

Not Gone With the Wind: Long-Run Impacts of Herbicidal Warfare in Vietnam^{*}

Gaku Ito[†] Duc Tran[‡] Yuichiro Yoshida[§]

August 28, 2023

Abstract

We investigate the legacies of herbicidal warfare on population size and growth in Vietnam in 2001–2020. The identification strategy exploits the flight-level military records and the plausibly exogenous hamlet-level variation in herbicide exposure around the spray-on, direction-change, and spray-off points of fixed-wing aircraft missions, combined with a fuzzy regression kink design. The elasticity estimate suggests that a 5% increase in herbicide exposure in the Vietnam War is associated with a decrease in population *size* of 0.80–0.99% in 2001 and 1.24–1.43% in 2020. We also find negative associations between herbicide exposure and overall and annual population *growth rates*.

Word Count

5,996 words

(excluding the title page, reference list, and exhibits)

Keywords

conflict, development, herbicide, historical legacies, Vietnam War

JEL Classifications

J1, N4, N9, R1

*We thank Masaaki Higashijima, Yoshiaki Kobayashi, Vally Koubi, Adeline Lo, Wakako Maekawa, session participants of the 2022 Annual Meeting of the American Political Science Association, the 2022 Annual Meeting of the Japan Society for International Relations, the 2023 Pacific International Politics Conference, and seminar participants at Hiroshima University, Kyoto University, and University of Toyama for valuable comments and suggestions. Financial support from the Japan Society for the Promotion of Science (JSPS) is gratefully acknowledged (Grant Numbers 18KT0054, 19K21681, and 20K01464). All remaining errors are ours.

[†]Associate Professor, Graduate School of Economics, Osaka Metropolitan University. 3–3–138 Sugimoto, Sumiyoshi, Osaka 558–8585, Japan. Email: gaku@omu.ac.jp. URL: <https://gaku-ito.github.io>.

[‡]Assistant Professor, Hitotsubashi Institute for Advanced Study, Hitotsubashi University. 2–1 Naka, Kunitachi, Tokyo 186–8601, Japan. Email: duc.tran@r.hit-u.ac.jp. URL: <https://ductran-data.github.io>.

[§]Professor, Graduate School of Humanities and Social Sciences, Hiroshima University. 1–5–1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739–8529, Japan. Email: yuichiro@hiroshima-u.ac.jp.

1 Introduction

Does a temporary shock of large-scale violence leave lasting effects on socioeconomic outcomes? In some instances, exposure to violence persistently depletes human capital and undermines interpersonal and institutional trust (e.g., Grasse, 2023; Lichter, Löffler and Siegloch, 2020; Nunn and Wantchekon, 2011). Elsewhere, forms of violence improve political engagement, social cohesion, and altruism (e.g., Bauer et al., 2016; Berman, Clarke and Majed, 2023; Blattman, 2009). Several studies also highlight a rapid recovery or catch-up growth of economic and population outcomes in post-war societies, with the impact of wartime destruction remaining temporary (e.g., Brakman, Garretsen and Schramm, 2004; Davis and Weinstein, 2002, 2008; Miguel and Roland, 2011; but see Bosker et al., 2007).

This article joins the growing literature with new evidence from the herbicide spray in the Vietnam War. While earlier works reveal a powerful recovery from the destruction (e.g., Miguel and Roland, 2011), recent literature finds persistent effects of U.S. bombing on development and health outcomes in Cambodia (e.g., Lin, 2022), Laos (e.g., Riaño and Caicedo, 2021; Yamada and Yamada, 2021), and Vietnam (e.g., Palmer et al., 2019; Singhal, 2019). Most closely related to this article, Appau et al. (2021), Le, Pham and Polacheck (2022), and Yamashita and Trinh (2022) examine the legacies of herbicidal warfare. For example, Appau et al. (2021) underline a persistent negative association between the district-level herbicide exposure and the household agricultural productivity, trust levels, and economic production. In a similar vein, Le, Pham and Polacheck (2022) find a commune-level positive association between herbicide exposure and the immediate and detrimental prevalence of health disease and mobility disability in the cohort born before the spray mission ended in 1971.

While the existing insights are valuable, there remain challenges for causal identification primarily due to the commonly used, district-level or commune-level geographic aggregation. This is unfortunate, as the distribution of herbicide exposure is partly a function of micro-level factors including climate and wind conditions, local geography, and instantaneous decision of aircraft pilots besides initial mission plans. The commonly-employed geographic

aggregation risks masking such micro-level variation that would be valuable for causal identification. An alternative, straightforward empirical strategy is to exploit the micro-level variation with geographically less intensive aggregation, which we adopt in this article.

Here, we combine the fine-grained archival data previously unused at the original, flight-level scale with a fuzzy regression kink (RK) design to address the identification challenge arising from the nonrandom nature of herbicide exposure. Indeed, historical records suggest that Operation Ranch Hand (1962–1971) missions intended to prevent directly damaging densely populated areas and croplands not under Viet Cong control (MACV, 1969), which would introduce bias into naive comparisons. To surmount such identification challenges, our RK strategy leverages the *uncontrollable*, natural experimental variation in the hamlet-level exposure to herbicide around the spray-on, direction-change, and spray-off points of fixed-wing aircraft missions. To capture the micro-level variation in herbicide missions, we rely on the spray flight-level records documented in the Stellman-National Academy of Sciences version of the Herbicide Report System file (S-NAS HERBS, Stellman et al., 2003*a,b*).

The RK analysis reveals lasting legacies of herbicidal warfare. The elasticity estimate suggests that a 5% increase in herbicide exposure, which is observed within a 1 km distance from the spray start, turn, and end points, is associated with a 1.24–1.43% decrease in population size in 2020. The analysis also reveals negative associations between the herbicide shock and overall and annual population *growth rates* in 2001–2020. While the discrepancy might arise from different units of analysis, the previously unseen patterns run counter to the conventional wisdom of a powerful recovery of population from wartime destruction.

2 Data

2.1 Stellman-National Academy of Sciences HERBS File

To measure the landscape of herbicide spray, we rely on the military archival records of the Stellman-National Academy of Sciences version of the Herbicide Report System file (S-NAS

HERBS) database (Stellman et al., 2003*a,b*). The S-NAS HERBS is a compilation of military mission records of the HERBS file originally developed by the U.S. Department of Defense with corrections, containing 9,141 reports of the spray missions with several spray methods (i.e., fixed-wing aircraft, helicopters, and ground-spraying) in 1961–1971. A record in the database corresponds to a single mission with one or more spray paths. Fixed-wing aircraft missions, which dispersed approximately 95% of herbicides (Stellman et al., 2003*a*, 681–682), constitutes a majority of the records, and the following analysis relies on the aircraft missions.¹ Figure 1 shows the geographical and temporal distributions of the mission records.

A noteworthy aspect of the S-NAS HERBS is its geocoding accuracy. Besides mission dates, spray methods, agent, and gallonage information, the database documents “the actual flight paths taken by Ranch Hand aircraft as they carried out their spray missions,” with the “locations at which the aircraft switched directions or turned off and on their spray nozzles” (Stellman et al., 2003*b*, 323), thereby providing a detailed landscape of spray missions.²

2.2 Treatment: Herbicide Exposure

We use hamlets in South Vietnam as the unit of analysis in the following analysis, based on the geocoordinates (points) recorded in the Vietnam Hamlet Evaluation System (HES) Gazetteer Data (Douglass, 2011). To quantify the hamlet-level exposure to herbicide, we broadly follow the distance-weighted approach of Stellman and Stellman (1986, 309) as:

$$\text{HERB}_i = \sum_{j \in \mathcal{H}} G_j \cdot e^{-\lambda D_{ij}}, \quad (1)$$

where i indexes hamlets and j spray flight “legs” of the fixed-wing aircraft missions. Following Stellman et al. (2003*b*, 325), we first split individual flight paths (contiguous lines) into

¹We discard 91 errant entries of “fixed-wing aircraft” records with single (not multiple) geocoordinates.

²To navigate aircraft and record spray-on, direction-change, and spray-off points, the tactical air navigation (TACAN) system distance measuring equipment (DME) was well-developed in Southeast Asia by the 1960s. During mission flights, TACAN/DME continuously offered geographical navigation using the (short) ultra high frequency (UHF) radio range (Rowley, 1975). An essentially same navigation system guided civic aviation until the mid-1990s or before the civic use of the global positioning system (GPS).

