

Mass Repression and Political Loyalty: Evidence from Stalin's 'Terror by Hunger'

ONLINE APPENDICES

Arturas Rozenas
New York University
arturas.rozenas@nyu.edu

Yuri Zhukov
University of Michigan
zhukov@umich.edu

CONTENTS

| | | |
|-----|--|----|
| 1 | Rayon-level data | 2 |
| 1.1 | Descriptive Statistics. | 2 |
| 1.2 | Geographic Matching | 4 |
| 1.3 | Coding of Parties in 2002. | 7 |
| 2 | Survey Design | 8 |
| 2.1 | Sample Selection | 8 |
| 2.2 | Questionnaire | 10 |
| 2.3 | Territorial Control. | 11 |
| 3 | Selecting Optimal Instruments | 13 |
| 3.1 | Manually Selected Instruments | 13 |
| 3.2 | Bayesian Model Selection | 14 |
| 3.3 | LASSO-Based Instruments | 15 |
| 3.4 | Bayesian Model Averaging | 15 |
| 3.5 | Estimation Workflow | 16 |
| 4 | Weather and Famine: Additional Evidence | 17 |
| 4.1 | First-Stage Results. | 17 |
| 4.2 | Qualitative Evidence | 19 |
| 5 | Attribution of Famine Deaths and Repression. | 20 |
| 5.1 | Attribution of Famine Deaths | 20 |
| 5.2 | Exposure to Acts of Repression | 22 |
| 6 | Additional IV Results | 24 |
| 6.1 | Full Second-Stage Output | 24 |
| 6.2 | Results without normalized outcomes | 25 |

| | | |
|------|--|----|
| 7 | Robustness Checks for IV Results | 26 |
| 7.1 | Alternative Specifications | 26 |
| 7.2 | Alternative Instruments | 27 |
| 7.3 | Alternative Measures of WWII Partisan Behavior | 29 |
| 7.4 | Measurement Error in the Instrument Construction | 30 |
| 7.5 | Sensitivity Check | 32 |
| 8 | A Placebo Test: Western Ukraine | 33 |
| 9 | Robustness Checks for Survey Results | 36 |
| 9.1 | Alternative Measures of DNR/LNR Exposure | 36 |
| 9.2 | Violations of the Linearity Assumption. | 37 |
| 9.3 | Alternative Estimators and Individual Level Demographics | 39 |
| 9.4 | Sample Attrition | 41 |
| 10 | Disaggregated second-stage results | 45 |
| 11 | Evidence from battlefield behavior in WWII | 46 |
| 12 | Alternative Explanations for Soviet Elections | 48 |
| 12.1 | Propaganda | 48 |
| 12.2 | Electoral Mobilization | 49 |
| 12.3 | Election Fraud | 50 |
| 12.4 | Famine-Relief and Post-Famine Economic Investments. | 50 |
| 13 | Alternative Explanations for Post-Soviet Period | 53 |
| 13.1 | Propaganda and Mobilization. | 53 |
| 13.2 | Post-Famine Ethnic Composition | 57 |
| 14 | Mechanisms of Persistence | 61 |
| 14.1 | Bias Due to Misreporting of Family Exposure | 61 |

1. RAYON-LEVEL DATA

1.1. Descriptive Statistics

Table 1.1 reports summary statistics for the community-level variables we use in the main analysis. Units of analysis are 1933-era rayons in Soviet Ukraine.

Table 1.1: Summary statistics for the main variables.

| Statistic | N | Mean | St. Dev. | Min | Max |
|---------------------------------------|-----|--------|----------|-------|--------|
| Famine mortality (percent) | 386 | 11.37 | 5.97 | 1.60 | 32.82 |
| Famine mortality (percent logged) | 386 | 2.26 | 0.63 | 0.47 | 3.49 |
| Opposition to Red partisans | 386 | -2.24 | 0.85 | -6.07 | -0.54 |
| Anti-Soviet votes (percent) | 386 | 0.30 | 0.15 | 0.03 | 0.81 |
| Anti-Soviet protests (count) | 386 | 1.35 | 10.44 | 0 | 188 |
| Anti-Russian votes (percentage) | 386 | 65.64 | 19.10 | 12.80 | 88.63 |
| Anti-Russian protests (count) | 386 | 19.69 | 115.01 | 0 | 1,812 |
| Ukrainians in 1926 (proportion) | 386 | 0.82 | 0.15 | 0.23 | 0.99 |
| Russians in 1926 (proportion) | 386 | 0.07 | 0.11 | 0.002 | 0.65 |
| Rural population in 1926 (proportion) | 386 | 0.88 | 0.14 | 0.19 | 1.00 |
| Forestation | 386 | 0.08 | 0.15 | 0.00 | 1.00 |
| Manufacturing | 386 | 0.39 | 0.35 | 0.00 | 1.00 |
| Dominant crop: dairy | 386 | 0.08 | 0.28 | 0 | 1 |
| Dominant crop: potato | 386 | 0.12 | 0.33 | 0 | 1 |
| Dominant crop: wheat | 386 | 0.60 | 0.49 | 0 | 1 |
| Population in 1926 (log) | 386 | 11.06 | 0.57 | 8.71 | 13.34 |
| Distance to Russian border (km) | 386 | 276.71 | 190.47 | 0.00 | 690.98 |

1.2. Geographic Matching

Since our data come from various time periods when different administrative structures were in place, considerable care had to be exercised to match these data to the common unit of analysis, the rayon borders in 1933, during the time of famine. Some of the matching had to be conducted manually with the help of the specialists from Institute for Demography and Social Studies (IDSS) in Ukraine, and some of it was conducted using automated geographic matching.

For the 1926 census data (from which we use pre-famine population, and distribution of nationalities with the rayon, as well as proportion of rural vs urban population), the following administrative-territorial structure was in place at the time of the census: six economic regions (Poliska, Pravoberezhna, Liboberezhna, Stepova, Girnichopromislova, Dniprovska); each region was divided into okruhy (40 okruhy and the Moldova Autonomous Soviet Socialist Republic (ASSR), and each okruh divided into rayons and its urban rada (capital of the okruha), with a total of 624 rayons. The administrative-territorial structure in 1939 consisted of 15 oblasts and the Moldova ASSR, with each oblast divided into rayons, with a total of 504 rayons.

We first recalculated the 1926 data to the 1939 (as the 1939 structure followed closer 1933 rayon borders) administrative structure according to the following steps. First, we separated the 1926 rayons with the same name and borders in 1939. Next, we listed the rayons that changed names but kept the same borders. The remaining 1926 rayons were redistributed into existing 1939 rayons. This was done by comparing maps dated January 1, 1927 and January 15, 1939.

As a first step of this redistribution, we used calculations performed by the Central Statistical Administration of the Ukrainian SSR¹, where urban and rural populations in 1926 were recalculated separately into the rayon structure of 1939. Comparing populations according this redistribution with the respective population in 1926, it was found that for the rural population in 90% of rayons in 1939 they did not match the respective populations in 1926. This was due to the fact that in most cases borders of rayons changed by additions or subtractions of territory. These differences were corrected by a painstaking process estimating land proportions of these changes by comparing the two maps and estimating respective populations that had to be added or subtracted.

Other datasets were geographically matched using automated methods. Figure 1.1 illustrates our geographic matching and spatial aggregation procedure, where raw data are (a) points, like polling station locations, and (b) polygons or grid cells, like weather.

