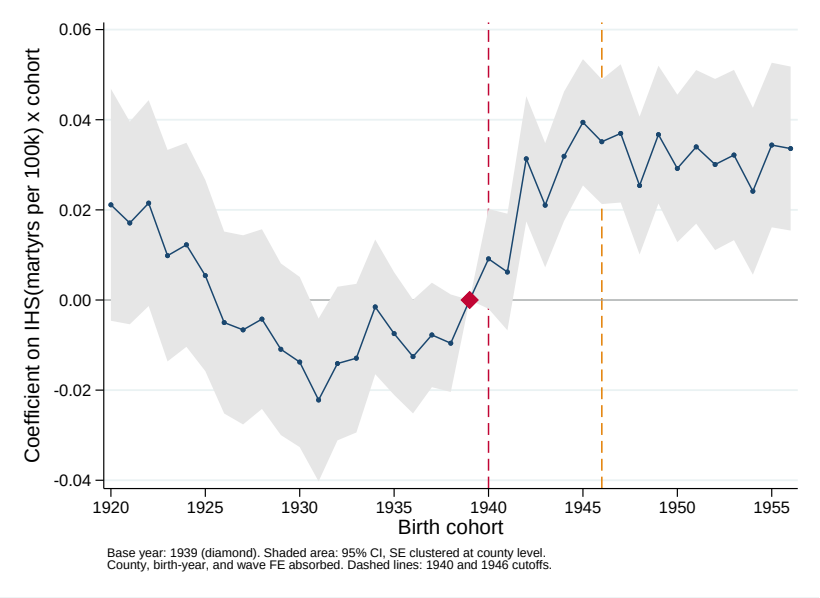
Figures

Figure 1: Geographic Distribution of Martyr Intensity, 1931–1945



Notes: County-level counts of recognized Anti-Japanese War martyrs (1931–1945) assigned to each martyr's native place (*jiguan*). Counts are plotted on a logarithmic color scale. Counties with no recorded martyrs appear in light grey. Xinjiang, Tibet, and Qinghai are hatched to indicate that they are excluded from the analysis. The lower-right inset reproduces the standard nine-dash line. Appendix B reports the same map alongside the place-of-sacrifice variant and three additional intensity scales.

Figure 2: Event Study: Cohort-by-Cohort Interaction Coefficients



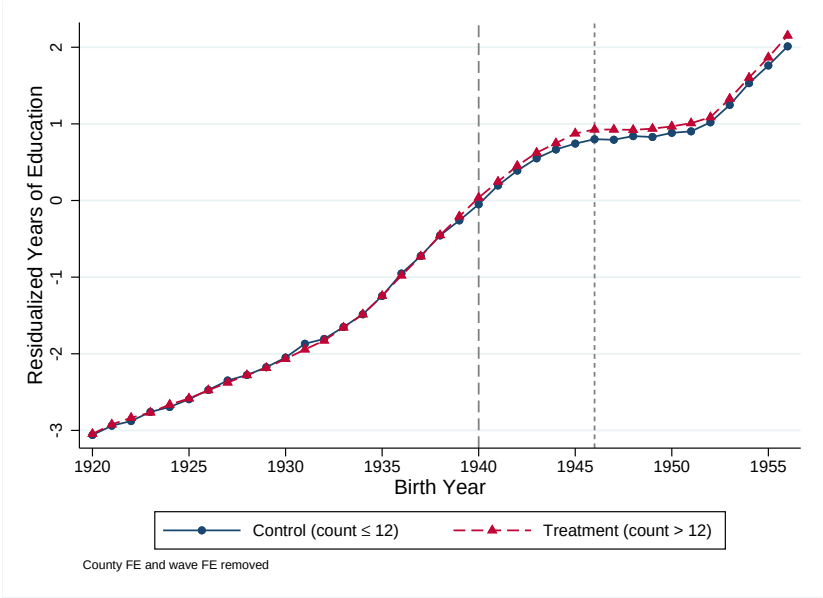
(a) IHS(martyrs per 100k)



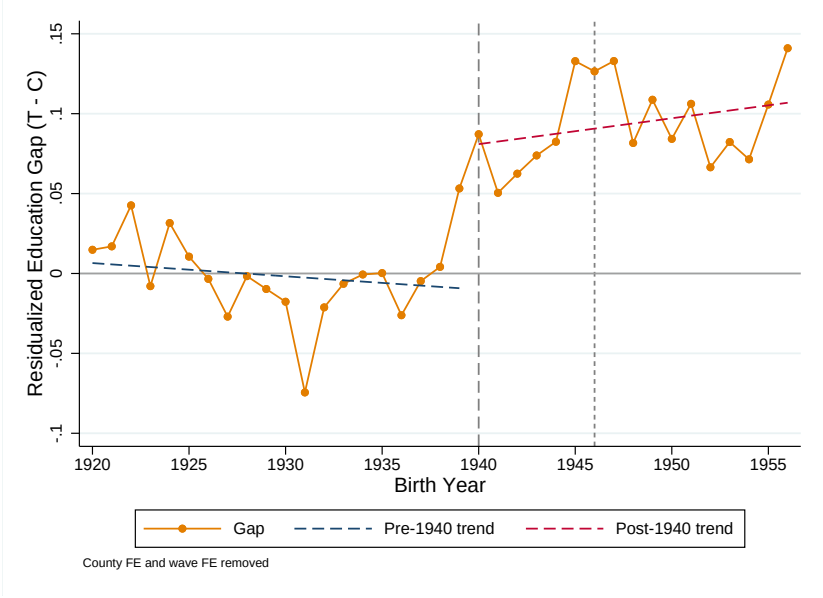
(b) Above-median dummy

Notes: Each point plots the coefficient on the interaction between a birth-cohort indicator and the treatment variable from a single pooled regression using the 1982, 1990, and 2000 censuses. Panel (a) uses IHS(martyrs per 100,000); panel (b) uses a dummy for above-median county martyr count. The base cohort is 1939 (normalized to zero, shown as diamond). County, birth-year, and census-wave fixed effects are absorbed; individual minority status is controlled. Standard errors clustered at the county level; shaded area shows the 95% confidence interval. Dashed lines mark the 1940 and 1946 cutoffs.

Figure 3: Cohort Profiles of Schooling Levels and Treatment Gaps



(a) Mean years of education by treatment status



(b) Education gap (treatment minus control)

Notes: Panel (a) plots the mean residualized years of education for each birth cohort by treatment status, pooling the 1982, 1990, and 2000 censuses after removing county and wave fixed effects. Panel (b) plots the corresponding treatment-control gap by cohort. Treatment is defined as county martyr count above the sample median. Vertical lines mark the 1940 and 1946 cutoffs; dashed lines in panel (b) are fitted separately for the pre-1940 and post-1940 cohorts.