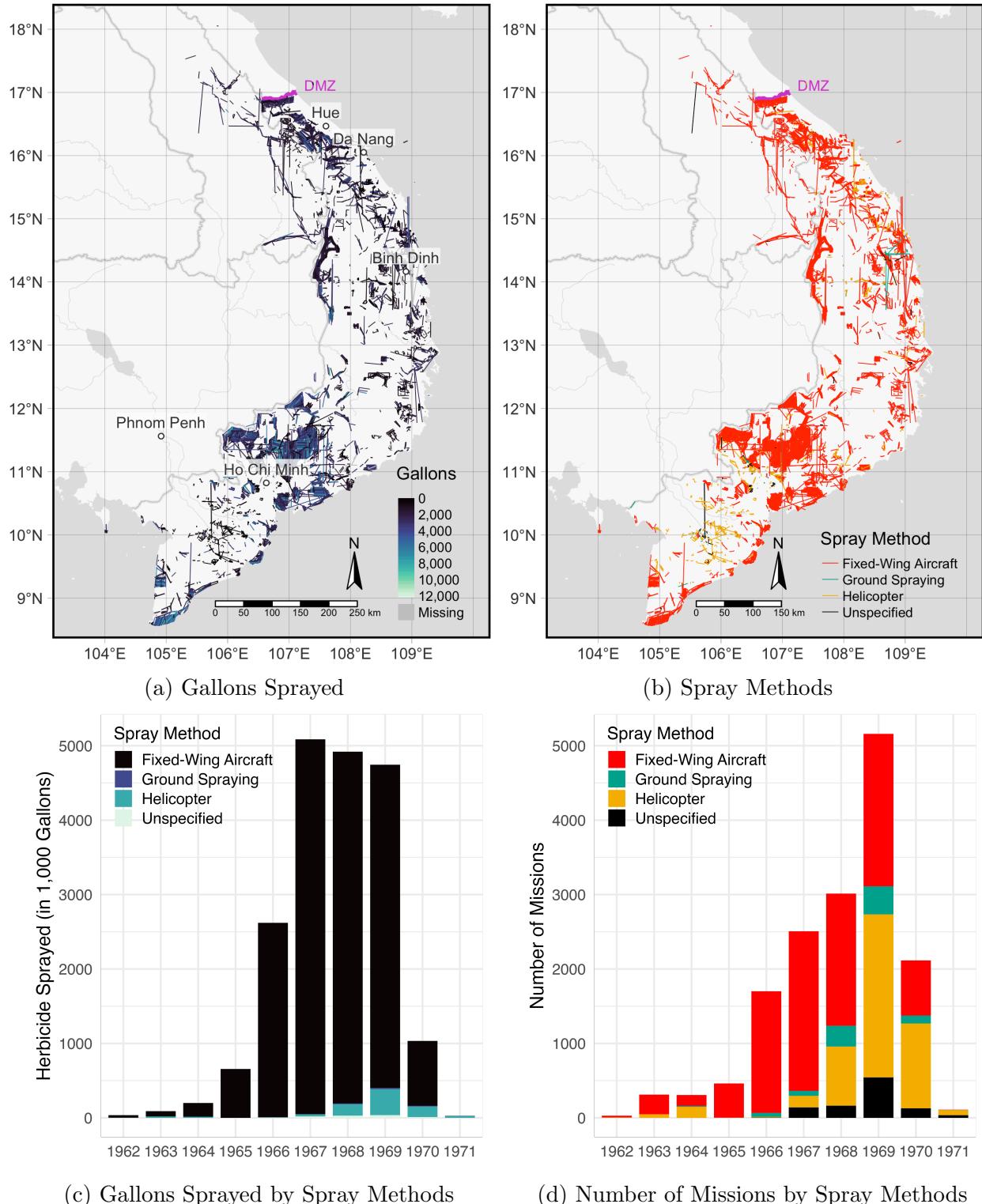


Figure 1: Herbicide Spray Missions

Notes: Segments in maps represent the spray paths. (a) Shading is proportional to the quantity of sprayed herbicide in gallons sprayed by individual spray legs (straight lines without a turn). (b) Colors indicate distinct spray methods. Thick gray lines indicate international borders. Thin gray lines and polygons represent major water bodies. (c), (d) Yearly distribution of herbicide spray (in 1,000 gallons) and spray methods, omitting four helicopter missions in 1961.

distinct flight legs indicated by the straight lines without turns. As the gallonage information is aggregated at the mission level, we divide and assign the total mission-level herbicide quantities to individual legs in proportion to its spray run length.³ G_j indicates the quantity of the herbicides sprayed in leg j in gallons, \mathcal{H} is a set of all spray flight legs, and D_{ij} represents the geodesic distance between hamlet i and flight leg j . $\lambda = \frac{\ln(2)}{D^{\text{Half}}}$ is an arbitrary parameter, where D^{Half} determines the distance at which the herbicide exposure decays to the half of a “direct” (zero-distance) hit with $D_{ij} = 0$, with the baseline of $D^{\text{Half}} = 500$ meters. For a robustness check purpose, Appendix C.2 reestimates the main regressions with alternative half-decay distance parameters of 100 meters, 250 meters, and 1 km.

2.3 Outcome: Population Size and Growth Rate

The following analysis focuses on contemporary population size and population growth rate as outcome variables. We rely on the WorldPop data, which provides geographically disaggregated records of population estimates at 100 meter level on an annual basis in the 2001–2020 period.⁴ To construct the population size variable, we first extract the annual records of population counts at the hamlet locations, and take the natural logarithm of the variable after adding 1. The (logged) overall and annual population growth rates, respectively, are measured as $\ln \text{Population}_{2020} - \ln \text{Population}_{2001}$ and $\ln \text{Population}_t - \ln \text{Population}_{t-1}$.

2.4 Covariates: Key Spray Targets and Related Attributes

To facilitate the empirical analysis, we combine archival sources and several geographical databases to construct three sets of covariates. The first set of covariates measures the proximity to the key spray targets. Operation Ranch Hand involved two primary objectives, forest defoliation and destruction of enemy food supplies, with key targets including

³For example, if a mission includes one flight with four legs (three turns) with an equal leg length sprayed 1,000 gallons in total, we assign 250 gallons to each flight leg.

⁴Available at: <https://www.worldpop.org>, accessed August 6, 2021. Due to errant entries, we exclude the population estimates in 2000 from the analysis.

“base camps and fire support bases...lines of communication, enemy infiltration routes, and enemy base camps” (Institute of Medicine, 1995, 85). The mission authorization process also intended to prevent damage to densely populated areas and crops not under Viet Cong (VC) control (MACV, 1969). To measure proximity to the key spray targets, we measure the prevalence of VC control in 1967–1969 (HES, McCormick, 2021), geodesic distances to suspected areas of North Vietnam Army (NVA) bases (Enemy Base Area File), U.S. Air Force and Navy bases, U.S. Army and Marine troops in 1961–1971 (S-NAS-HERBS), and roads (including trails, *Indochina Atlas*), average hamlet population size in 1967–1969 (HES, Douglass, 2011), and dummy variables for the presence of rice croplands and slash and burn cultivation (*Indochina Atlas*), with the detailed described in Appendix A.4.⁵

Second, geographic covariates include mean elevation and its standard deviation as a proxy of terrain ruggedness (USGS, 1996), soil suitability for rice cultivation (Zabel, Putzenlechner and Mauser, 2014), distance to rivers, flow accumulation (Lehner, Verdin and Jarvis, 2008), a dummy variable for forest presence (*Indochina Atlas*), and average precipitation and wind speed (1970–2000, Fick and Hijmans, 2017). Finally, historical covariates are the number of hamlets within a 30 km radius,⁶ distance to railways (*Indochina Atlas*) and international borders (as of the Vietnam War period), and the average of the annual minimum distance to aerial bombing drop points in 1965–1971 (Defense Digital Service, 2016).

3 Identification Strategy: Fuzzy Regression Kink

To explore the evidence of causal effects, we rely on a fuzzy regression kink (RK) design that exploits the exogenous fluctuations in herbicide exposure within geographically small areas around the spray start, turn, and end points. Recall that the S-NAS-HERBS database

⁵Where multiple hamlet-month observations are available, VC control measures the average of a dummy variable indicating VC control. To measure road proximity and cropland presence, we georeferenced and image-processed the maps of *Indochina Atlas* compiled by the Central Intelligence Agency (CIA) in 1970. For the locations of NVA bases, we rely on the Enemy Base Area File (BASFA), July 1, 1967–July 1, 1971.