¹CSA USSR (Central Statistical Administration of the USSR [Tsentralnoie statisticheskoe upravlenie SSSR]).

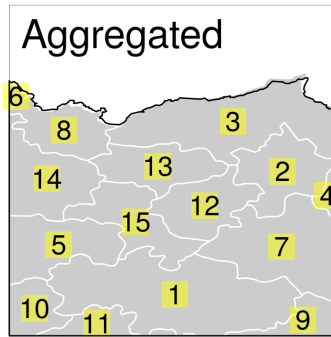
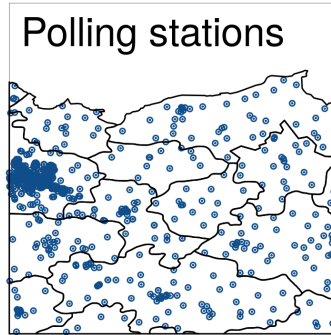
The upper left overlays raw data with administrative units of interest (here, 1933-era rayons around Kharkiv oblast). The lower left displays aggregated measures for each unit, and the right pane shows the same in tabular form.

For points (Figure 1.1a), we identify administrative units that contain each event, and generate local summary statistics (e.g. average ‘anti-Russian’ vote share, protest event counts) at each time interval.² Because polygon borders do not always align (Figure 1.1b), we aggregate such data with area-weighted means. Formally, area-weighted means are $x_i = \sum_{j_i} w_{j_i} x_{j_i}$, where x is the variable of interest (e.g. monthly rainfall), $i \in \{1, \dots, N\}$ indexes destination spatial units (e.g. 1933 rayons), $j \in \{1, \dots, J\}$ indexes source units (e.g. grid cell), and J_i is the subset of source units that overlap with i . Weights $w_{j_i} = \frac{A_{i \cap j}}{A_i}$ represent the proportion of i 's area (A_i) covered by the intersection of i and j ($A_{i \cap j}$).

²If the original source did not include geographic coordinates, but only an address (e.g. name of city or village), we use the latter information to geo-code events to the most precise location possible, with a suite of mapping APIs from Google, Yandex and MapQuest.

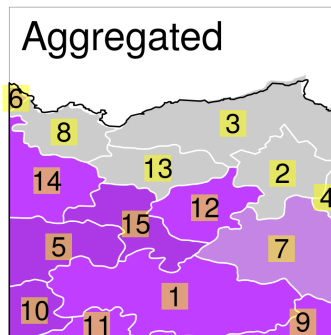
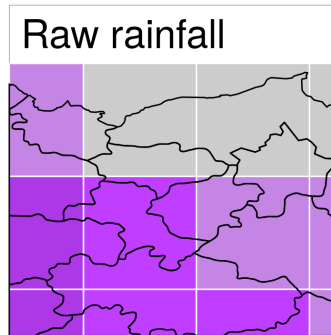
Figure 1.1: SPATIAL AGGREGATION

(a) Point-to-polygon aggregation



| | RAYON (1933) | VOTE |
|----|----------------------|-------|
| 1 | Балаклійський | 24.52 |
| 2 | Велико-Бурлуцький | 26.64 |
| 3 | Вовчанський | 25.35 |
| 4 | Двурічанський | 22.68 |
| 5 | Зміївський | 30.18 |
| 6 | Золочівський | 28.12 |
| 7 | Куп'янський | 25.84 |
| 8 | Липецький | 22.83 |
| 9 | Ізюмський | 26.6 |
| 10 | Олексіївський | 24.8 |
| 11 | Петрівський | 21.04 |
| 12 | Печенізький | 31.53 |
| 13 | Старо-Салтівський | 31.02 |
| 14 | Харківська міськрада | 43.88 |
| 15 | Чугуївський | 22.02 |

(b) Polygon-to-polygon aggregation



| | RAYON (1933) | RAIN |
|----|----------------------|-------|
| 1 | Балаклійський | 8.75 |
| 2 | Велико-Бурлуцький | 6.52 |
| 3 | Вовчанський | 5.82 |
| 4 | Двурічанський | 6.87 |
| 5 | Зміївський | 10.29 |
| 6 | Золочівський | 7.33 |
| 7 | Куп'янський | 7.54 |
| 8 | Липецький | 6.74 |
| 9 | Ізюмський | 8.13 |
| 10 | Олексіївський | 9.4 |
| 11 | Петрівський | 8.84 |
| 12 | Печенізький | 8.25 |
| 13 | Старо-Салтівський | 6.57 |
| 14 | Харківська міськрада | 8.75 |
| 15 | Чугуївський | 9.06 |

1.3. *Coding of Parties in 2002*

The significance of the 2004 Orange Revolution as a turning point in Ukraine's political institutions and culture has been a matter of some debate (Kuzio, 2005; Katchanovski, 2008). For our purposes, however, the events of 2004 – mass protests culminating in the nullification of fraudulent presidential election results, and a re-vote – marked the beginning of a ten-year period in which Ukrainian electoral politics crystallized more clearly along pro-Russian/pro-Western lines (i.e. Party of Regions, Opposition Bloc and Communists on the pro-Russian side, and Our Ukraine, Yulia Tymoshenko Bloc, Petro Poroshenko Bloc and various nationalist parties on the pro-western side). For 2002 elections, we code the following list of parties as 'pro-Russian,' based on their programmatic announcements and expressed support for Yanukovich or Yuschchenko:

- Communist Party of Ukraine
- Democratic Union
- Blok Natalii Vitrenko
- Communist Party of Ukraine (renewed)
- Party block 'For United Ukraine'
- Peasant Party of Ukraine
- Social-Democratic party of Ukraine (United)
- Party block 'ZUBR (For Ukraine, Belarus, Russia)
- 'Russian Block'

The Communist Party of Ukraine and the bloc 'For United Ukraine' (predecessor of the later 'Party of Regions') account for most of the votes in this category of parties.

2. SURVEY DESIGN

Our individual-level data are from an original survey in eastern Ukraine (Donets'ka, Luhans'ka, and Kharkivs'ka oblasts), implemented by Center "Social Indicators" in collaboration with Kyiv International Institute of Sociology. The survey's principal investigators were Alexander Kupatadze (King's College, London), Thomas Zeitzoff (American University), and Yuri Zhukov (University of Michigan). The field stage lasted from September 1 to October 8, 2017, following a pilot test in May 21 – June 11, 2017.

The pilot test included 200 pretest interviews, including 100 in the city of Kharkiv and 100 in the city of Severodonetsk, Luhans'ka oblast. The field stage comprised 1246 face-to-face interviews by locally-based enumerators, with a response rate of 39.7%, across 100 primary sampling units (PSU's) (settlements and city districts) in Donetska, Luhanska, and Kharkivska oblasts. Following the field stage, the survey firms conducted quality control tests of 55% of all interviews, including analysis of enumerators' diaries and verification of the fact and quality of interviews in each household. Survey responses that did not pass quality control tests (15 total) were dropped, and replaced with new interviews.