References

Akbulut-Yuksel, Mevlude, "Children of War: The Long-Run Effects of Large-Scale Physical Destruction and Warfare on Children," *Journal of Human Resources*, 2014, *49* (3), 634–662.

Akresh, Richard and Damien de Walque, "Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide," *World Bank Policy Research Working Paper*, 2008, (4606).

Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts, "Can War Foster Cooperation?," *Journal of Economic Perspectives*, 2016, *30* (3), 249–274.

Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong, "Unity in diversity? How intergroup contact can foster nation building," *American Economic Review*, 2019, *109* (11), 3978–4025.

Becker, Sascha O, Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya, "Forced migration and human capital: evidence from post-WWII population transfers," *American Economic Review*, 2020, *110* (5), 1430–63.

Bellemare, Marc F. and Casey J. Wichman, "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics*, 2020, *82* (1), 50–61.

Bellows, John and Edward Miguel, "War and Local Collective Action in Sierra Leone," *Journal of Public Economics*, 2009, *93* (11–12), 1144–1157.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Blattman, Christopher, "From Violence to Voting: War and Political Participation in Uganda," *American Political Science Review*, 2009, *103* (2), 231–247.

_ **and Edward Miguel**, "Civil War," *Journal of Economic Literature*, 2010, *48* (1), 3–57.

Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman, and Y Jane Zhang, "Curriculum and ideology," *Journal of Political Economy*, 2017, *125* (2), 338–392.

Chamarbagwala, Rubiana and Hilcías E. Morán, "The Human Capital Consequences of Civil War: Evidence from Guatemala," *Journal of Development Economics*, 2011, *94* (1), 41–61.

Chen, Shuo, Raymond Fisman, Xiaohuan Lan, Yongxiang Wang, and Qing Ye, "The Costs and Benefits of Clan Culture: Elite Control versus Cooperation in China," NBER Working Paper 32414, National Bureau of Economic Research 2024.

Chen, Yi, Ziying Fan, Xiaomin Gu, and Li-An Zhou, "Arrival of Young Talent: The Send-Down Movement and Rural Education in China," *American Economic Review*, 2020, *110* (11), 3393–3430.

Chen, Yuyu and Li-An Zhou, "The Long-term Health and Economic Consequences of the 1959–1961 Famine in China," *Journal of Health Economics*, 2007, *26* (4), 659–681.

de Walque, Damien, "The Socio-Demographic Legacy of the Khmer Rouge Period in Cambodia," *Population Studies*, 2006, *60* (2), 223–231.

Duflo, Esther, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 2001, *91* (4), 795–813.

Eastman, Lloyd E., *Seeds of Destruction: Nationalist China in War and Revolution, 1937–1949*, Stanford: Stanford University Press, 1984.

Fouka, Vasiliki, "Backlash: The unintended effects of language prohibition in US schools after World War I," *The Review of Economic Studies*, 2020, *87* (1), 204–239.

Gilligan, Michael J., Benjamin J. Pasquale, and Cyrus Samii, "Civil War and Social Cohesion: Lab-in-the-Field Evidence from Nepal," *American Journal of Political Science*, 2014, *58* (3),

604–619.

Glewwe, Paul and Michael Kremer, "Schools, Teachers, and Education Outcomes in Developing Countries," in Eric A. Hanushek and Finis Welch, eds., *Handbook of the Economics of Education*, Vol. 2, Elsevier, 2006, pp. 945–1017.

Hannum, Emily, "Political Change and the Urban-Rural Gap in Basic Education in China, 1949–1990," *Comparative Education Review*, 1999, *43* (2), 193–211.

Ichino, Andrea and Rudolf Winter-Ebmer, "The Long-Run Educational Cost of World War II," *Journal of Labor Economics*, 2004, *22* (1), 57–86.

Jia, Ruixue, "The Legacies of Forced Freedom: China's Treaty Ports," *Review of Economics and Statistics*, 2014, *96* (4), 596–608.

Johnson, Chalmers A., *Peasant Nationalism and Communist Power: The Emergence of Revolutionary China, 1937–1945*, Stanford: Stanford University Press, 1962.

Kremer, Michael, Gautam Rao, and Frank Schilbach, "Behavioral Development Economics," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics: Applications and Foundations, Volume 2*, North-Holland, 2019, pp. 345–458.

Lai, Brian and Clayton Thyne, "The Effect of Civil War on Education, 1980–97," *Journal of Peace Research*, 2007, *44* (3), 277–292.

Lary, Diana, *The Chinese People at War: Human Suffering and Social Transformation, 1937–1945*, Cambridge: Cambridge University Press, 2010.

León, Gianmarco, "Civil Conflict and Human Capital Accumulation: Long-term Consequences of Political Violence in Peru," *Journal of Human Resources*, 2012, *47* (4), 991–1022.

Meng, Xin, Nancy Qian, and Pierre Yared, "The Institutional Causes of China's Great Famine, 1959–1961," *Review of Economic Studies*, 2015, *82* (4), 1568–1611.

Miguel, Edward and Gerard Roland, "The Long-Run Impact of Bombing Vietnam," *Journal of Development Economics*, 2011, *96* (1), 1–15.

_ **and Mary Kay Gugerty**, "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 2005, *89* (11–12), 2325–2368.

Ministry of Veterans Affairs of the People's Republic of China, "中华英烈网 (Zhonghua Yinglie Wang; China Martyrs Website)," https://www.chinamartyrs.gov.cn/ 2026. County-level martyr records were scraped and harmonized by the authors.

Mitter, Rana, *Forgotten Ally: China's World War II, 1937–1945*, Boston: Houghton Mifflin Harcourt, 2013.

Nannicini, Tommaso, Andrea Stella, Guido Tabellini, and Ugo Troiano, "Social Capital and Political Accountability," *American Economic Journal: Economic Policy*, 2013, *5* (2), 222–250.

Nunn, Nathan, "Historical Development," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 2, Elsevier, 2014, pp. 347–402.

Ouyang, Difei and Weidi Yuan, "The Intergenerational Transmission of Historical Conflicts: An Application to China's Trade," *Journal of Comparative Economics*, 2021, *49* (3), 675–692.

Pepper, Suzanne, *Radicalism and Education Reform in 20th-Century China*, Cambridge: Cambridge University Press, 1996.

Qian, Nancy, "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance," *Quarterly Journal of Economics*, 2008, *123* (3), 1251–1285.