⁶The 30 km cutoff reflects the historical facts of Ranch Hand missions and standard error clustering in the regression estimation. See model specification section for details.

that documents the “actual flight paths” and the “locations at which the aircraft switched directions or turned off and on their spray nozzles” (Stellman et al., 2003b, 323).⁷ The core idea behind our RK strategy is that the locations at which aircraft makes turns and turned on and off spray nozzles and the herbicide dispersal were, at least partly, *uncontrollable*. Besides intended targets, the realized distribution of herbicide was driven by plausibly exogenous micro-level factors including climate conditions, wind, terrain, and turbulence from the aircraft as well as ground fire hits (Institute of Medicine, 1995, 86–87). Consistent with the identification idea, “[t]he responsibility for flying the C-123 during the crucial spraying part of each mission was shared between the pilot and the copilot,” with the pilot having “control of the switches which started and stopped the spray” (Buckingham, 1982, 37). As such, instantaneous decision of aircraft pilots and drift due to the disturbing factors jointly determined the spray dispersal, which inevitably deviated from the initial mission plans and generated haphazard, natural-experimental variation in the herbicide distribution.⁸

Consequently, within geographically small areas around the spray start, turn, and end points, the realized herbicide dispersal generates a discontinuous *slope change*, or a discontinuity in the first derivative, in the distribution of herbicide exposure, which remains uncorrelated with potential confounding forces. Our RK strategy leverages this natural experimental kink in the treatment function to derive causal identification. Intuitively, we compare hamlets that were barely covered by spray flights and received direct hits (i.e., “treated” hamlets with greater herbicide exposure due to direct hits) with hamlets that were sufficiently close to but located barely outside of the spray flight paths and received indirect hits (i.e., “control” hamlets with less herbicide exposure due to accidental hits). Located within geographically small areas, these treated and control hamlets should have

⁷See footnote 2 for the aircraft navigation system in the Vietnam War period. Also note that a fuzzy RK design allows incomplete manipulation while addressing measurement errors in the treatment and the running variable (Card et al., 2015b, 2467–2469). Appendix B investigates potential nonlinearity in the covariate distributions across the kink point, and Appendix C examines model dependence and how model specification influences the main findings.

⁸To illustrate, approximately 29% of aircraft sorties were intercepted by ground fire in 1966 (Institute of Medicine, 1995, 86), and the crop damage induced by drift on defoliation missions was greater than the damage by crop destruction missions (National Academy of Sciences, 1974, S-5).

similar (or more precisely, no kink in) geographic, historical, and socioeconomic attributes prior to herbicide exposure, which we empirically validate in the following section.⁹

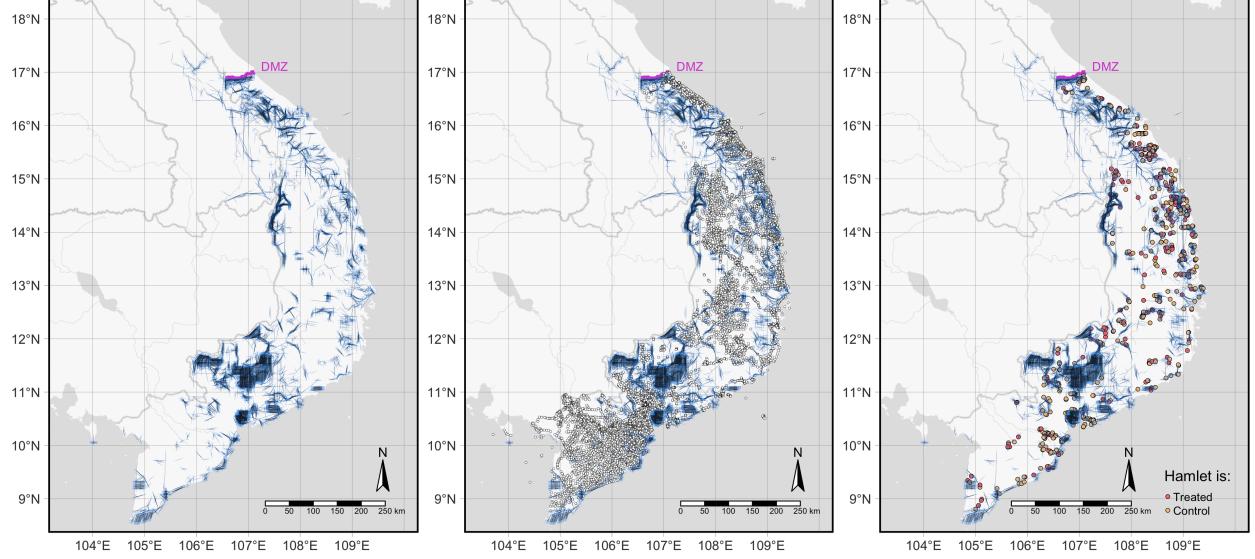
Here, we adopt a RK design rather than a regression discontinuity (RD) design given the probable drifts of herbicide outside the areas with direct hits. We rely on a fuzzy, not sharp, RK design given the unknown slope change parameter of the herbicide exposure (treatment) function at the kink point as well as probable measurement errors in the treatment and the running variable retrieved from archival records (Card et al., 2015*b*, 2464–2467).

3.1 Running Variable

The RK design requires a subset of sample hamlets located within sufficiently small geographic areas around the spray start, turn, and end points. The analysis also requires the corresponding distance measure between hamlet locations and the spray start, turn, and end points as the running variable as well as a dummy variable indicating actual spray hits.

The coding procedure involves several steps. First, we extend the recorded flight paths of fixed-wing aircraft missions from the flight start, turn, and end points by an arbitrary length of 0.05 degree \approx 5.56 km. Second, we add buffers of 0.001 degree \approx 111 meter width to both sides (222 meters) of the extended flight paths, as graphically illustrated by Figure 2(a). The baseline extension length of 5.56 km and the 222 meter buffer width mimic ordinary spray flights in Operation Ranch Hand. A routine fixed-wing aircraft mission involved multiple aircraft and dispersed herbicide at the airspeed of 130–150 knots (240–278 km/h) and an altitude of 150 feet, with each aircraft covering a (laterally contiguous) swath of 80 meter wide and 16 km long (Buckingham, 1982, 37, 132; Institute of Medicine, 1995, 25, 86–87; Stellman et al., 2003*b*, 327). An extension length of 5.56 km approximately corresponds to the “less-than-one-half-minute-away” distance from the observed kink point at the airspeed of 240–278 km/h ($= 4\text{--}4.63 \text{ km/m}$). A 222 meter buffer width similarly approximates the combined swath of a routine spray mission involving three airplanes. This 222 meter or

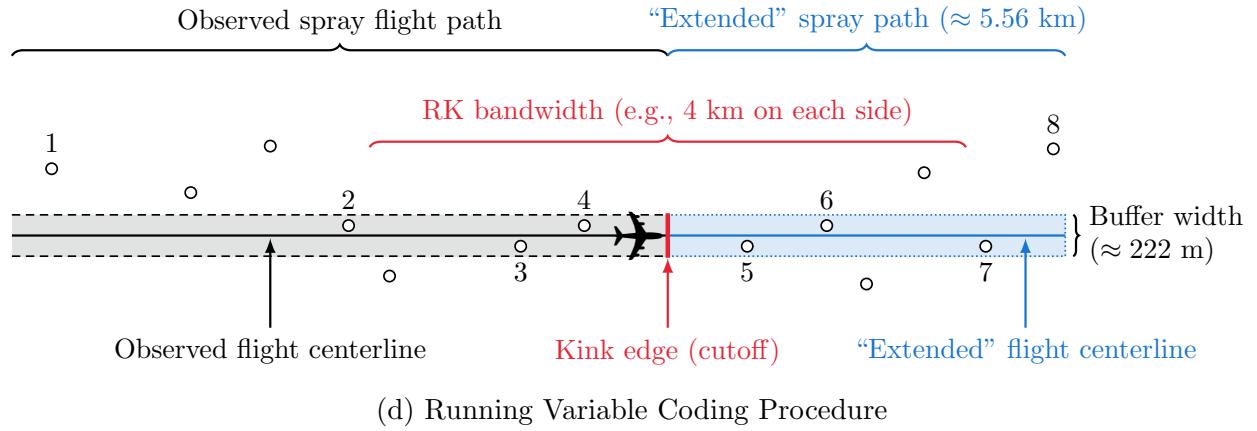
⁹As discussed below, a valid RK design requires that covariate distributions evolve smoothly across the kink point as well as the smoothness of the running variable distribution.



(a) Observed and “Extended” Spray Flight Paths

(b) Flight Paths and Hamlet Locations

(c) RK Sample Hamlets



(d) Running Variable Coding Procedure

Figure 2: Spray Flight Paths, Hamlet Locations, and the RK Running Variable Coding Procedure

Notes: (a) Black polygons represent observed flight paths, and blue polygons indicate the extended flight paths with a 0.05 degree (≈ 5.56 km) distance from the start/turn/end points with a 111 meter buffer on each side (222 meters in total). Of the 14,733 hamlets (Panel (b)), hamlets covered by observed flight polygons with a 222 meter buffer are categorized into the treatment groups (red dots), while the hamlets only covered by a blue polygon are categorized into the control group (yellow dots; Panel (c)). (d) Vertical thick segment indicates a kink edge. Dots and horizontal solid segments, respectively, represent hamlet locations and observed and “extended” spray flight paths. Dashed (dotted) segments and shades represent the buffer around the observed (extended) spray path.

three-aircraft approximation also reflects the historical fact that three C-123 airplanes were assigned to Operation Ranch Hand until 1964, and the number of C-123s increased to 12 in 1965 and then to 36 in 1967 (Institute of Medicine, 1995, 86).