2.1. *Sample Selection*

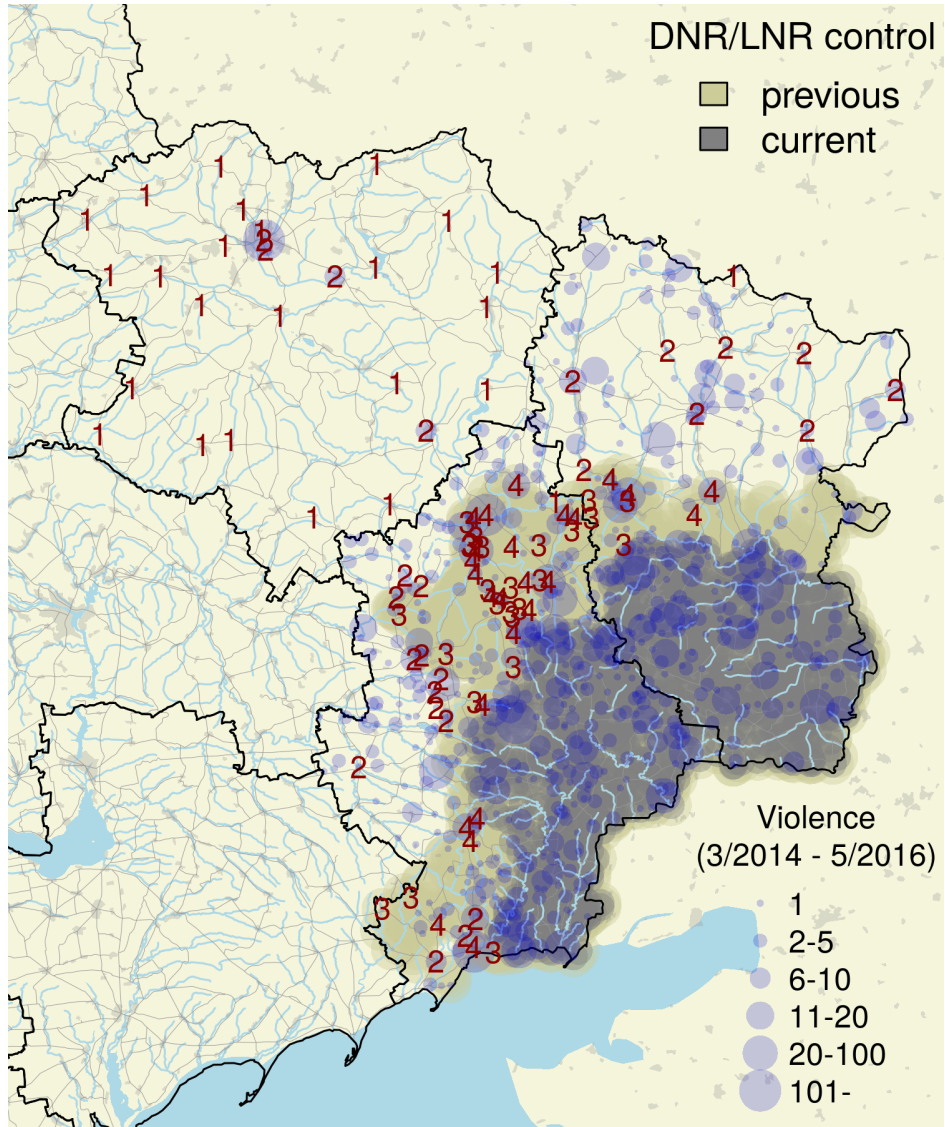
Recruitment of survey respondents relied on a multistage stratified sampling design, beginning with a stratified random sample of 100 government-controlled municipalities and city districts (PSUs), followed by simple random samples of electoral precincts within those municipalities, households within precincts, and individuals within households.

To ensure sufficient variation in quantities of theoretical interest, we stratified PSU's by exposure to separatist territorial control and violence. Figure 2.2 shows a map of the PSU's, in four strata: (1) no violence, no DNR/LNR control, (2) violence, no DNR/LNR control, (3) no violence, past DNR/LNR control, (4) violence, past DNR/LNR control.

The survey's secondary sampling units (SSUs) were 124 electoral precincts within the PSUs, excluding those located in hospitals and prisons. Within each settlement, the survey firms randomly selected 1-3 eligible voting precincts, depending on population size.

Within each precinct for which boundary information was available (streets and house numbers), the survey firm's regional office randomly selected a starting address (street and house number). If the selected starting address was block of flats, enumerators used a random number provided by the survey firm to select a starting apartment. Within precincts for which boundary information was not available (mainly, within villages and some small urban-type villages), enumerators visited the local post office or village council to obtain a list of streets. Using a random number provided by the survey firm, enu-

Figure 2.2: Survey's primary sampling units. Four strata: (1) no violence, no DNR/LNR control, (2) violence, no DNR/LNR control, (3) no violence, DNR/LNR control, (4) violence, DNR/LNR control.



merators then randomly selected streets, and starting apartments on those streets. Where there were no open post offices or village councils, enumerators stood in the center of the village, with their back facing the central building (i.e. closed post office or village council, shop, etc.), and counted streets from left to right. Using random numbers, enumerators then selected a street, and a starting dwelling on that street.

Starting from the first apartment or building, enumerators used random numbers to select new apartments or buildings to visit. In households that agreed to take part in the survey, enumerators selected a respondent randomly from the list of all eligible members of the household using a modified Kish Grid method.

Enumerators interviewed only one person within each household. If this person was not at home, the enumerator returned at another time. If this person refused to take part in the survey, the enumerator proceeded to the next household, rather than replacing the selected respondent with another family member. To avoid too many “closed apartments without contact,” the maximum number of such apartments within 1 precinct could not exceed 3 households with less than 3 attempts to establish a contact (main visit plus 2 call-backs). In other words, if an enumerator already had 3 “closed apartments without contact” they had to make a call-back to these “closed apartments without contact.” Only after these call-back visits were complete, could enumerators proceed to the next household. For each SSU, subject recruitment used soft quotas on sex and age, with no more than 65% female respondents, and no more than 50% aged 50 years or older.

2.2. Questionnaire

The survey covered multiple topics, most of which were unrelated to the current project. Questions of primary theoretical interest are the following:

- To the best of your knowledge, did any of your family members (parents, grandparents, siblings, uncles, cousins) die in the Ukrainian Famine, also known as the Holodomor, from 1932-1933?
 - Yes / No / Don’t know / Refuse to answer
- The Ukrainian famine of 1932-1933, also known as Holodomor, was a willful attempt by Soviet authorities to destroy the Ukrainian nation.
 - Strongly disagree / Disagree / Neutral / Agree / Strongly agree
- Please let us know whether you support or oppose [Russian Armed Forces / Donetsk and Luhans’k People’s Republics / Ukrainian Armed Forces / Right Sector] and how much?

– Strongly oppose / Oppose / Neutral / Support / Strongly support

We measure *opposition to pro-Russian armed groups* as the average of responses to last question above, flipping the sign for support of Russian Armed Forces and DNR/LNR:

$$\text{Opposition to pro-Russian groups} = \frac{(6 - \text{Rus. army}) + (6 - \text{DNR/LNR}) + \text{Ukr. army} + \text{R.S.}}{4}$$

In addition, the survey collected basic demographic information about respondents' age, sex, education, marital status, household size, profession and income.

Table 2.2 reports summary statistics for these and other survey responses.