Shemyakina, Olga, "The Effect of Armed Conflict on Accumulation of Schooling: Results from Tajikistan," *Journal of Development Economics*, 2011, *95* (2), 186–200.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

van de Ven, Hans, *China at War: Triumph and Tragedy in the Emergence of the New China, 1937–1952*, Cambridge, MA: Harvard University Press, 2017.

Voigtländer, Nico and Hans-Joachim Voth, "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," *Quarterly Journal of Economics*, 2012, *127* (3), 1339–1392.

Voors, Maarten J., Eleonora E. M. Nillesen, Erwin H. Bulte, Robert Lensink, Philip Verwimp, and Daan P. Van Soest, "Violent Conflict and Behavior: A Field Experiment in Burundi," *American Economic Review*, 2012, *102* (2), 941–964.

Walder, Andrew G., "Cultural Revolution in China, 1966–1971: Data from County Annals," Harvard Dataverse 2017. Dataset.

Wang, Zheng, "National Humiliation, History Education, and the Politics of Historical Memory: Patriotic Education Campaign in China," *International Studies Quarterly*, 2008, *52* (4), 783–806.

Wei, Lingnan, Thomas Talhelm, Jiaqing Zhu et al., "A Collectivism Index for Investigating Cultural Variation in China across Regions and Time," *Scientific Data*, 2026.

Online Appendix for "War and Human Capital Accumulation: Evidence from the Sino-Japanese War"

Jinyuan LIU

Contents

A. The Martyr Registry and Data Construction

Figure A1. Sample Martyr Profile on the Chinese Martyrs Website

B. Sample and Treatment Descriptives

Table B1. County-Level Treatment Distribution

Figure B1. Distribution of County Martyr Count

Figure B2. Cohort Size by Treatment Status

Figure B3. Martyr Intensity Under Alternative Color Scales

C. Sacrifice-Place Robustness

Table C1. Alternative Treatment Measures (Cutoff 1940)

Table C2. Stricter Cohort Cutoff (Cutoff 1946)

Table C3. Historical County Controls

Table C4. Long March and Korean War as Additional Controls

Table C5. Individual Heterogeneity

Table C6. County Heterogeneity

Table C7. Basic-School Completion

D. Functional-Form Robustness

Table D1. OLS versus Poisson Pseudo-Maximum Likelihood

E. Migration Robustness

Table E1. Baseline DID on Non-Mover Subsample

F. Schooling-Margin Decomposition

Table F1. Schooling-Margin Effects (Cutoff 1940)

Table F2. Schooling-Margin Effects (Cutoff 1946)

Table F3. Schooling-Margin Effects by Sex

Table F4. Schooling-Margin Effects with Historical Controls

Appendix A: The Martyr Registry and Data Construction

The treatment variable in this paper is built from individual-level martyr profiles published on the Chinese Martyrs Website (*Zhonghua Yinglie Wang*, https://www.zhonghuayinglie.cn), the national registry maintained by the Ministry of Veterans Affairs of the People's Republic of China. The registry is the official canonical record of individuals whom the state has formally recognized as martyrs (*lieshi*). Formal recognition is a bureaucratic act by the central government: a service member or civilian killed in a qualifying incident, after review, is entered into the registry, and the entry becomes the legal basis for survivor benefits, commemorative inclusion, and posthumous honors. The registry is therefore not a casualty count. It is the state's own list of losses it has chosen to acknowledge, and in this paper serves as a proxy for the intensity of local loss that the post-war state later sought to redress.

A.1 Registry content and profile structure

Each registry entry is a structured profile page for a single recognized martyr. The fields that are consistently populated across profiles are the martyr's name, gender, date of birth, native place (*jiguan*), former work unit or military affiliation, date of sacrifice, and location of sacrifice. Some entries include a short biographical citation and photograph. Native place is reported as a nested administrative string (province, prefecture, county) that corresponds to the individual's registered household origin. Location of sacrifice is reported at the same hierarchy but refers to the geographic location of the qualifying incident rather than the individual's place of origin. Our baseline treatment uses native place, because the post-war political-recognition and commemorative response operated through the communities that the martyr came from rather than the battlefield where he or she died. The sacrifice-place alternative is used in Appendix C. Figure A1 reproduces a representative profile.

A.2 Crawling workflow

Profiles are served from a paginated JSON API that returns lightweight index pages of profile identifiers. Each identifier resolves to a dedicated profile URL from which the full record can

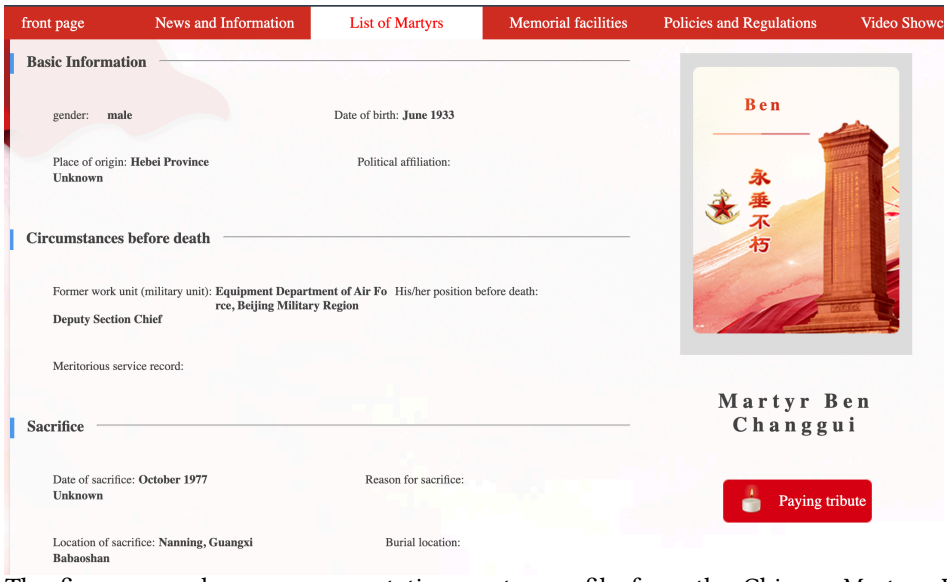
be retrieved. We implemented a scripted crawler that (i) enumerates profile identifiers by paging through the index endpoint, (ii) requests each profile page, (iii) parses the structured fields, and (iv) writes the result to a flat file with one row per martyr. The crawler throttled requests to stay within reasonable load on the server, persisted progress to disk after each page, and retried transient failures with exponential backoff. We ran the crawler until no new identifiers were returned, then compared the recovered count against the site's reported totals to confirm coverage.