Third, we overlay the extended flight path polygons onto hamlet locations. The hamlets

covered by the flight path polygons without the 5.56 km extension are coded as “treated,” while those only covered by the extended parts are categorized into the control group (Figures 2(b) and (c)). For example, in Figure 2(d), among the hamlets located within the 5.56 km distance from the kink edge, polygon edges corresponding to the spray start, turn, end points, hamlets 2–4 are categorized into the treatment group while hamlets 5–8 are categorized into the control group. We then measure the geodesic distance between hamlet locations and the nearest kink edges to assign running variable to individual hamlets.¹⁰

As in Figure 2(c), the full RK sample includes the hamlets that are geographically covered by the observed or extended flight path polygons with the kink edge distance smaller than 5.56 km extension length. The geoprocessing leaves us a sample of 716 hamlets (4.86% of 14,733 hamlets in panel (b)) in 125 districts (38 provinces) in South Vietnam.

3.2 Model Specification

Our RK estimation builds upon the following two-stage model:

$$\begin{aligned} \ln \text{HERB}_{hd} &= \gamma \text{EdgeDist}_{hd} + \delta \text{EdgeDist}_{hd} \times \mathbb{1}[\text{EdgeDist}_{hd} \geq 0] + \zeta \mathbb{1}[\text{EdgeDist}_{hd} \geq 0] \\ &\quad + \mathbf{X}_{hd}^\top \boldsymbol{\beta} + \eta_d^{\text{District}} + \kappa_{a[h]}^{\text{Agent}} + \theta_{k[h]}^{\text{End-edge}} + \iota_{k[h]}^{\text{Pre-1967}} + f_1(\text{Lon}_{hd}, \text{Lat}_{hd}) + e_{hd}, \quad (2) \\ Y_{hd} &= \tau \ln \widehat{\text{HERB}}_{hd} + \lambda \text{EdgeDist}_{hd} + \nu \mathbb{1}[\text{EdgeDist}_{hd} \geq 0] + \mathbf{X}_{hd}^\top \boldsymbol{\xi} + \pi_d^{\text{District}} + \rho_{a[h]}^{\text{Agent}} \\ &\quad + \phi_{k[h]}^{\text{End-edge}} + \psi_{k[h]}^{\text{Pre-1967}} + f_2(\text{Lon}_{hd}, \text{Lat}_{hd}) + u_{hd}, \text{ if } |\text{EdgeDist}_{hd}| \leq b \quad (3) \end{aligned}$$

where h , d , a , and k , respectively, index hamlets, districts, herbicide agents, and kink edges. HERB denotes the herbicide exposure score, EdgeDist is the geodesic distance from the kink edges in kilometers recentered at zero, and $\mathbb{1}[\text{EdgeDist} \geq 0]$ is a dummy variable which takes one if $\text{EdgeDist} \geq 0$ (treatment group) and zero otherwise (control group).¹¹ \mathbf{X} is

¹⁰The illustration here assumes that single spray flights cover single hamlets. When more than one spray flight covers single hamlets, we modify the coding rule for the treated hamlets not to categorize the hamlets located close to one kink edge (e.g., 100 meters) but far from another edges (e.g., 10 km > 5.56 km threshold) into the treatment group. See Appendix A.3 for details.

¹¹The specification includes $\mathbb{1}[\text{EdgeDist}_{hd} \geq 0]$ and thereby allows discontinuity in the treatment function at the kink point. When the kernel is symmetric (e.g., a uniform kernel), the asymptotic bias and variance

a vector of hamlet-level covariates, and $f_1(\text{Lon}, \text{Lat})$ and $f_2(\text{Lon}, \text{Lat})$ are two-dimensional cubic polynomials of hamlet geocoordinates to screen out spatial trends.¹² η^{District} and π^{District} are (South Vietnam) district fixed effects, and κ and ρ are herbicide agent fixed effects. $\theta^{\text{End-edge}}$ and $\phi^{\text{End-edge}}$ are end-edge fixed effects which take one if a hamlet's running variable is distance to a spray flight end edge and zero otherwise (i.e., a flight start or turn edge); and $\iota^{\text{Pre-1967}}$ and $\psi^{\text{Pre-1967}}$ are fixed effects indicating a hamlet's running variable measured as the distance to the flight leg edges of pre-1967 missions, before the increase of the number of aircraft from 12 to 36 assigned to Operation Ranch Hand mentioned above.¹³ The panel setup replaces the outcome with annual population growth rate and adding a year fixed effect to the right hand side.¹⁴ Following Card et al. (2012, 2015b, 2017), the RK estimation relies on an uniform kernel, and the baseline setup uses a bandwidth of $b = 4$ kilometers (522 hamlet observations). A 4 km bandwidth approximates to the “one-minute away” distance at the typical airspeed of 4–4.63 km/m (= 240–278 km/h) of fixed-wing aircraft missions. To ensure that the fixed bandwidth choice does not drive the results, we replicate the RK estimates with alternative bandwidth sizes.

The parameter of interest is τ in the second-stage (equation 3), which captures the average effect of a marginal increase in ln HERB on the outcome at the cutoff. Formally, τ can be interpreted as the treatment-on-the-treated (TT) effect (Florens et al., 2008) or the local average response (LAR) of herbicide exposure (Altonji and Matzkin, 2005), instrumented by $\text{EdgeDist} \times \mathbb{1}[\text{EdgeDist} \geq 0]$ in the first stage (Card et al., 2015b).

Several aspects of the specification are worth explanations. Following the asymptotic results of Pei et al. (2022), we adopt a local linear specification given the relatively small

of the RK estimand are not affected by the continuity imposition (Card et al., 2012, 2015b).

¹²Linear, quadratic, and fifth-order polynomials of longitude and latitude yield qualitatively similar results.

¹³The herbicide agent fixed effect has five categories: “Blue,” “Orange,” “Purple,” “White,” and “Others.” Given the small numbers of observations, we categorize “Pink” and “Unknown” into “Others.”

¹⁴With abuse of notations, the panel version of the second-stage model is:

$$Y_{hdt} = \tau \ln \widehat{\text{HERB}}_{hd} + \lambda \text{EdgeDist}_{hd} + \nu \mathbb{1}[\text{EdgeDist}_{hd} \geq 0] + \mathbf{X}_{hd}^\top \boldsymbol{\xi} + \pi_d^{\text{District}} + \rho_{a[h]}^{\text{Agent}} \\ + \phi_{k[h]}^{\text{End-edge}} + \psi_{k[h]}^{\text{Pre-1967}} + \eta_t + f_2(\text{Lon}_{hd}, \text{Lat}_{hd}) + e_{hdt},$$

where η_t denotes year fixed effect, with the corresponding first-stage model specified analogously.

sample size. We also rely on district, agent, end-edge, and pre-1967 mission fixed effects instead of the ideal flight or leg fixed effects due to limited sample size. Nonetheless, the specification reflects the historical fact that South Vietnam Government both at the national and local levels and the U.S. jointly controlled the herbicide missions (Buckingham, 1982, 36–38; Institute of Medicine, 1995, 86).¹⁵ District fixed effects subsume the regional variation in the authorization process and other district-level differences of counterinsurgency strategies; and the remaining fixed effects absorb the types of chemical shocks (agent fixed effects), aircraft directions (end-edge fixed effects), and overall intensity of herbicide missions and potential military significance of nearby targets (pre-1967 fixed effects).

Throughout the analysis, we rely on the two-stage least-squares (2SLS) estimator. To account for the increased error due to the two-stage estimation and potential error dependence across space, we report Conley’s (1999) standard errors robust to spatial clustering with a 30 km cutoff.¹⁶ Appendix A reports descriptive statistics of the variables.

3.3 Identification Assumption

A valid fuzzy RK design hinges on the smoothness assumption, which yields two testable implications: First, the density of the running variable is sufficiently smooth, or continuously differentiable at the cutoff; and second, predetermined covariates evolve smoothly around the kink point (Card et al., 2012, 2015*a*, 2017).

Following literature (Bana, Bedard and Rossin-Slater, 2020; Card et al., 2015*a,b*; Landais, 2015), we validate the smoothness assumption with the running variable and covariate distributions in three ways. First, Figure B.1 in the Appendix examines the continuity of the running variable distribution using the polynomial estimator of Cattaneo, Jansson and Ma

¹⁵The authorization process also involved the U.S. Ambassador, the U.S. Military Assistance Command, and the Corps Tactical Zones (CTZs). Dell and Querubin (2018) use the CTZ (Corps I-II) boundary as one of the sources for causal identification. As the CTZ boundaries follow the province boundaries, district fixed effects subsume the difference in counterinsurgency strategy across CTZs and local governments.

¹⁶We use a 30 km cutoff to reflect the combat range of aircraft missions in which a C-123 airplane covered a swath of 80 m wide and 16 km long (Buckingham, 1982, 132). A 30 km cutoff approximately reflects the 16-km range as a radius. The sample mean (median) of the aircraft spray legs is 18 km (16.48 km).