Table 2.2: Summary statistics: survey data

| Variable | Range | Mean | SD |
|--|---------|-------|-------|
| <i>Famine exposure</i> | | | |
| Descendant of Holodomor victim | (0,1) | 0.20 | 0.44 |
| <i>Political preferences</i> | | | |
| Opposition to pro-Russian groups | (1,5) | 2.83 | 1.16 |
| Support Ukrainian Army (Likert) | (1,5) | 2.83 | 1.38 |
| Support Ukrainian Army (dummy) | (0,1) | 0.33 | 0.49 |
| Support Right Sector (Likert) | (1,5) | 1.61 | 0.94 |
| Support Right Sector (dummy) | (0,1) | 0.04 | 0.20 |
| Support Russian Army (Likert) | (1,5) | 1.86 | 0.91 |
| Support Russian Army (dummy) | (0,1) | 0.02 | 0.17 |
| Support DNR/LNR (Likert) | (1,5) | 1.80 | 0.89 |
| Support DNR/LNR (dummy) | (0,1) | 0.02 | 0.17 |
| <i>Historical views</i> | | | |
| Holodomor: due to state policy (Likert) | (1,5) | 3.33 | 1.38 |
| Holodomor: due to state policy (dummy) | (0,1) | 0.40 | 0.50 |
| Holodomor: due to natural causes (Likert) | (1,5) | 2.62 | 1.24 |
| Holodomor: due to natural causes (dummy) | (0,1) | 0.18 | 0.42 |
| <i>Demographics</i> | | | |
| Age | (18,90) | 51.06 | 17.02 |
| Profession (categories) | (1,12) | 6.66 | 3.30 |
| Income (categories) | (2,7) | 5.26 | 0.68 |
| Higher education | (0,1) | 0.22 | 0.41 |
| Male | (0,1) | 0.45 | 0.50 |
| Married | (0,1) | 0.55 | 0.50 |
| Ukrainian-speaking | (0,1) | 0.06 | 0.24 |
| Number of other adults in household | (0,5) | 0.42 | 0.75 |
| Likert variables range from "Strongly disagree/oppose" (1) to "Strongly agree/support" (5). Dummy variables are 1 if response in "Agree/Strongly agree" or "Support/Strongly support". | | | |

2.3. Territorial Control

We geographically matched the survey's PSUs with data on territorial control, particularly whether a populated place was under the control of pro-Russian separatists (DNR/LNR)

on a given day. For this information, we draw on three sets of sources. First are official daily situation maps publicly released by Ukraine's National Security and Defense Council (RNBO). Second are daily maps assembled by the pro-rebel bloggers 'dragon_first.1' and 'kot_ivanov.' Third are Facebook posts on rebel checkpoint locations prior to June 18, 2014 – the earliest date for RNBO and pro-rebel maps. For RNBO and pro-rebel blog maps – released to the public as high-resolution image files – we georeferenced and vectorized each daily map into spatial polygons. To construct polygons from user-reported checkpoint locations, we used the geographic convex hull of the coordinates of observed checkpoints on each day.

We coded a municipality as being under *DNR/LNR control* if, on a given day, it fell inside the rebel control polygons from one of the map collections. We created separate daily indicators for each of the three collections, and two combined indicators – DNR/LNR control according to at least one map collection, and control according to both RNBO and the bloggers. Prior to June 18, 2014, the two measures are equal, as there is only one source (Facebook rebel checkpoints.) We used this overlay to measure the number of consecutive days that each PSU was under DNR/LNR control in 2014-2016. Figure 2.2 shows the spatial distribution of past and present territorial control relative to our PSU locations.

3. SELECTING OPTIMAL INSTRUMENTS

Here we provide more details on how we constructed the ‘optimal’ scalar instrument for famine mortality – weather adversity index – and how we constructed alternative instruments that we use in our robustness analyses. The basic idea behind our approaches follow the previous work in (Belloni, Chernozhukov and Hansen, 2011; Belloni et al., 2012) who provide a general treatment of instrumental variables regressions when the number of instruments is large.

Let $\mathbf{w} = (w_{i1}, \dots, w_{i,48})$ denote all weather shocks (potential instruments), let \mathbf{x}_i denote all exogenous covariates, including oblast fixed effects, and cubic splines for distance to Russia. Then the first-stage model is given by:

$$\ln(\text{Famine}_i) = D(\mathbf{w}) + \mathbf{x}'_i\beta + \epsilon_i, \quad (1)$$

where D is an unknown function that maps from the set of all weather shocks to famine outcomes, after adjusting for covariates. To make the problem more tractable, we assume that D is a linear additive function, $D(\mathbf{w}) = \sum_{k=1}^{48} \gamma_k w_{ik}$. Even under this restriction we have an impractically large set of instruments. We therefore use several approaches to select a more feasible subset of these instruments.

3.1. Manually Selected Instruments

Our first approach is to manually select weather variables that have been mentioned in the historical literature, mainly (Davies and Wheatcroft, 2009; Tauger, 2001) who base their discussion on extensive reading of archival materials and back-of-the-envelope comparisons of temporal weather patterns. Our reading of this literature suggests the following weather events as potentially important:

1. Cold weather in March and April of 1931 delayed spring sowing, which made the grain more vulnerable to summer heat.
2. This cold spell preceded unusually hot temperatures in July of 1931.
3. Early frost in September-October 1931 led to germination failures of the winter grain.
4. Cold weather in March and April of 1932 again interfered with spring sowing, just as it did in 1931.

5. A cold spring again preceded a hot summer in 1932, with July temperatures adversely affecting the harvest.
6. The lack of rain in July 1932 led to further deterioration in crop yield according to [Davies and Wheatcroft \(2009\)](#), but [Tauger \(2001\)](#) argues that the rainfall during the same period in some areas precipitated infectious diseases.

In cases where studies mention temperature or rainfall in two subsequent months, we select only one of them to reduce the number of instruments and avoid collinearity.

While this is the most intuitive approach to instrument-selection, it has potential drawbacks. The historical literature is most likely to mention weather events that affected crop output and famine across the entire Ukraine, whereas our identification strategy relies on weather shocks that explain local, within-regional variation in famine outcomes. Also, important weather variables might simply be overlooked by these historical accounts, thus the next batch of methods of instrument selection use principled data-driven approaches.

3.2. Bayesian Model Selection

Our first data-driven approach is to use Bayesian model selection (BMS) ([Raftery, Madigan and Hoeting, 1997](#); [Clyde and George, 2004](#)). It estimates a large set of models using different combinations of weather variables. Variables are then selected based on their contribution in explaining the variation famine outcomes. Formally, let m denote a subset of instruments and let \mathcal{M} denote all such subsets – all 2^{48} possible combinations of weather variables. Then, the posterior probability that the model $m \in \mathcal{M}$ captures the data generating process is given by

$$\Pr(m|Y) = \frac{\pi(Y|m) \Pr(m)}{\Pr(Y)} = \frac{\pi(Y|m) \Pr(m)}{\sum_{m \in \mathcal{M}} \pi(Y|m) \Pr(m)}, \quad (2)$$

where Y represent all the data, $\pi(Y|m)$ represent the likelihood of the data given the subset m of instruments, and $\Pr(m)$ is the prior probability of model m , which we assume to be uniform, $\Pr(m) = 1/|\mathcal{M}|$ for all $m \in \mathcal{M}$. The optimal model m^* is the one with the highest posterior probability $\Pr(m^*|Y)$.³

³We use R package BMA ([Raftery et al., 2018](#)) to implement this procedure.