A.3 Sample construction and geocoding

We restrict the crawled panel to individuals whose recorded date of sacrifice falls between 1931 and 1945, the Anti-Japanese War window. Restricting on the date of sacrifice rather than the date of birth yields a sample whose deaths are unambiguously attributable to the war rather than to a later conflict. We drop profiles with missing or uninterpretable native-place or sacrifice-place strings, and deduplicate on the pair of name and date of sacrifice to eliminate duplicate entries created by multi-source record ingestion. The remaining profiles are geocoded against a crosswalk of historical and contemporary Chinese administrative names. Each profile's native-place and sacrifice-place strings are matched to a modern county code. Where a string resolves only to the prefecture level, we apply a population-weighted allocation to constituent counties using the 1953 population census. County-level counts are then aggregated into the native-place and sacrifice-place measures used in the analysis.

The final panel covers 2,296 counties. The construction script, the raw crawled fields, and the cleaned county-level counts are available at the project repository.

Figure A1: Sample Martyr Profile on the Chinese Martyrs Website



Notes: The figure reproduces a representative martyr profile from the Chinese Martyrs Website (*Zhonghua Yinglie Wang*), the national registry maintained by the Ministry of Veterans Affairs of the People's Republic of China. Each profile records the martyr's name, gender, date of birth, native place (*jiguan*), former work unit, date of sacrifice, and location of sacrifice. Our analysis uses the date of sacrifice to restrict the sample to deaths between 1931 and 1945, and the native-place string to assign each martyr to a modern county.

Appendix B: Sample and Treatment Descriptives

Table B1 summarizes the county-level treatment distribution, computed once per county across the pooled 1982, 1990, and 2000 census sample. The raw martyr count is highly right-skewed, which motivates the inverse-hyperbolic-sine transformation used in the baseline. The median county records 12 recognized martyrs; the top one percent records more than twenty times that number.

Figure B1 plots the cross-county distribution of the raw martyr count, winsorised at the 99th percentile, with the median marked. The mass is concentrated at the low end, consistent with a small set of heavy-loss counties driving the right tail.

Figure B2 plots the pooled cohort-size distribution by treatment status. The 1920–1956 window is balanced across treatment and control across nearly all cohorts. Sample size rises with birth year because younger cohorts are more likely to be alive and captured in the 1982, 1990, and 2000 censuses, but the treatment-control composition is stable.

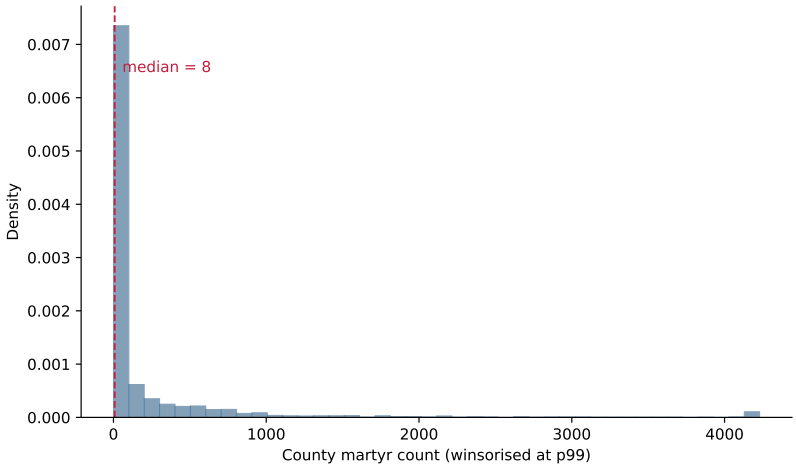
Figure B3 maps the two county-level intensity measures under four color scales. Each row corresponds to one transformation (log, inverse-hyperbolic-sine, raw count winsorised at the 99th percentile, and above-median binary). The left column assigns each martyr to his or her registered native place (*jiguan*); the right column assigns each martyr to the county where the qualifying death occurred. Native-place intensity is concentrated in the central and southern revolutionary-base regions where the Eighth Route Army and New Fourth Army drew most of their recruits. Place-of-sacrifice intensity shifts mass toward the northern and eastern theaters where major battles were fought. The two measures are nearly indistinguishable at the county level (Pearson correlation 0.98 in raw counts), reflecting that most recognized martyrs died within their home province. The right tail dominates the linear and IHS variants, so the log scale offers the cleanest visual summary; that is the scale used in Figure 1 of the main text.

Table B1: County-Level Treatment Distribution (Pooled Sample)

	Mean	SD	p25	p50	p75	p99
Martyr count	241.1	910.4	0	8	115	4,230
IHS(martyrs per 100k)	3.027	2.398	0.942	2.597	5.017	
Above-median dummy	0.482					
Counties				2,640		

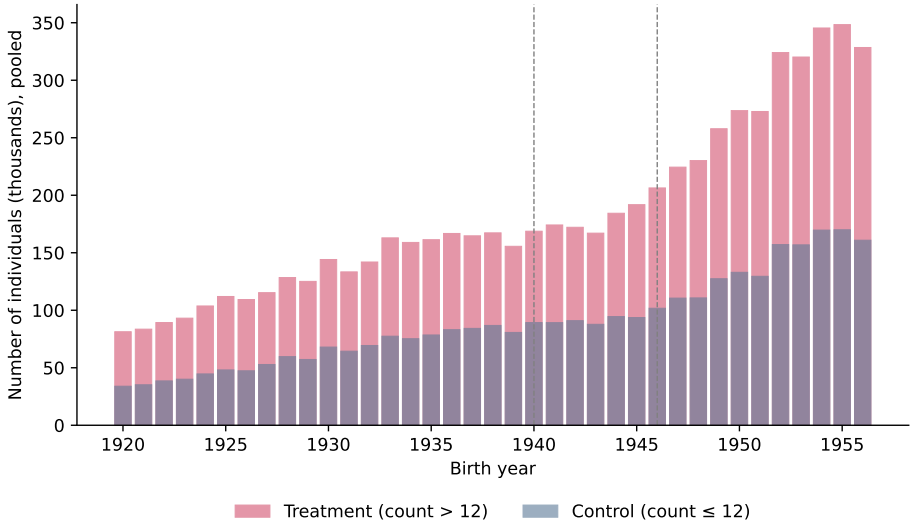
Notes: The unit of observation is a county, taken once across the pooled 1982, 1990, and 2000 census sample. "Martyr count" is the total number of recognized Anti-Japanese War martyrs whose native place (*jiguan*) maps to the county. "IHS(martyrs per 100k)" is the inverse hyperbolic sine of martyr count per 100,000 county residents using the 1953 population as the denominator. "Above-median dummy" is an indicator for a martyr count above the county-level sample median.