(2020), and fails to detect statistically significant discontinuity at the kink point ($t = -0.209$ and $p = 0.835$). Second, we test for the kink in the running variable distribution using polynomial regressions.¹⁷ As reported in Figure B.2 in the Appendix, the polynomial regressions fail to detect a slope change in the running variable distribution around the kink point.

Finally, Figure B.3 in the Appendix presents a series of placebo kink estimates with the covariates, *including* hamlet population in 1967–1969, as the left-hand-side variable using a reduced-form version of the RK specification. Again, and consistent with the smoothness assumption, most of the placebo regressions does not reveal a discernible slope change at the kink point. Exceptions are the substantively small but statistically significant slope change in wind speed and the distance to U.S. bases plausibly arising from random chance or sampling error. Indeed, as reported in Appendix B.2, the same randomization inference exercise introduced in the next section fails to negate that the covariate kink arises by chance. Moreover, Appendix C.1 finds little evidence of model dependence, suggesting the minor role of the marginally significant covariate kink in influencing the RK estimates.

4 Results

Figure 3 displays the distributions of herbicide exposure (first-stage association; panel (a)) and population size in 2020 (reduced-form association; panel (b)), given hamlet geocoordinates and fixed effects.¹⁸ As the fuzzy RK estimand can be written as the ratio of the reduced-form and first-stage associations, the co-evolving slope change provides a graphical

¹⁷Formally, and with abuse of notation, we first aggregate the hamlet observations using bins with different sizes based on the running variable and then estimate the following regression model:

$$N_b^{\text{obs.}} = \beta \mathbb{1}[\overline{\text{EdgeDist}}_b \geq 0] + \sum_{p=1}^P \left[\gamma_p \overline{\text{EdgeDist}}_b^p + \delta_p \overline{\text{EdgeDist}}_b^p \cdot \mathbb{1}[\overline{\text{EdgeDist}}_b \geq 0] \right] + e_b,$$

where b indexes bins, $N_b^{\text{obs.}}$ reflects the number of observations in each bin, $\overline{\text{EdgeDist}}$ is the midpoint of EdgeDist of each bin, and P is the polynomial order. δ_1 captures the slope change of the probability density function of the running variable at the kink point. Not to violate the smoothness assumption, δ_1 should remain indistinguishable from zero.

¹⁸The RK estimand, τ_{RK} , can be written as the ratio of the reduced-form slope change (i.e., slope change in the outcome function) relative to the first-stage slope change (i.e., slope change in the treatment function)

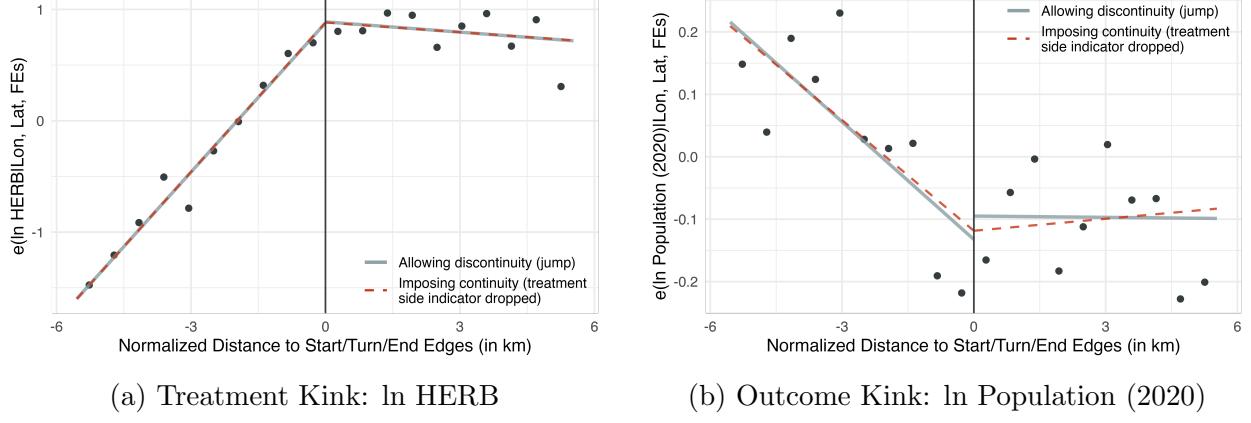


Figure 3: Slope Change in the Treatment and Outcome Functions at the Kink Point

Notes: Dots represent the binned frequencies, and solid and dashed lines show linear regression fits adjusting for longitude, latitude, and agent and district fixed effects. Solid line allows discontinuity (as equation 2), while the dashed line imposes continuity at the kink point by dropping the treatment-side dummy variable. The bin size is selected by the integrated mean squared error (MSE)-optimal mimicking variance evenly-spaced method using spacing estimators (Calonico, Cattaneo and Titiunik, 2015).

but initial empirical evidence of herbicide legacies on contemporary population outcome. In the following, we first report the cross-sectional RK estimates with population size as the outcome. We then extend the analysis with alternative growth rate and panel specifications across different bandwidth sizes, followed by a brief summary of robustness checks.

4.1 Population Size Estimates

Table 1 reports the RK estimation results with population size in 2001 (Panel A) and 2020 (Panel B) as the outcome, along with the first-stage estimates. Model (1) is the baseline model without additional controls besides the two-dimensional polynomial of longitude and latitude and fixed effects. Models (2) to (4) consecutively add the proximity to the key targets, geographical characteristics, and historical attributes as covariates. To address potential spillover effects, Model (5) adjusts for the spatially-lagged treatment with the 30

at the kink point as (Card et al., 2012, 2015b):

$$\tau_{\text{RK}} = \frac{\frac{\lim_{v_0 \uparrow 0} d(\mathbb{E}[Y|V=v])}{dv} \Big|_{v=v_0} - \frac{\lim_{v_0 \uparrow 0} d(\mathbb{E}[Y|V=v])}{dv} \Big|_{v=v_0}}{\frac{\lim_{v_0 \uparrow 0} d(\mathbb{E}[D|V=v])}{dv} \Big|_{v=v_0} - \frac{\lim_{v_0 \uparrow 0} d(\mathbb{E}[D|V=v])}{dv} \Big|_{v=v_0}},$$

where V denotes the running variable, v_0 the kink point, and D the treatment.

km neighbor cutoff in addition to the full set of covariates in model (4). Note that given the limitation in data availability and the historical facts, some of the covariates can partly be posttreatment (e.g., bombing point distance, 1965–1971). The models with covariate adjustments, therefore, do not necessarily provide conservative estimates due to possible posttreatment bias of unknown directions and should be interpreted with caution.

Table 1 confirms the graphical evidence of Figure 3 and underlines the herbicide legacies. Across model specifications, the first-stage, treatment kink estimates are substantively and statistically significant (Panel A). The second-stage coefficients on $\ln \text{HERB}$ are also consistently signed negative and retain the statistical significance at the conventional 5% level (Panels B and C). Note also that uninstrumented ordinary least squares (OLS) estimates reported in Appendix Table A.2 substantively underestimate or even fail to reveal the negative association, suggesting that nonrandom herbicide assignment invites bias into naive comparisons. Moreover, the coefficient stability suggests that the estimates are unlikely to be driven by an arbitrary covariate adjustment choices, which is of a particular concern in RK applications (Ando, 2017) and we further investigate in Appendix C.1. The negative association also remains visible across the outcomes, population size in 2001 and 2020.

As an initial robustness check, Table 1 also reports three randomization inference results.¹⁹ First, an important concern for RK applications is the misspecification of underlying nonlinearity such that one falsely specifies a quadratic relationship with no kink as a discontinuous slope change (Ganong and Jäger, 2018). To guard against the misspecification bias, Table 1 reports kink point (KP) randomization p -value obtained from the reduced-form permutation test (Ganong and Jäger, 2018).²⁰ For the exercise, we first generate 10,000 placebo kink points randomly drawn from a uniform distribution, $\mathcal{U}(-4 \text{ km}, 4 \text{ km})$. We then consecutively estimate the reduced-form version of the RK model with the subsample within

¹⁹Randomization exercise relies on the reduced-form and first-stage models as randomization inference cannot be applied without additional assumptions to unobserved subgroups such as compilers.

²⁰If our RK specification correctly captures a discontinuous slope change, we would be unlikely to see a slope change with placebo kink points, and thus have a small KP randomization p -value. If, on the other hand, we falsely specify a quadratic function with no kink as a discontinuous slope change, we would observe discernible kink estimates with placebo kink points with a large KP randomization p -value.