3.3. LASSO-Based Instruments

Our second approach is to use LASSO regression as in (Belloni et al., 2012), which solves the following optimization problem

$$\min_{\gamma} \sum \left(\ln(\text{Famine}_i) - \sum_{k=1}^{48} \gamma_k w_{ik} \right) + \lambda \|\gamma\|_1. \quad (3)$$

LASSO regression induces sparsity in the estimated coefficient vector $\hat{\gamma}$ by shrinking some coefficients to be identically zero. The number of selected variables depends on the penalty term λ . A standard approach in LASSO is to choose λ via cross-validation, but we find that this results in too many instruments being selected many of which end up being unimportant in the first-stage regression (something that is also documented in (Belloni et al., 2012)). Thus, we adjust the penalty λ so that the number of selected instruments does not exceed 16 (the number of weather instruments selected by the BMS).

3.4. Bayesian Model Averaging

We also construct a scalar index of *weather adversity* that succinctly summarizes the mapping between various weather variables to famine outcomes. This is driven by several motivations: First, we conduct a number of mechanism tests and placebo tests for instruments, and we would face an insurmountable number of issues with a large number of instruments. Second, both of the variable selection procedures have a known shortcoming in that they select an optimal set of variables, but ignore alternative sets of variables that can be very plausible though could be slightly less optimal ().

We circumvent both of these problems by constructing a scalar index of weather adversity that considers potential value of variables not selected in the single optimal model. The weather adversity index is defined as

$$\text{Weather}_i = \sum_{k=1}^{48} \hat{\gamma}_k w_{i,k}, \quad (4)$$

where the coefficients $\hat{\gamma}_k$ measure the model-weighted posterior average effects of weather shock w_k on famine mortality:

$$\hat{\gamma}_k = E(\gamma_k|Y) = \sum_{m \in \mathcal{M}} \pi(\gamma_k|m, Y) \Pr(m|Y). \quad (5)$$

The weather adversity index is a linear combination of individual weather shocks each

weighted by how well it predicts famine mortality, after adjusting for covariates and fixed effects.

3.5. Estimation Workflow

Using manually selected instruments in IV regressions is straightforward – we just include them in the first stage equation. For instruments selected by other methods, there are several options. First, one could calculate predicted famine mortality using the coefficients from the BMS, BMA, or LASSO regressions and then plug them into the second stage. Second, one could take the instruments selected by the BMS, BMA, or LASSO regressions and re-run the first stage using those instruments.

The latter approach (while perhaps less intuitive) has several important advantages over the first approach and has been recommended in the literature (Belloni et al., 2012). Since LASSO coefficients are generally biased downwards, Belloni, Chernozhukov and Hansen (2011) argue that it is more appropriate to re-estimate the first stage with OLS using only the instruments selected by LASSO (they call this procedure ‘post-LASSO’). The same issue applies to instruments selected by BMS or BMA. Furthermore, if we used the predicted values from BMS, BMA, or LASSO, the standard errors of the second-stage coefficients would need to be manually adjusted to account for the estimation uncertainty in the first-stage. If instead we use the second approach and regress selected instruments on the treatment using OLS, we can estimate the TSLS using standard statistical software that calculate correct standard errors.⁴ For these reasons, we use the second approach, identical to the post-LASSO procedure in (Belloni et al., 2012), with the following two-stage specification:

$$\ln(\text{Famine}_i) = \boldsymbol{\theta}'\text{Weather}_i + \hat{\beta}'\mathbf{x}_i + \epsilon_i, \quad (6)$$

$$y_i = \alpha \cdot \widehat{\ln(\text{Famine}_i)} + \boldsymbol{\zeta}'\mathbf{x}_i + u_i. \quad (7)$$

When there are multiple instruments selected by BMS or LASSO, the variable Weather_i is a vector of multiple instruments (and so $\boldsymbol{\theta}$ is a vector of parameters). When we use weather adversity index, then the variable Weather_i is a scalar, and so is θ .

When we use weather adversity index as an instrument, we need to adjust the calculated first-stage F statistic. The first-stage heteroscedasticity robust F statistic is given by

⁴We use *R* package *BMA* (Raftery et al., 2018) to fit the Bayesian model averaging, package *glmnet* to fit LASSO regressions (Simon et al., 2011), package *spdep* to implement the semi-parametric spatial filtering via Moran eigenvectors (Bivand, Hauke and Kossowski, 2013), and package *lfe* (Gaure, 2018) to estimate the IV regressions.

$F = \frac{\hat{\pi}V^{-1}\hat{\pi}}{k}$, where V is heteroscedasticity robust covariance matrix and k is the number of restrictions (instruments). When the weather adversity index is used, the software infers that $k = 1$, even though the index is constructed from many variables. To account for that, we recalculate the first stage F statistic returned by the software by setting $k = \sum_{k=1}^{48} \Pr(\gamma_k \neq 0|Y)$, the number of variables that were used in constructing the index weighted by their posterior weights (resulting in smaller F statistics than returned by the software).⁵

4. WEATHER AND FAMINE: ADDITIONAL EVIDENCE

4.1. First-Stage Results

Table 4.3 reports regression coefficients and robust standard errors for each instrument constructed by each of the four methods. The first stage F-statistics (that account for clustering at oblast level) indicate that the weather instruments are relevant, with the instruments selected by LASSO having the poorest performance. There is also some significant overlap with the instruments selected by different procedures, which is re-assuring.

The selected shocks are consistent with what historical accounts and agronomy suggest. To take a few examples, BMA and LASSO both select precipitation in July 1932 as an important predictor of famine. The sign of this coefficient suggests that precipitation increased famine mortality, consistent with [Tauger \(2001\)](#)'s claim that rainfall caused fungal infestation in crops, resulting in a poor harvest. Cold weather in March and April of 1931 also emerges as an important predictor: both BMA and LASSO procedures selected these instruments, and the coefficients are negative and statistically significant in both cases. This is consistent with the claim in [\(Davies and Wheatcroft, 2009\)](#) that cold weather in Spring of 1931 interfered with early sowing and also resulted in output losses.

A recent study by [Naumenko \(2017\)](#) on the causes of famine (not its political impact) argues that the onset of famine cannot be explained by weather, in that there was no major weather shock in 1931-1932 in Ukraine. We see nothing contradictory between our first-stage results and the latter findings. Our first-stage results do not imply that weather caused the *onset* of famine, only that famine mortality was worse in some places due to more adverse weather.