Figure B1: Distribution of County Martyr Count (Pooled Sample)



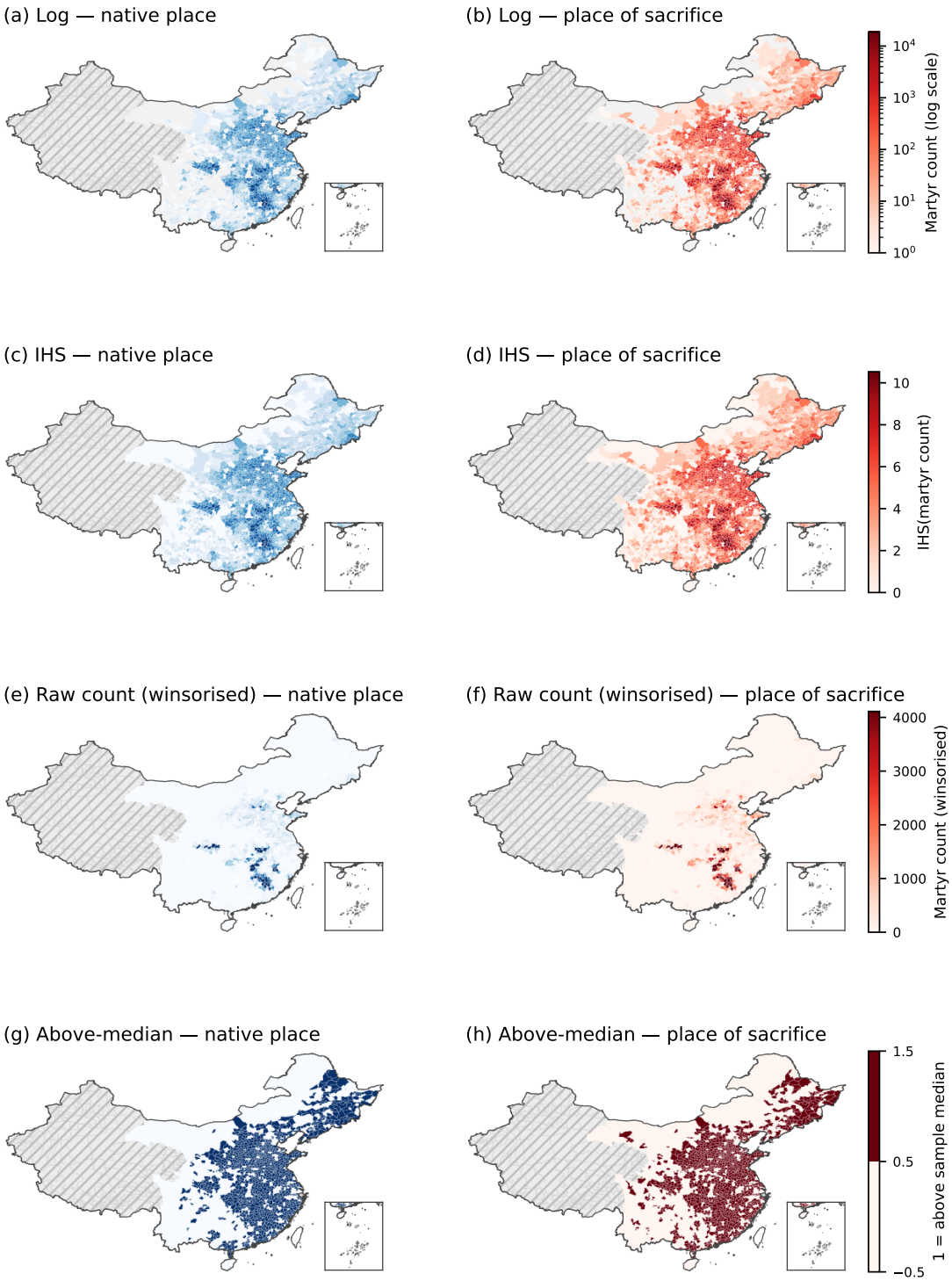
Notes: Cross-county histogram of the raw martyr count across the pooled 1982, 1990, and 2000 census sample. The count is winsorised at the 99th percentile for visual clarity. The dashed vertical line marks the county-level sample median, the threshold used to define the above-median indicator in the main analysis.

Figure B2: Cohort Size by Treatment Status (Pooled Sample)



Notes: Each bar reports the number of individuals in the pooled 1982, 1990, and 2000 census sample at a given birth year, split by above-median treatment status of the county of residence. Vertical dashed lines mark the 1940 and 1946 cohort cutoffs.

Figure B3: Geographic Distribution of Martyr Intensity Under Alternative Color Scales



Notes: County-level counts of recognized Anti-Japanese War martyrs (1931–1945) plotted under four color scales: log (top row, used in Figure 1 of the main text), inverse-hyperbolic-sine, raw count winsorised at the 99th percentile, and above-median binary (bottom row, the alternative treatment). The left column uses native place (*jiguan*) and is shown in blue; the right column uses place of sacrifice and is shown in red. Counties with no recorded martyrs appear in light grey. Xinjiang, Tibet, and Qinghai are hatched to indicate that they are excluded from the analysis. The two geographies are nearly identical at the county level (Pearson correlation 0.98 in raw counts and 0.91 in log(1 + count)). The lower-right inset of each panel reproduces the standard nine-dash line.

Appendix C: Sacrifice-Place Robustness

The baseline treatment in the main text measures wartime loss by a martyr's *native place* (*jiguan*). Native place is the most direct proxy for the community that bore the human cost of the war and that later became the object of post-war political recognition, commemorative investment, and resource claims. Each registry profile, however, also records the county where the martyr died. The two geographies can differ when units were raised locally but deployed elsewhere, or when individuals enlisted away from home and later died in their native county. This appendix asks whether the main results depend on using native place rather than place of sacrifice. To do so, we rebuild the treatment from sacrifice-place counts while holding fixed the 1953 population denominator, cohort window, fixed-effects structure, and clustering scheme used in the baseline. The rest of the appendix re-runs the baseline difference-in-differences, the main robustness checks, the heterogeneity cuts, and the supply-side mechanism regressions with this alternative treatment. Tables C1 through C7 report the results.

Figure 1 (main text) and Figure B3 (Appendix B) map the two intensity measures and motivate the comparison. The robustness regressions below ask whether substituting place of sacrifice for native place changes the substantive conclusions of the main text.