Table 1: Herbicide Exposure and Population Size in 2001 and 2020

	Panel A: ln HERB				
	(1)	(2)	(3)	(4)	(5)
First stage					
EdgeDist $\times \mathbb{1}[\text{EdgeDist} \geq 0]$	-0.693*** (0.095)	-0.660*** (0.089)	-0.629*** (0.087)	-0.633*** (0.086)	-0.632*** (0.084)
F-statistic (weak instrument)	53.132	54.585	51.893	54.03	56.504
Adjusted R ²	0.620	0.652	0.663	0.666	0.677
Running variable (RV) randomization <i>p</i> -value	0.000	0.000	0.000	0.000	0.000
Spatial noise (SN) randomization <i>p</i> -value	0.000	0.000	0.000	0.000	0.000
Panel B: ln Population (2001)					
	(1)	(2)	(3)	(4)	(5)
Second stage					
ln HERB	-0.203*** (0.069)	-0.177*** (0.065)	-0.168*** (0.063)	-0.165*** (0.063)	-0.165*** (0.064)
5% treatment increase effect size	-0.99%	-0.86%	-0.82%	-0.80%	-0.81%
Average outcome	1.315	1.315	1.315	1.315	1.315
Reduced form					
Kink point (KP) randomization <i>p</i> -value	0.019	0.021	0.004	0.003	0.000
RV randomization <i>p</i> -value	0.003	0.003	0.006	0.007	0.007
SN randomization <i>p</i> -value	0.004	0.004	0.005	0.006	0.006
Panel C: ln Population (2020)					
	(1)	(2)	(3)	(4)	(5)
Second stage					
ln HERB	-0.294*** (0.076)	-0.264*** (0.072)	-0.258*** (0.072)	-0.254*** (0.072)	-0.255*** (0.072)
5% treatment increase effect size	-1.43%	-1.29%	-1.26%	-1.24%	-1.24%
Average outcome	1.467	1.467	1.467	1.467	1.467
Reduced form					
KP randomization <i>p</i> -value	0.013	0.007	0.008	0.009	0.004
RV randomization <i>p</i> -value	0.000	0.000	0.000	0.000	0.001
SN randomization <i>p</i> -value	0.000	0.000	0.000	0.000	0.000
Observations	522	522	522	522	522
Avg. N neighbors (Conley SE cluster size)	27.4	27.4	27.4	27.4	27.4
Key target covariates	✓	✓	✓	✓	✓
Geographic covariates		✓	✓	✓	✓
Historical covariates			✓	✓	✓
ln spatially-lagged HERB				✓	✓
Fixed effects and <i>f</i> (Lon, Lat)	✓	✓	✓	✓	✓

Notes: **p* < 0.1; ***p* < 0.05; ****p* < 0.01. Conley (1999) standard errors adjusted for spatial clustering with a 30 km cutoff and a Bartlett kernel are in parentheses. Key target covariates: NVA base distance, population (1967–1969), U.S. base distance, U.S. troop distance, rice cropland, road distance, slash and burn cropland, Viet Cong control prevalence. Geographic covariates: Precipitation, wind speed, elevation, flow accumulation, forest presence, rice suitability, river distance, ruggedness. Historical covariates: Bombing point distance, border distance, number of neighbor hamlets, railway distance. ln spatially-lagged HERB is the logged average HERB of the neighbor hamlets with a 30 km cutoff. Fixed effects: Agent fixed effect, district fixed effect, end-edge fixed effect, pre-1967 mission fixed effect. Randomization inference: KP randomization *p*-value is computed by the permutation test of Ganong and Jäger (2018) with 10,000 placebo kink points drawn from uniform distribution $\mathcal{U}(-4 \text{ km}, 4 \text{ km})$. RV randomization *p*-value is obtained by randomly assigning the running variable to the sample hamlets for 10,000 times. SN randomization *p*-value is obtained by 10,000 synthetic spatial noise simulations of Kelly (2021) with the outcome replaced by randomly generated spatial noise.

baseline 4 km bandwidth around each placebo kink point to obtain the empirical distribution of placebo estimates. The two-sided p -value is computed by doubling the minimum of the fraction of placebo estimates not smaller or greater than the actual reduced-form coefficient. Second, another intuitive approach of randomization inference is to randomly assign the running variable (RV) and the corresponding treatment-side indicator variable to the sample hamlets while holding the kink point fixed (e.g., Dell and Querubin, 2018, 30–33; see also, Cattaneo, Frandsen and Titiunik, 2015). The RV randomization two-sided p -values report the share of the absolute 10,000 placebo coefficients that are larger than the absolute actual reduced-form kink coefficient. Finally, to address potential spatial curve-fitting and residual spatial autocorrelation, we use the spatial noise (SN) randomization inference procedure developed by Kelly (2021). The SN randomization inference procedure replaces the outcome by synthetic noise with the same estimated spatial structure of the observed outcome partialled out by covariates. If the estimation is not an artifact of spatial trends, the observed variable should not explain spatial noise. The SN randomization p -value indicates the fraction of 10,000 spatial noise simulations which yield absolute t values greater than the absolute t value estimated with observed data.

For both treatment and outcomes, the randomization inference results are consistent with asymptotic inference. The three series of randomization p -values remain smaller than the conventional 5% threshold, suggesting that the actual estimates are unlikely to arise from misspecification of the underlying nonlinearity (KP randomization), random chance (RV randomization), or spatial autocorrelation and curve-fitting (SN randomization).

Turning to magnitude, the coefficient estimates can readily be interpreted as elasticity given the log-log specification. The estimates suggest that, measured at the 100 meter grid scale, a 5% increase in herbicide exposure is persistently followed by a 1.24% ($-0.254 \times \ln(1.05) = 0.0124$, Model 4, with all covariates) to 1.43% (Model 1, without covariates) decrease in population size in 2020 (Panel B), and a 0.80% to 0.99% decrease in 2001 (Models 1 and 4 in Panel C). Note that the 5% increase reference does not overstate

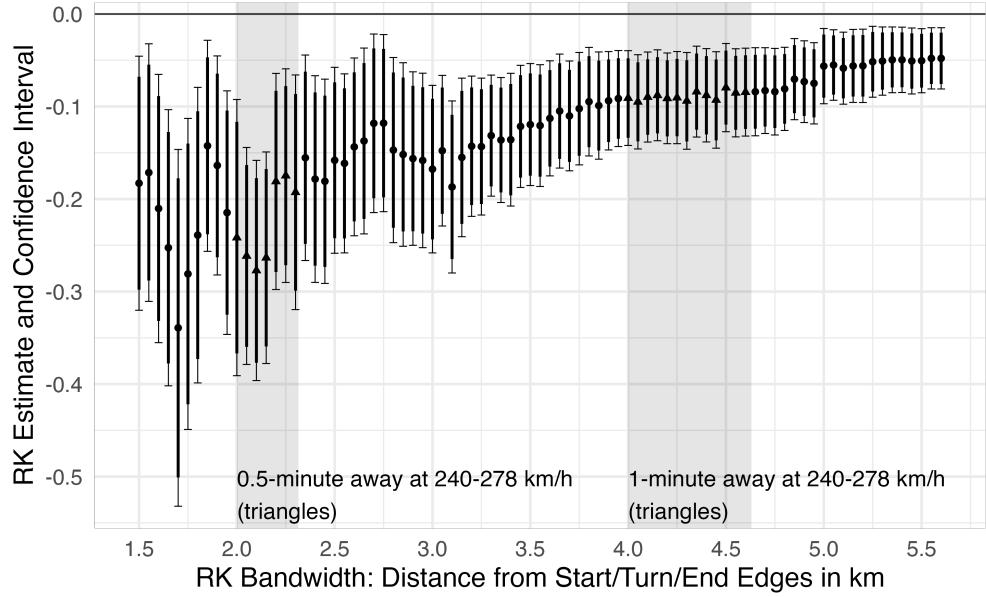
the magnitude of the treatment effect, as such an increase in herbicide exposure is observed within a geographically small area around the kink edges. Specifically, a 5% increase in the treatment approximately corresponds to the difference in $\ln \text{HERB}$ between the average of the control hamlets within a 1 km distance from the kink edge with the average of the treated hamlets within the same distance from the kink edge ($\frac{8.22}{7.802} = 1.054$).

Besides the baseline population growth during the 2001–2020 period (from 1.315 to 1.467 in the logarithm scale, Table 1), the growing, rather than fading away, coefficient size implies that herbicide legacies shape not only population size but also population growth rate. This leads to a testable implication that a similar negative association is also present in local population growth rate, which contrasts the conventional wisdom of rapid recovery of population from wartime destruction.