[Naumenko \(2017\)](#) employs a different dataset on famine mortality than the one we use in our analyses. Unlike [Naumenko \(2017\)](#), the excess mortality estimates in *Mapa* dataset

⁵One issue with data-driven approaches to instrument selection and construction is that the construction procedure itself can introduce measurement error. We discuss the robustness of our results to such measurement error in [Appendix 7.4](#)

Table 4.3: First stage estimates for different instruments, adjusted for oblast fixed effects and the covariates. Standard errors in parentheses clustered by oblast.

| | Instrument selection method: | | | |
|------------------------|------------------------------|-----------------|-----------------|----------------|
| | Index (1) | Manual (2) | BMA (3) | LASSO (4) |
| Weather index | 0.36*** (0.04) | | | |
| Precipitation: 1931-02 | | | -0.20*** (0.07) | |
| Precipitation: 1931-05 | | | -0.14** (0.06) | |
| Temperature: 1931-03 | | -0.45*** (0.11) | -1.44*** (0.16) | |
| Temperature: 1931-08 | | -0.71*** (0.26) | | |
| Temperature: 1932-03 | | 0.16 (0.13) | | |
| Temperature: 1932-07 | | 0.16 (0.12) | | |
| Temperature: 1931-04 | | | -0.40*** (0.15) | |
| Temperature: 1931-07 | | | 0.61*** (0.12) | |
| Temperature: 1931-12 | | | 0.51*** (0.11) | |
| Temperature: 1932-06 | | | 0.79*** (0.16) | |
| Temperature: 1932-08 | | | -0.33*** (0.12) | |
| Temperature: 1932-09 | | | -0.52*** (0.12) | |
| Precipitation: 1931-04 | | | | 0.06* (0.04) |
| Precipitation: 1931-06 | | | | -0.0001 (0.04) |
| Precipitation: 1932-02 | | | | -0.08 (0.11) |
| Temperature: 1932-11 | | | 0.19*** (0.05) | 0.26 (0.18) |
| Precipitation: 1931-07 | | -0.01 (0.01) | -0.18*** (0.04) | |
| Precipitation: 1931-12 | | | -0.21*** (0.05) | |
| Precipitation: 1932-05 | | | 0.22*** (0.04) | 0.05* (0.03) |
| Precipitation: 1932-07 | | 0.15*** (0.03) | 0.34*** (0.06) | 0.13*** (0.03) |
| Precipitation: 1932-11 | | | -0.14** (0.07) | |
| Precipitation: 1932-09 | | | | -0.04 (0.06) |
| Precipitation: 1932-12 | | | 0.24*** (0.07) | 0.11 (0.09) |
| Temperature: 1931-06 | | | | 0.004 (0.18) |
| Temperature: 1931-10 | | | | -0.35 (0.24) |
| Temperature: 1932-04 | | | | -0.27 (0.17) |
| Robust F -statistic | 193.63 | 13.88 | 16.29 | 7.59 |
| Observations | 386 | 386 | 386 | 386 |
| Adjusted R^2 | 0.85 | 0.80 | 0.85 | 0.80 |

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

rely on adjusted 1926, 1937 and 1939 census figures. In particular, the latter dataset corrects for the redistribution of armed forces and prisoners, compensates for census undercounts of small children and certain urban residents, and adjusts vital statistics for under-registration during the intercensal period. According to [Wolowyna et al. \(2016, 5-6\)](#), without these adjustments, the raw census data artificially boost population numbers in rural areas that were decimated by Holodomor. Based on our personal correspondence with the authors of these two studies, the main disagreement between the datasets are in their estimates of *total* famine mortality, Ukraine-wide, not cross-sectional variation. At the rayon level, which is the only thing that matters for our analysis, correlation between the two measures of famine mortality is 0.98.

4.2. *Qualitative Evidence*

We now document some additional qualitative evidence, in addition to the evidence provided in [Tauger \(2001\)](#) and [Davies and Wheatcroft \(2009\)](#), suggesting different ways in which weather might have contributed to famine mortality.

At the time of famine, the Central Committee of the Communist Party of the Soviet Union (Bolsheviks) (CC CPSU) understood that weather was one of multiple causes of the event, not only in Ukraine, but also in other parts of the USSR. On November 2, 1932, the director of the All-Union grain trust of Ukraine and Crimea A. Burkov reported to CC CPSU that

“[...] as a result of heavy rains, weeds started to kill crops. [...] We gathered 72% of crops manually instead of gathering 70% with combine-harvesters and only 30% manually. Why did we have to harvest not with combine-harvesters, but mostly manually? The main reason is that the fields are heavily clogged with weed and we could not start combine-harvesters at these weeded fields” ([Burkov, November 2, 1932](#)).

Burkov’s main claim was that heavy rain impeded the use of agricultural technologies.

Other local executives also mentioned bad weather conditions. For example, on August 20, 1932, the first secretary of the North Caucasus regional committee of the CPSU B. Sheboldaev wrote to Stalin the following:

“Undoubtedly, the poor farming of collective and state farms (late sowing, lack of crop rotation, poor cultivation of fields) which gave a huge contamination and thinned grain affected the grain supply, but there were also special weather conditions (‘fuse’, ‘stack’, ‘rust’, ‘fog’, etc., as these phenomena are called by agronomists and practitioners), which sharply reduced the harvest, and in some areas brought it to naught (Georgievsky, Vorontsovo-Alexandrovsky, Mozdoksky, former Prokhladnensky and other regions). Finally, the harvesting conditions (rains for one and a half months) also led to some losses (germination)” ([Sheboldaev, August 20, 1932](#)).

Similarly, on November 28, 1932, secretary of the regional committee of the CPSU (b) of the Russian Federation I. Vareikis reported that

“In the presence of an average yield for other grain crops, in case of oats and wheat the yield was significantly lower, and we lost a lot of oats due to the dry wind and rust that swept us during the filling (the last phase of ripening - AV). A few days of sharp dry wind in early August led to the fact that oats gave a decrease in yields of 2 or even 3 times, and in some state farms oats were gathered for cattle feed, because there was no point in threshing” (Vareikis, November 28, 1932).

Reports of inclement weather were sent even from Siberia: “a number of unfavorable conditions in the current year (prolonged sowing, prolonged rains, floods, frosts) led to a significant decrease in yields,” claimed the first secretary of the East Siberian Regional Committee of the CPSU (b) F. Leonov in his letter to Stalin on October 22, 1932 (Leonov, October 22, 1932).

5. ATTRIBUTION OF FAMINE DEATHS AND REPRESSION

In the paper, we argue that weather-induced famine is a theoretically plausible treatment with which we can estimate the effects of state repression. Here we provide additional qualitative evidence to support this claim. Using memoirs of famine survivors and archival materials, we make two related arguments. First, we show that, empirically, places with worse weather had more famine deaths. The current section provides illustrative examples that witnesses interpreted these famine deaths as having resulted from the coercive policies of the regime, even though weather contributed to the process. Second, we show that locally bad weather precipitated worse grain output, which exposed communities to conspicuous *acts* of state repression, in the form of food and cattle confiscation, beatings, and mass arrests. In sum, places hit by bad weather saw more famine deaths because they were exposed to more punitive measures.

5.1. Attribution of Famine Deaths

Memoirs of Holodomor eyewitnesses overwhelmingly show that they saw famine deaths as resulting from how the regime acted prior to the onset of famine and how it responded to food shortages and starvation. First, survivors often mention *collectivization* as a form of repression that directly (and intentionally) caused mass starvation and deaths. Here are some common statements to that effect:

“People were forced to join collective farms, but people did not want to. It was like committing suicide. So they organized a brigade that took away everything we had. And people were left

without anything and got swollen from hunger” (Natsional’nyy tsentr narodnoyi kul’tury, 2018).

“They came and forced us to join the collective farm. But how can we join if the father is not at home? So they took the cow and the inventory: plows, harrows, and mowers, and a steam engine – everything...” (Natsional’nyy tsentr narodnoyi kul’tury, 2018).

“Who did not go to the collective farm, they took everything to the grain from them. When everything was taken, the spring came, but there was nothing to eat” (Golodomor 1932-1933, N.d.).