C.1 Baseline Difference-in-Differences

Tables C1 and C2 re-estimate the pooled baseline specification of Section IV under the two cohort cutoffs. Every coefficient on *Post* × sacrifice intensity is positive. The log and cohort-size-controlled specifications remain significant at conventional levels under both cutoffs, and the IHS(count) specification with cohort-size control carries a coefficient of 0.020 that is significant at the one-percent level. The two unconditional binary columns and the IHS(rate) coefficient under the stricter 1946 cutoff attenuate toward insignificance. This is informative rather than troubling. Native place and place of sacrifice overlap substantially at the county level, so the two treatments capture a common underlying wartime shock. At the same time, the post-war rewards emphasized in the main text, including revolutionary-base designation,

veteran benefits, and commemorative infrastructure, were more closely tied to where recruits originated than to the battlefield where they fell. Relative attenuation under the sacrifice-place treatment is therefore what one would expect if native place is the more conceptually precise measure for the post-war resource-response channel.

Table C1: Sacrifice-Place Robustness: Alternative Treatment Measures (Cutoff 1940)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Binary (above median)			Inverse hyperbolic sine			Log
Post × High sacrifice count	0.0238	0.0709**					
	(0.0327)	(0.0323)					
Post × High sacrifice rate			0.0511				
			(0.0326)				
Post × IHS(sac count)				0.0089	0.0200***		
				(0.0075)	(0.0075)		
Post × IHS(sac rate)						0.0146*	
						(0.0082)	
Post × ln(sac per 100k)							0.0158**
							(0.0079)
Cohort size control		Yes			Yes		
Minority control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment start	1940	1940	1940	1940	1940	1940	1940
Observations	6,062,582	6,062,582	5,949,799	6,062,582	6,062,582	5,949,799	5,949,799
R-squared	0.300	0.301	0.300	0.300	0.301	0.300	0.300
Counties	2,044	2,044	2,044	2,044	2,044	2,044	2,044

Notes: Individual-level regressions pooling the 1982, 1990, and 2000 population censuses. The treatment variable now counts martyrs by *place of death* ("sacrifice place") rather than by native place (*jiguan*). The dependent variable is years of education (winsorised at the 1st and 99th percentiles). Columns (1)–(3) use binary treatment: "High sacrifice count" and "High sacrifice rate" are dummies for counties above the sample median sacrifice count and sacrifice-martyrs per 100,000 residents in 1953, respectively. Columns (4)–(6) use inverse hyperbolic sine (IHS) transformations. Column (7) reports the log specification for reference. Cohort size control is ln(county × birth year × wave cell count). All specifications include county fixed effects, birth-cohort fixed effects, wave fixed effects, and a minority-status control. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table C2: Sacrifice-Place Robustness: Stricter Cohort Cutoff (Cutoff 1946)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Binary (above median)			Inverse hyperbolic sine			Log
Post × High sacrifice count	0.0238	0.0513*					
	(0.0301)	(0.0287)					
Post × High sacrifice rate			0.0467				
			(0.0306)				
Post × IHS(sac count)				0.0137**	0.0208***		
				(0.0066)	(0.0064)		
Post × IHS(sac rate)						0.0126	
						(0.0079)	
Post × ln(sac per 100k)							0.0129*
							(0.0076)
Cohort size control		Yes			Yes		
Minority control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment start	1946	1946	1946	1946	1946	1946	1946
Observations	6,031,238	6,031,238	5,919,095	6,031,238	6,031,238	5,919,095	5,919,095
R-squared	0.300	0.300	0.299	0.300	0.300	0.299	0.299
Counties	2,044	2,044	2,044	2,044	2,044	2,044	2,044

Notes: Same specification as Table C1, except that the post-war cohort indicator $Post_i$ equals one for individuals born in 1946 or later. Year 1945 is omitted as the boundary cohort. This cutoff excludes children who may have been exposed to the war in utero or as infants. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

C.2 Robustness: Historical Controls and Placebo Events

Table C3 augments the sacrifice-place specification with the historical county controls used in Table 5. Table C4 reports the Long March and Korean War placebo counterparts of Table 6, again using sacrifice-place intensities for each conflict. The sacrifice-place coefficients remain positive once these controls are introduced, and the placebo coefficients remain small and statistically indistinguishable from zero. The pattern mirrors the native-place evidence in the main text.

Table C3: Sacrifice-Place Robustness: Historical County Controls

	Cutoff 1940		Cutoff 1946	
	No ctrls	+Hist ctrls	No ctrls	+Hist ctrls
Post × IHS(sac rate)	0.0146*	0.0217***	0.0126	0.0188***
	(0.0082)	(0.0076)	(0.0079)	(0.0068)
Post × High sacrifice count	0.0238	0.0855***	0.0238	0.0592**
	(0.0327)	(0.0311)	(0.0301)	(0.0292)
Historical controls	No	Yes	No	Yes
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Observations	5,949,799	5,949,799	5,919,095	5,919,095

Notes: Individual-level regressions pooling the 1982, 1990, and 2000 censuses. The treatment is martyr intensity measured at the sacrifice place. "+ Hist ctrls" augments the baseline with sent-down-youth density, instrumented famine intensity, ln(1 + Cultural Revolution victims), ln(1 + grain output), and urbanization in 1964. Each historical control enters in levels and interacted with *Post*. All specifications include county, birth-cohort, and wave fixed effects, and a minority-status control. Standard errors clustered at the county level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table C4: Sacrifice-Place Robustness: Long March and Korean War as Additional Controls

	(1) Baseline	(2) + Long March	(3) + Korea	(4) + Both
Post × IHS(sac rate)	0.0146*	0.0198**	0.0134	0.0200**
	(0.0082)	(0.0085)	(0.0082)	(0.0084)
Post × IHS(Long March)		-0.1563***		-0.2025***
		(0.0386)		(0.0391)
Post × IHS(Korea)			0.0672***	0.0775***
			(0.0137)	(0.0136)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Observations	5,949,799	5,949,799	5,949,799	5,949,799

Notes: Pooled 1982, 1990, and 2000 censuses; cutoff 1940. Column (1) reports the sacrifice-place IHS(rate) baseline. Columns (2)–(4) add IHS-transformed per-capita martyr counts for the Long March (1934–1936) and Korean War (1950–1953), each interacted with *Post*, as additional controls. The coefficient on the sacrifice-place treatment is stable across specifications; the Long March and Korean War coefficients are not consistently signed or significant, consistent with the main-text finding that the sacrifice effect is specific to the Anti-Japanese War rather than to general military service. Standard errors clustered at the county level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

C.3 Heterogeneity

Tables C5 and C6 re-run the individual-level and county-level heterogeneity cuts of Section IV using the sacrifice-place treatment. The same qualitative pattern reappears. The effect is larger for males, in counties with weaker pre-war educational capacity, in counties without ancestral clans, and in non-treaty-port prefectures. The heterogeneity evidence therefore does not depend on how county-level sacrifice is assigned geographically.