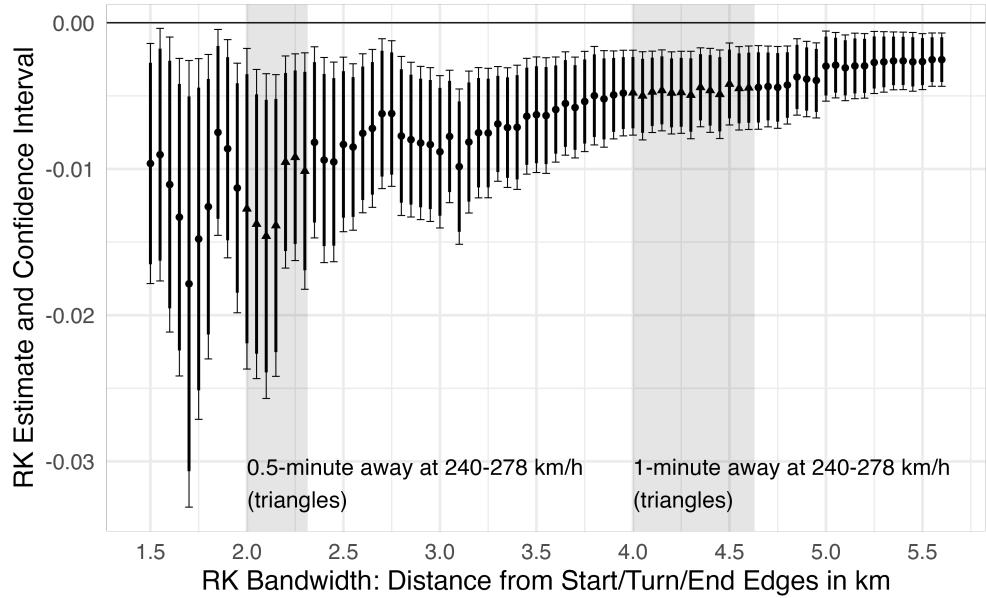
4.2 Population Growth Rate Estimates

To examine the testable implication of cumulatively increasing legacies of herbicide exposure, we reestimate the RK models with the overall and annual population growth rates during the 2001–2020 period as the outcomes, using both cross-sectional and panel setups. This growth-rate specification also surmounts the robustness concern that a small number of volatile and less persistent population count observations in 2001 and 2020 drive the estimation results. For the panel specification, we extend the baseline cross-sectional specification by replacing the outcome by the annual growth rate and adding a year fixed effect as an additional control.²¹ In the panel setup, we report robust standard errors clustered at the hamlet level. Another robustness issue remaining in Table 1 is that the results are based on a fixed bandwidth of 4 km, leaving potential concern for the bandwidth choice driving the results. To examine the herbicide legacies on population growth rate and address these robustness concerns, we report the RK estimates for overall and annual population growth rates across different RK bandwidths ranging from 1.5 km to 5.6 km with an increment of 50 meters.

²¹The results remain qualitatively unchanged when further adjusting for the lagged outcome, $Y_{hd,t-1}$.



(a) Overall Population Growth Rate, 2001–2020



(b) Annual Population Growth Rate, 2002–2020

Figure 4: Herbicide Exposure and Population Growth Rate

Notes: Symbols and thin (thick) vertical segments with horizontal ticks represent the RK estimates and the corresponding 95% (90%) confidence intervals based on Conley (1999) standard errors with spatial clustering (Panel (a)) and robust standard errors with hamlet-level clustering (Panel (b)). The model specification in Panel (a) follows model (1) in Table 1. Panel (b) adds a year fixed effect to the baseline specification. Grey shades represent “0.5-minute-away” (2–2.315 km) and “one-minute-away” distances (4–4.63 km) from the kink point at the typical airspeed of Ranch Hand aircraft of 240–278 km/h, and triangles indicate the corresponding point estimates. Annual growth estimate in Panel (b) covers the 2002–2020 period as the outcome is defined as $\ln \text{Population}_t - \ln \text{Population}_{t-1}$ and our dataset covers the 2001–2020 period.

Figure 4 presents the RK estimates for overall (Panel (a)) and annual (Panel (b)) population growth rates across bandwidth sizes, painting a picture that is consistent with the population size results. Regardless of the bandwidth settings and outcomes, the growth rate estimates reveal a negative association consistent with the population size estimates.²² The effect size of the persistent association is also discernible given the average outcome of 0.151 (or $e^{0.151} = 1.163$, 16.3% overall population growth) and the geographically small level of outcome measurement of 100 meter grid. The cross-sectional RK estimate of $\tau = -0.091$ with the baseline bandwidth of 4 km (a “one-minute-away” distance) suggests that a 5% increase in herbicide exposure translates into a 0.44% ($-0.091 \times \ln(1.05) = -0.0044$) decrease in the overall (logged) population growth rate in 2001–2020 (Figure 4(a)). Given $\ln \text{Population}_{2020} - \ln \text{Population}_{2001} \approx \frac{\text{Population}_{2020}}{\text{Population}_{2001}} - 1$, a 0.44% decrease can roughly be interpreted as the effect size in the percentage point scale. The negative association is also visible in the panel specification with the annual population growth rate as the outcome (Figure 4(b)), revealing the herbicide legacies remaining in the dynamics as well as the snapshots of population outcomes in Vietnam in the present day.

4.3 Robustness Check and Sensitivity Analysis

Robustness and sensitivity concerns remaining in the main estimates include (1) the robustness to the parameter choice to construct the RK sample and the exposure score, HERB, (2) confounding kink in observed covariates and model dependence, and (3) unobserved confounding kink. Appendix C addresses these robustness concerns by leveraging alternative combinations of flight buffer width and half-decay distance parameters, all possible 2^N covariates $= 2^{20} = 1,048,576$ model specifications per outcome, a jackknife approach and the known historical differences in counterinsurgency strategies across four U.S. Corps Tac-

²²Figure 4 follows the specification of model (1) in Table 1 (without covariate adjustments). Note that, as reported in Appendix C.1, the growth rate specification is remarkably robust to different model specifications (different combinations of adjusted covariates) and exhibits little model dependence. The empirical distribution of the RK estimates of all possible 2^N covariates $= 2^{20} = 1,048,576$ model specifications is almost normally distributed with the mean and median estimates of 0.091 (Appendix Figure C.1), which is identical to the point estimate reported in the text and Figure 4(a) ($\hat{\tau} = 0.091$).

tical Zones in South Vietnam (see, e.g., Dell and Querubin, 2018), and a sensitivity analysis approach to quantify how severe unobserved confounding forces would need to be to eliminate the main estimates (Cinelli and Hazlett, 2020). Reassuringly, none of the robustness checks yield results that would invalidate or overturn the main findings.

5 Conclusion

This article combined the historical records of herbicide missions with a fuzzy RK design, and revealed the negative legacies of herbicidal warfare on contemporary population size and growth rate in Vietnam. The empirical analysis suggests that the temporary shock of herbicidal warfare left lasting, rather than temporary, effects on contemporary population outcomes. While the reported results tell us little about the underlying mechanisms, there are at least three distinct, although not mutually exclusive, explanations for the revealed persistent association, which also opens up pathways for future research. First, the negative association might reflect decreased birthrates and life expectancy. The reported negative association between herbicide exposure and health outcomes (e.g., Le, Pham and Polachek, 2022) makes the heavily sprayed areas persistently suffer lower birthrates and decreased life expectancy. Second, and relatedly, the lower population size and growth rates can arise from deteriorated agricultural productivity to sustain local population (e.g., Appau et al., 2021). Finally, domestic migration preferences and patterns can also shape population dynamics. Increased out-migration from the herbicide affected areas coupled with decreased in-migration into the damaged areas can also account for the lasting negative associations.

References

- Altonji, Joseph G., and Rosa L. Matzkin.** 2005. “Cross section and panel data estimators for nonseparable models with endogenous regressors.” *Econometrica*, 73(4): 1053–1102.
- Ando, Michihito.** 2017. “How much should we trust regression-kink-design estimates?” *Empirical Economics*, 53(3): 1287–1322.
- Appau, S, S A Churchill, R Smyth, and T A Trinh.** 2021. “The long-term impact of the Vietnam War on agricultural productivity.” *World Development*, 146(October): 105613.
- Bana, Sarah H., Kelly Bedard, and Maya Rossin-Slater.** 2020. “The Impacts of Paid Family Leave Benefits: Regression Kink Evidence from California Administrative Data.” *Journal of Policy Analysis and Management*, 39(4): 888–929.
- Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts.** 2016. “Can War Foster Cooperation?” *Journal of Economic Perspectives*, 30(3): 249–274.
- Berman, Chantal, Killian Clarke, and Rima Majed.** 2023. “From Victims to Dissidents: Legacies of Violence and Popular Mobilization in Iraq (2003-2018).” *American Political Science Review*, forthcoming.
- Blattman, Christopher.** 2009. “From Violence to Voting: War and Political Participation in Uganda.” *American Political Science Review*, 103(2): 231–247.
- Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm.** 2007. “Looking for multiple equilibria when geography matters: German city growth and the WWII shock.” *Journal of Urban Economics*, 61(1): 152–169.
- Brakman, Steven, Harry Garretsen, and Marc Schramm.** 2004. “The strategic bombing of German cities during World War II and its impact on city growth.” *Journal of Economic Geography*, 4(2): 201–218.
- Buckingham, William A.** 1982. *Operation Ranch Hand: The Air Force and Herbicides in Southeast Asia, 1961-1971*. Washington, DC: Office of Air Force History.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik.** 2015. “Optimal Data-Driven Regression Discontinuity Plots.” *Journal of the American Statistical Association*, 110(512): 1753–1769.
- Card, David, Andrew Johnston, Pauline Leung, Alexandre Mas, and Zhuan Pei.** 2015a. “The effect of unemployment benefits on the duration of unemployment insurance receipt: New evidence from a regression kink design in Missouri, 2003-2013.” *American Economic Review*, 105(5): 126–130.
- Card, David, David Lee, Zhuan Pei, and Andrea Weber.** 2012. “Nonlinear Policy

Rules and the Identification and Estimation of Causal Effects in a Generalized Regression Kink Design.”