Second, survivors often mention that Soviet authorities not only failed to alleviate their plight, but instead made it even worse by taking away their last bits of food and cattle, which were often their last source of nutrition during the winter of 1933 (the most deadly episode):

“The members of the Bread Procurement Commission continued in their task, relentlessly tramping from house to house, confiscating everything edible they could find in their attempt to meet the state quotas. Even the smallest amounts of grain and meat were forcibly seized from the villagers. [...] A notice came that within twenty hours we had to deliver about 250 pounds of meat to the state in the form of livestock. This meant we had to give up our cow. We never cried so much as we did on that day. It was as if we were losing our very lives, which indeed was not far from the truth” (Dolot, 1985, p. 174, 222).

It is quite understandable that, in these conditions, people would not think that their plight had been caused by weather. Without forced collectivization, grain requisition, food confiscation, and mass arrests, they would not have witnessed the extent of death that they did, irrespective of weather. They overwhelmingly interpreted famine as having political roots, since “the party ripped people off, cheated them” (Pyrig, 2007, p. 115). It is only natural that people would then interpret famine as a genocidal act with a political intention to annihilate them not just as individuals, but as members of a community (Ukrainians):

“That’s all because of the orders of Stalin, who did not like Ukrainians and wanted to massively kill them.” (Mykola Khmelnyk Golodomor 1932-1933, N.d.)

“I can not keep silent, it is necessary for our children and grandchildren to know the truth about the terrible fate of the Ukrainians who lived in that horrible year, 1933. Maybe it was the punishment of the Lord for our sins! But no! Rather, the revenge of the devil, because the Lord could not be so cruel. It is difficult to remember, but it is necessary for the commemoration of the deceased.” - Kateryna Chnmut⁶

“Nothing could save a simple person. After all, the policy of genocide was conducted by the state’s repressive and punitive system.” - O. Kuzmenko⁷

⁶Testimony of the Holodomor Witnesses, <https://tinyurl.com/y9x91ymd>.

⁷Testimony of the Holodomor Witnesses, <https://tinyurl.com/y9x91ymd>.

5.2. Exposure to Acts of Repression

From the archival evidence we can also infer that bad weather not only increased local exposure to famine deaths (due to reduction in available food, discussed earlier), but that it also exposed local populations to conspicuous *acts* of state repression. Most pressure was exerted on rayons that were not able to fulfill their quotas – and rayons exposed to worse weather were naturally more likely to end up in that position (this follows from the generally existing relationship between weather and grain output in Ukraine as well as the evidence we documented earlier).

There is substantial evidence – from all levels of government – that the main response to failures to deliver grain quotas was to apply direct punitive measures. On November 5, 1932, Vyacheslav Molotov – who was sent by Stalin to Ukraine to oversee grain procurements – sent a telegram to authorities in Odessa oblast, in which requested to “immediately begin to put into practice rigid restrictions on the imports of manufactured goods (with the exception of such goods as salt, matches, kerosene) to villages and collective farms that are unsatisfactorily carrying out the plan of grain procurement, and also to apply the transfer of goods from these villages to villages that successfully fulfill the plan for grain delivery” (Kondrashin, 2012, p. 172-173).

Internal reports from Soviet authorities inside Ukraine admitted that weather in some locations was inhibiting growth, sowing, and collection of grain, as we discussed earlier in Appendix 4.2. However, the Central Committee of the Ukrainian Communist Party explicitly disallowed the use of weather and other exigencies as excuses for missing quotas, and instructed local authorities to treat shortages as acts of political sabotage:

“This sharp decline [in procurements] cannot be explained by any objective causes like rain and so on. The Central Committee believes that the main reason is bad organizational work in the area of grain procurement... This shows that the measures specified in the resolution of the Central Committee in November 18 are not yet truly enforced and that the resistance and the sabotage organized by kulak counter-revolutionary elements and their conspirators is not yet broken, *the repressive measures against them are applied insufficiently and too hesitantly.*”⁸

The November 18 decree referred to here as not “applied insufficiently” mandated the creation of a “black list” of collective farms that had failed to fulfill their quotas. When a collective farm was black-listed, then, among other things, the state would immediately halt deliveries of goods, close stores, prohibit all trade with such collective farms, any deny access to credit and advances. Such farms would also be inspected and “cleaned”

⁸Decree No. 17 of Central Committee of the CP(b) of Ukraine “On the application of repressive measures against collective farms that sabotage grain procurement,” November 27, 1932, Central State Archives of Public Organizations of Ukraine, F.1, Op.6, Sp. 238, Arch. 21-22.

of its “counter-revolutionary” elements, and the territory would be sealed off by OGPU detachments. ⁹ Essentially, this would mean “ a death sentence was imposed on the population of the given kolkhoz, village, or raion” (Wolowyna et al., 2016, p. 193).

Outside party-level decrees, OGPU field offices were instructed to apply repressive measures widely in response to grain shortages. “[OGPU] documents testify that at the peak of famine, repressive measures were the authorities’ key priority” (Berelovich, 2005, p. 24). OGPU operatives on the ground deliberated about repression in the following terms: “Without administrative pressure we will not be able to retrieve bread, therefore, it is fine if we go over the top” (Berelovich, 2005, p. 205). ‘Administrative pressure’ (*administrativnoe vozdeystvie*, Soviet bureaucratic slang for repression) included arrests, beatings, mass searches, and confiscation of all property, including household items.

Even by the admission of the authorities, arrests during the famine were highly arbitrary. For example, a regional party secretary from Dnipropetrovsk oblast reported to Stalin (Kondrashin, 2012, p. 657, emphasis added):

“Anyone who wanted to arrest was arresting: not just every policeman, not only a commissioner for grain procurements, but also letter-carriers [...] *In a very large percentage, people were arrested who should not have been arrested at all, and actual enemies and saboteurs escaped arrest.*”

Party officials described conditions in detention facilities as “inexcusable. The detainees we placed in unheated premises (in sheds), without clothes, pressed one against another, and could not even sit down” (Kondrashin, 2012, p. 657).

⁹Decree No. 15 of Central Committee of the CP(b) of Ukraine “On strengthening grain procurement,” November 27, 1932, Central State Archives of Public Organizations of Ukraine, F.1, Op. 6, Spr. 237, Arch. 207-216.