Table C5: Sacrifice-Place Robustness: Individual Heterogeneity

Panel A: Gender				
	Male		Female	
	IHS(rate)	Dummy	IHS(rate)	Dummy
Post × Sacrifice	0.0688***	0.1869***	-0.0168	-0.0964*
	(0.0123)	(0.0454)	(0.0111)	(0.0499)
Observations	1,869,395	1,906,841	1,779,927	1,815,634
Panel B: Pre-war educational base				
	Low base		High base	
	IHS(rate)	Dummy	IHS(rate)	Dummy
Post × Sacrifice	0.0397***	0.1614***	0.0071	-0.0195
	(0.0105)	(0.0490)	(0.0143)	(0.0454)
Observations	3,001,887	3,053,833	2,947,912	3,008,749

Notes: Pooled 1982, 1990, and 2000 censuses; cutoff 1940. Panel A splits the sample by gender. Panel B splits by county-level pre-war educational base, measured as the county-specific mean years of schooling for cohorts born before 1940. "High base" indicates above-median pre-war schooling. Each cell reports the coefficient on *Post*×sacrifice intensity. All specifications include county, birth-cohort, and wave fixed effects, and a minority-status control. Standard errors clustered at the county level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table C6: Sacrifice-Place Robustness: County Heterogeneity

Panel A: Ancestral clan presence				
	No clans		Has clans	
	IHS(rate)	Dummy	IHS(rate)	Dummy
Post × Sacrifice	0.0181	0.0212	0.0191**	0.0636
	(0.0169)	(0.0555)	(0.0093)	(0.0402)
Observations	2,117,441	2,153,199	3,832,358	3,909,383
Panel B: Treaty-port prefecture				
	Non-treaty		Treaty port	
	IHS(rate)	Dummy	IHS(rate)	Dummy
Post × Sacrifice	0.0123	0.0139	0.0301	0.0426
	(0.0089)	(0.0369)	(0.0212)	(0.0697)
Observations	4,589,953	4,686,940	1,359,846	1,375,642

Notes: Pooled 1982, 1990, and 2000 censuses; cutoff 1940. Panel A splits counties by whether any ancestral clan (genealogy) is recorded in Chen et al. (2024). Panel B splits counties by whether the enclosing prefecture contains a late-Qing treaty port. Each cell reports the coefficient on *Post*×sacrifice intensity. All specifications include county, birth-cohort, and wave fixed effects, and a minority-status control. Standard errors clustered at the county level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

C.4 Mechanisms

Table C7 re-runs the supply-side mechanism regressions of Section V with the sacrifice-place treatment: long-run fiscal education spending, primary completion, and junior-high completion. The coefficients retain their positive sign and remain statistically meaningful under the alternative treatment. We do not reproduce the CGSS and CHIP demand-side regressions here because those outcomes are conceptually tied to respondents' county of origin rather than to any battlefield. For those specifications, native place remains the theoretically appropriate treatment.

Table C7: Sacrifice-Place Robustness: Basic-School Completion

	Primary completion (≥ 6 yr)		Junior-high completion (≥ 9 yr)	
	IHS(rate)	Dummy	IHS(rate)	Dummy
Post × Sacrifice	0.0024*	0.0041	0.0006	0.0020
	(0.0012)	(0.0044)	(0.0015)	(0.0063)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Observations	5,949,799	6,062,582	5,949,799	6,062,582

Notes: Pooled 1982, 1990, and 2000 censuses; cutoff 1940. Dependent variables are indicators for completion of primary schooling (≥ 6 years) and junior-high schooling (≥ 9 years). The treatment is martyr intensity measured at the sacrifice place. The pattern mirrors the native-place mechanism results in Table 12 in the main text: gains concentrate at the basic-school margin targeted by post-war educational policy. All specifications include county, birth-cohort, and wave fixed effects, and a minority-status control. Standard errors clustered at the county level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Appendix D: Functional-Form Robustness

Table D1 reports Poisson pseudo-maximum-likelihood (PPML) estimates of the baseline individual-level education specification alongside the OLS benchmark discussed in the main text. Years of education is non-negative and has a substantial mass at low values, so it is useful to verify that the positive baseline result does not depend on a linear functional form. The PPML estimates preserve the sign and significance pattern of the OLS results under both the log and the IHS treatment transformations, and the sign pattern holds within each census wave. The substantive conclusion therefore does not hinge on functional-form assumptions.

Table D1: OLS versus Poisson Pseudo-Maximum Likelihood (PPML)

	IHS(rate)		Above-median dummy	
	OLS	PPML	OLS	PPML
Post × Treatment	0.0338***	0.0146	0.0743	0.0461
	(0.0077)	(0.0038)	(0.0488)	(0.0120)
Minority control	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Observations	7,260,153	7,260,143	7,375,105	7,375,095

Notes: Pooled individual-level regressions of years of education on *Post*×war exposure, using the 1982, 1990, and 2000 censuses. Columns (1) and (3) report OLS (`reghdfejl`); columns (2) and (4) report Poisson pseudo-maximum likelihood (`ppmlhdfe`). Columns (1)–(2) use the IHS of martyrs per 100,000; columns (3)–(4) use a binary indicator for above-median county martyr count. The 1940 cutoff is used throughout and the 1939 boundary cohort is dropped. All specifications include county, birth-cohort, and wave fixed effects, and the minority-status control. Standard errors, clustered at the county level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Appendix E: Migration Robustness

The treatment in this paper is assigned at the county of residence at the time of the census. If individuals born in heavy-loss counties later migrated out of those counties, the treatment effect would attenuate, and if individuals born in low-loss counties later migrated into heavy-loss counties, the treatment effect would be diluted. Either pattern would weaken our coefficients. A more demanding concern is that schooling-related migration could endogenously sort better-educated individuals into specific counties and bias the coefficient. We assess this concern by re-estimating the baseline on the subsample of non-migrants identified through the household-registration variable.

The non-migrant flag is built from the household-registration questions in the 1990 and 2000 censuses. In 1990, *regstatu* reports whether the respondent's hukou is registered in the current location. In 2000, *r061* carries an analogous registered-here indicator. We code an individual as a non-migrant if the respective variable equals one. The 1982 IPUMS file does not release an individual-level non-mover flag, so we treat the entire 1982 wave as in-sample. Across the two waves with the variable, 97.4 percent of pooled observations qualify as non-movers, leaving the non-mover subsample only 0.13 percent smaller than the full sample.