- Card, David, David S. Lee, Zhuan Pei, and Andrea Weber.** 2015b. “Inference on Causal Effects in a Generalized Regression Kink Design.” *Econometrica*, 83(6): 2453–2483.
- Card, David, David S. Lee, Zhuan Pei, and Andrea Weber.** 2017. “Regression Kink Design: Theory and Practice.” In *Regression Discontinuity Designs: Theory and Applications*. , ed. Matias D. Cattaneo and Juan Carlos Escanciano. Bingley: Emerald Publishing.
- Cattaneo, Matias D., Brigham R. Frandsen, and Rocío Titiunik.** 2015. “Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate.” *Journal of Causal Inference*, 3(1): 1–24.
- Cattaneo, Matias D., Michael Jansson, and Xinwei Ma.** 2020. “Simple Local Polynomial Density Estimators.” *Journal of the American Statistical Association*, 115(531): 1449–1455.
- Cinelli, Carlos, and Chad Hazlett.** 2020. “Making sense of sensitivity: Extending omitted variable bias.” *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 82(1): 39–67.
- Conley, Timothy G.** 1999. “GMM estimation with cross sectional dependence.” *Journal of Econometrics*, 92(1): 1–45.
- Davis, Donald R., and David E. Weinstein.** 2002. “Bones, Bombs, and Break Points: The Geography of Economic Activity.” *American Economic Review*, 92(5): 1269–1289.
- Davis, Donald R., and David E. Weinstein.** 2008. “A search for multiple equilibria in urban industrial structure.” *Journal of Regional Science*, 48(1): 29–65.
- Defense Digital Service.** 2016. “Theater History of Operations Data (THOR): Vietnam War.” Available at: <https://data.world/datamil/vietnam-war-thor-data>.
- Dell, Melissa, and Pablo Querubin.** 2018. “Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies.” *Quarterly Journal of Economics*, 133(2): 701–764.
- Douglass, Rex W.** 2011. “Data of the Vietnam War v1.0 (Vietnam Hamlet Evaluation System Gazetteer Data).” Available at: <https://esoc.princeton.edu/data/vietnam-hamlet-evaluation-system-gazetteer-data>, accessed July 7, 2021.
- Fick, Stephen E., and Robert J. Hijmans.** 2017. “WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas.” *International Journal of Climatology*, 37(12): 4302–4315.
- Florens, J.P., J.J. Heckman, C Meghir, and E. Vytlacil.** 2008. “Identification of Treatment Effects Using Control Functions in Models With Continuous, Endogenous

- Treatment and Heterogeneous Effects.” *Econometrica*, 76(5): 1191–1206.
- Ganong, Peter, and Simon Jäger.** 2018. “A Permutation Test for the Regression Kink Design.” *Journal of the American Statistical Association*, 113(522): 494–504.
- Grasse, Donald.** 2023. “State Terror and Long-Run Development: The Persistence of the Khmer Rouge.” *American Political Science Review*, forthcoming.
- Institute of Medicine (Committee to Review the Health Effects in Vietnam Veterans of Exposure to Herbicides).** 1995. *Veterans and Agent Orange: Health Effects of Herbicides Used in Vietnam*. Washington, D.C.: National Academy Press.
- Kelly, Morgan.** 2021. “Persistence, Randomization, and Spatial Noise.” *CEPR Discussion Paper*, DP16609.
- Landais, Camille.** 2015. “Assessing the welfare effects of unemployment benefits using the regression kink design.” *American Economic Journal: Economic Policy*, 7(4): 243–278.
- Le, Duong Trung, Thanh Minh Pham, and Solomon Polacheck.** 2022. “The long-term health impact of Agent Orange: Evidence from the Vietnam War.” *World Development*, 155: 105813.
- Lehner, Bernhard, Kristine Verdin, and Andy Jarvis.** 2008. “New global hydrography derived from spaceborne elevation data.” *Eos*, 89(10): 93–94.
- Lichter, Andreas, Max Löffler, and Sebastian Siegloch.** 2020. “The Long-Term Costs of Government Surveillance: Insights from Stasi Spying in East Germany.” *Journal of the European Economic Association*, 19(2): 741–789.
- Lin, Erin.** 2022. “How War Changes Land: Soil Fertility, Unexploded Bombs, and the Underdevelopment of Cambodia.” *American Journal of Political Science*, 66(1): 222–237.
- McCormick, John T.** 2021. “The Hamlet Evaluation System Reevaluated.” Available at: https://github.com/jtmccorm/HES_Reevaluated, accessed July 7, 2021.
- Miguel, Edward, and Gérard Roland.** 2011. “The long-run impact of bombing Vietnam.” *Journal of Development Economics*, 96(1): 1–15.
- Military Assistance Command, Vietnam (MACV).** 1969. “Military Operations: Herbicide Operations (U).” *APO San Francisco 96222, Accession Number: AD0779793*. Available at: <https://apps.dtic.mil/sti/citations/AD0779793>, accessed July 10, 2023.
- National Academy of Sciences.** 1974. *The Effects of Herbicides in South Vietnam: Part A Summary and Conclusions*. Washington, DC: National Academy of Sciences.
- Nunn, Nathan, and Leonard Wantchekon.** 2011. “The slave trade and the origins of Mistrust in Africa.” *American Economic Review*, 101(7): 3221–3252.
- Palmer, Michael, Cuong Viet Nguyen, Sophie Mitra, Daniel Mont, and Nora Ellen Groce.** 2019. “Long-lasting consequences of war on disability.” *Journal of Peace Research*, 56(6): 860–875.

- Pei, Zhuan, David S. Lee, David Card, and Andrea Weber.** 2022. “Local Polynomial Order in Regression Discontinuity Designs.” *Journal of Business and Economic Statistics*, 40(3): 1259–1267.
- Riaño, Juan Felipe, and Felipe Valencia Caicedo.** 2021. “Collateral Damage: The Legacy of the Secret War in Laos.” *CEPR Discussion Paper*, DP15349.
- Rowley, Ralph A. (Office of Air Force History).** 1975. “The Air Force in Southeast Asia: FAC Operations, 1965–1970.” *Office of Air Force History, May 1975, available at: https://media.defense.gov/2011/Mar/24/2001330117/-1/-1/0/AFD-110324-008.pdf*.
- Singhal, Saurabh.** 2019. “Early life shocks and mental health: The long-term effect of war in Vietnam.” *Journal of Development Economics*, 141(March): 102244.
- Stellman, Jeanne Mager, Steven D. Stellman, Richard Christian, Tracy Weber, and Carrie Tomasallo.** 2003a. “The extent and patterns of usage of agent orange and other herbicides in Vietnam.” *Nature*, 422(6933): 681–687.
- Stellman, Jeanne Mager, Steven D. Stellman, Tracy Weber, Carrie Tomasallo, Andrew B. Stellman, and Richard Christian.** 2003b. “A geographic information system for characterizing exposure to Agent Orange and other herbicides in Vietnam.” *Environmental Health Perspectives*, 111(3): 321–328.
- Stellman, Steven D., and Jeanne M. Stellman.** 1986. “Estimation of exposure to agent orange and other defoliants among american troops in vietnam: A methodological approach.” *American Journal of Industrial Medicine*, 9(4): 305–321.
- U.S. Geological Survey (USGS).** 1996. “Global 30-Arc-Second Elevation Data, GTOPO30.”
- Yamada, Takahiro, and Hiroyuki Yamada.** 2021. “The long-term causal effect of U.S. bombing missions on economic development: Evidence from the Ho Chi Minh Trail and Xieng Khouang Province in Lao P.D.R.” *Journal of Development Economics*, 150(October): 102611.
- Yamashita, Nobuaki, and Trong Anh Trinh.** 2022. “Long-Term Effects of Vietnam War: Agent Orange and the Health of Vietnamese People After 30 Years.” *Asian Economic Journal*, 36(2): 180–202.
- Zabel, Florian, Birgitta Putzenlechner, and Wolfram Mauser.** 2014. “Global agricultural land resources: A high resolution suitability evaluation and its perspectives until 2100 under climate change conditions.” *PLoS ONE*, 9(9): 1–12.