Table E1 re-runs the baseline DID on both samples under both cutoffs. The full-sample coefficients reproduce the main text. The non-mover coefficients are essentially identical: 0.0345 versus 0.0342 under the IHS-rate treatment with the 1940 cutoff, and 0.0298 versus 0.0299 under the 1946 cutoff. The above-median binary coefficients are similarly stable. Migration into and out of high-loss counties does not drive the schooling effect.

Table E1: Baseline DID on Non-Mover Subsample

	Cutoff 1940		Cutoff 1946	
	Full	Non-mover	Full	Non-mover
Panel A: IHS(rate) treatment				
Post × IHS(rate)	0.0342***	0.0345***	0.0299***	0.0298***
	(0.0069)	(0.0070)	(0.0061)	(0.0062)
Panel B: Binary (above median) treatment				
Post × High count	0.0899**	0.0894**	0.0705**	0.0663*
	(0.0419)	(0.0435)	(0.0348)	(0.0370)
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes

Notes: County-cohort-wave panel pooling the 1982, 1990, and 2000 censuses, weighted by cell population. The dependent variable is the within-cell mean of years of completed schooling. "Full" uses every individual; "Non-mover" restricts to individuals whose household registration is local in the relevant census wave (*regstatu*=1 in 1990; *r061*=1 in 2000). The 1982 IPUMS file does not release a non-mover flag, so the 1982 wave appears in both columns. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Appendix F: Schooling-Margin Decomposition

The main results report the effect of wartime loss on years of education. Years of schooling pools several distinct margins: whether a child ever attends school, whether she completes primary, whether she completes junior high. The mechanism evidence in Section V predicts that the gains are concentrated at the lowest rungs, since the early PRC state targeted basic literacy and primary universalisation rather than elite training. This appendix decomposes the years-of-schooling effect at those two margins, then re-runs the decomposition under the same heterogeneity and robustness checks used for the main outcome. We focus on primary completion and junior-high completion. We do not report senior-high or college margins because the mechanism literature predicts no effect at those rungs and the data confirm: wartime loss has no statistically meaningful effect on senior-high completion and a small negative effect on college, neither of which is informative about the basic-schooling channel.

We construct each rung from the categorical schooling variable in each census wave directly, rather than from years-of-schooling thresholds. The 1982 IPUMS variable *educcn* carries verbal labels (illiterate, primary, junior, secondary, some college, graduated) that we decode to the cumulative indicators used here. The 1990 and 2000 raw *educ* codes map cleanly onto the same indicators. We then aggregate to county-cohort-wave cells weighted by cell population, and run the same difference-in-differences specification as in the main text. The dependent variable in each column is the within-cell share whose highest schooling level reached is at least primary or at least junior high.

Table F1 reports the pooled baseline under the 1940 cutoff. The effect of wartime loss is concentrated on the primary margin: a one-unit increase in IHS martyr intensity raises the within-cell share with at least primary schooling by 0.56 percentage points, significant at the one-percent level. The corresponding above-median estimate is 1.4 percentage points. The junior-high estimate is positive but small and statistically indistinguishable from zero. Table F2 repeats the exercise under the stricter 1946 cutoff and reproduces the same pattern. The schooling response to wartime loss is a basic-literacy response. The pattern matches the policy history of the early PRC and supports the supply-side mechanism evidence in Section V.

Table F3 re-runs the baseline separately by sex. The primary-school gain is larger and

more precisely estimated for males than for females, the same pattern as for years of schooling in Table 7. The junior-high estimates are noisier and smaller, with no clean gender contrast. Table F4 adds the historical county controls of Section B (sent-down youth density, Great Leap famine intensity, per-capita Cultural Revolution victims, 1965 grain output, 1964 urbanisation rate), each interacted with the post-war cohort indicator. The primary-margin coefficient is essentially unchanged. No single twentieth-century shock absorbs the basic-literacy effect.

Table F1: Schooling-Margin Effects (Cutoff 1940)

	Primary	Middle
Panel A: IHS(rate) treatment		
Post × IHS(rate)	0.0056***	0.0012
	(0.0009)	(0.0012)
Panel B: Binary (above median) treatment		
Post × High count	0.0142***	0.0059
	(0.0048)	(0.0063)
County FE	Yes	Yes
Cohort FE	Yes	Yes
Wave FE	Yes	Yes
Weights (cell pop.)	Yes	Yes

Notes: County-cohort-wave panel pooling the 1982, 1990, and 2000 censuses, weighted by cell population. The dependent variable in each column is the within-cell share whose highest schooling level reached is at least the indicated rung, constructed directly from the categorical schooling variable in each wave (*educcn* in 1982 IPUMS; raw *educ* in 1990 and 2000). Panel A reports the IHS-rate continuous treatment; Panel B reports the above-median binary treatment. All specifications include county, birth-cohort, and wave fixed effects, and a minority-share control. Standard errors clustered at the county level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table F2: Schooling-Margin Effects (Cutoff 1946)

	Primary	Middle
Panel A: IHS(rate) treatment		
Post × IHS(rate)	0.0050***	-0.0001
	(0.0008)	(0.0011)
Panel B: Binary (above median) treatment		
Post × High count	0.0126***	0.0017
	(0.0042)	(0.0052)
County FE	Yes	Yes
Cohort FE	Yes	Yes
Wave FE	Yes	Yes
Weights (cell pop.)	Yes	Yes

Notes: Same specification as Table F1, except that *Post* equals one for individuals born in 1946 or later and the boundary cohort is 1945. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table F3: Schooling-Margin Effects by Sex

	Primary		Middle	
	Female	Male	Female	Male
Panel A: IHS(rate) treatment				
Post × IHS(rate)	0.0026**	0.0072***	-0.0067***	0.0079***
	(0.0012)	(0.0012)	(0.0017)	(0.0011)
Panel B: Binary (above median) treatment				
Post × High count	0.0061	0.0177***	-0.0171**	0.0242***
	(0.0059)	(0.0052)	(0.0077)	(0.0081)
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes

Notes: The 1940-cutoff specification of Table F1 is estimated separately for females and males, using the sex variable in each census wave. Standard errors clustered at the county level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table F4: Schooling-Margin Effects with Historical County Controls

	Primary		Middle	
	Baseline	With controls	Baseline	With controls
Post × IHS(rate)	0.0056***	0.0027***	0.0012	0.0066***
	(0.0009)	(0.0010)	(0.0012)	(0.0012)
Hist. controls × Post	No	Yes	No	Yes
County FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes

Notes: Historical controls follow the main paper specification: sent-down youth density, Great Leap famine intensity, per-capita Cultural Revolution victims, 1965 grain output, and 1964 urbanisation rate. Each control is interacted with the post-war cohort indicator. The treatment is IHS(martyrs per 100k). Standard errors clustered at the county level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.