6. ADDITIONAL IV RESULTS

6.1. Full Second-Stage Output

| | Opposition to Red partisans | Anti-Soviet vote | Anti-Soviet protests | Anti-Russian vote | Anti- Yanukovich protests |
|-------------------------|--------------------------------|---------------------|-------------------------|----------------------|---------------------------------|
| | (1) | (2) | (3) | (4) | (5) |
| Famine | 0.77*** (0.09) | -0.45*** (0.14) | 0.45*** (0.15) | 0.36*** (0.06) | 0.50*** (0.17) |
| Ukrainians (1926) | -0.45** (0.20) | 1.51*** (0.35) | -0.61 (0.56) | 0.48** (0.20) | -0.58 (0.77) |
| Russians (1926) | 0.88** (0.42) | 0.99 (0.76) | -0.21 (1.04) | -0.82** (0.34) | 0.10 (1.22) |
| Rural population (1926) | 0.24* (0.14) | -0.94*** (0.24) | -2.08*** (0.69) | -0.63*** (0.17) | -3.33*** (0.97) |
| Forestation | 0.07 (0.08) | 0.15 (0.13) | -0.22 (0.15) | 0.15** (0.06) | -0.46* (0.27) |
| Industrialization | -0.08 (0.06) | 0.01 (0.08) | -0.09 (0.12) | -0.07** (0.03) | -0.20 (0.17) |
| Dominant crop: dairy | -0.32*** (0.12) | 0.08 (0.19) | 0.58*** (0.18) | 0.13** (0.06) | 0.70* (0.37) |
| Dominant crop: potato | 0.30*** (0.08) | -0.88*** (0.17) | 0.08 (0.24) | -0.22*** (0.06) | 0.08 (0.35) |
| Dominant crop: wheat | -0.09 (0.07) | -0.41*** (0.14) | -0.29* (0.15) | -0.07* (0.04) | -0.40 (0.28) |
| Population (1926) | 0.05 (0.05) | -0.14** (0.06) | 0.41*** (0.09) | 0.09** (0.03) | 0.67*** (0.13) |
| Moran's I (p-value) | 0.49 | 0.99 | 0.99 | 0.85 | 0.98 |
| Observations | 386 | 386 | 386 | 386 | 386 |
| Adjusted R ² | 0.94 | 0.82 | 0.97 | 0.97 | 0.93 |

*p<0.1; **p<0.05; ***p<0.01.

Table 6.4: Second-stage IV estimates with pre-famine covariates. Weather adversity index is used as instrument. Coefficients for Oblast fixed effects, cubic spline for distance to Russia, and spatial Moran eigenvectors are not shown. Standard errors adjust for clustering at oblast level.

6.2. Results without normalized outcomes

Table below shows results of IV regression identical to the baseline specifications reported in the paper, except that here we do not normalize the outcome variables.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|---|--------------------------|----------------------------|------------------------|
| Opposition to Red partisans (1941-1944) | 0.66 (0.08)** | 12.40 | 0.49 |
| Anti-Soviet vote (1946-1958) | -0.07 (0.02)** | 13.07 | 0.99 |
| Anti-Soviet protests (1987-1991) | 4.70 (1.60)** | 11.72 | 0.99 |
| Anti-Russian vote (2002-2014) | 6.90 (1.09)** | 13.48 | 0.85 |
| Anti-Yanukovich protests (2009-2013) | 58.01 (19.48)** | 12.17 | 0.98 |

The estimates are for logged famine mortality instrumented by weather adversity index. All specifications control for pre-famine rural population, proportion of Ukrainians, proportion of Russians, forestation, industrialization, dominant crop dummies, oblast fixed effects, synthetic spatial covariates, and cubic spline for the distance to Russian border. Significance levels (two-tailed): $\dagger p < 0.1$; $*p < 0.05$; $**p < 0.01$.

Table 6.5: IV RESULTS WHEN DEPENDENT VARIABLES ARE NOT NORMALIZED.

We see that the effects are quite substantial. For example, doubling famine mortality increased protest count by about 4 protests during the late Soviet period and by 58 protests in the post-Soviet period (there were many more protests during the post-Soviet period to start with). Also, the vote-percentage of anti-Russian parties increased by about 7 points as a result of doubling of famine deaths. The effect of famine on anti-Soviet votes is small in magnitude (0.07 percent), but this is because the overall variation in this variable was small; in fact, the 0.07 percent raw effect constitutes about one standard deviation effect.

7. ROBUSTNESS CHECKS FOR IV RESULTS

7.1. Alternative Specifications

In the current section, we evaluate the robustness of our results to several alternative model specifications. Table 7.6 reports Two-Stage Least Squares (TSLS) estimates for each outcome without pre-famine population weights and with additional covariates controlling for contemporaneous weather.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|--|--------------------------|----------------------------|------------------------|
| <i>Without population weights</i> | | | |
| Opposition to Red partisans (1941-1944) | 0.58 (0.12)** | 13.60 | 0.97 |
| Anti-Soviet vote (1946-1958) | -0.41 (0.14)** | 11.21 | 0.94 |
| Anti-Soviet protests (1987-1991) | 0.16 (0.16) | 6.22 | 0.00 |
| Anti-Russian vote (2002-2014) | 0.38 (0.05)** | 10.64 | 0.45 |
| Anti-Yanukovich protests (2009-2013) | 0.13 (0.16) | 6.22 | 0.00 |
| <i>Without Moran eigenvectors</i> | | | |
| Opposition to Red partisans (1941-1944) | 0.36 (0.47) | 12.03 | 0.00 |
| Anti-Soviet vote (1946-1958) | -0.57 (0.17)** | 12.03 | 0.00 |
| Anti-Soviet protests (1987-1991) | 0.46 (0.15)** | 12.03 | 0.99 |
| Anti-Russian vote (2002-2014) | 0.41 (0.08)** | 12.03 | 0.00 |
| Anti-Yanukovich protests (2009-2013) | 0.50 (0.18)** | 12.03 | 0.97 |
| <i>Controlling for contemporaneous weather</i> | | | |
| Opposition to Red partisans (1941-1944) | 0.72 (0.10)** | 10.17 | 0.78 |
| Anti-Soviet vote (1946-1958) | -0.43 (0.12)** | 15.53 | 0.97 |
| Anti-Soviet protests (1987-1991) | 0.44 (0.15)** | 11.75 | 0.99 |
| Anti-Russian vote (2002-2014) | 0.38 (0.06)** | 13.41 | 0.84 |
| Anti-Yanukovich protests (2009-2013) | 0.51 (0.18)** | 12.34 | 0.97 |

Note: All specifications control for pre-famine rural population, proportion of Ukrainians, proportion of Russians, forestation, industrialization, crop suitability, oblast fixed effects, synthetic spatial covariates, and cubic spline for the distance to Russia. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $*p < 0.05$; $**p < 0.01$.

Table 7.6: Additional IV results.

Estimates reported in the top panel of Table 7.6 suggest that excluding pre-famine population weights yields quite similar conclusions. The estimates for *Red partisans*, *Anti-Soviet vote*, and *Anti-Russian vote* are nearly identical in magnitude and significance to baseline estimates. However, while in the same direction, point estimates for the two

protest variables are smaller in magnitude, and not significant at 95 percent level.

The smaller point-estimates in non-weighted regressions makes intuitive sense. Since popular protests tend to occur in large urban areas, it is reasonable to expect the treatment effect to attenuate when places where protests cannot happen (irrespective of the underlying political preferences) carry the same weight in the estimation as places where protests *can* happen.

In the second panel of Table 7.6, we adjust for contemporaneous weather. Since weather is temporally autocorrelated and future weather may impact voting, protests and other outcomes, there is a concern that this temporal dependence might violate the exclusion restriction.¹⁰ For each outcome in the analysis, we calculate a contemporaneous average weather adversity index and add it as a control variable. For example, for *Red partisans* we estimate average weather adversity for the years 1941-1944. We do so by first calculating weather adversity for each year from 1941 to 1944 – using the same approach as for weather adversity in 1933 – and then averaging the computed indexes. The results remain very similar to the baseline IV regressions.

7.2. *Alternative Instruments*

Table 7.7 reports IV coefficients using three alternative methods to select instruments. All specifications control for the same covariates and fixed effects as the model in the main text. The estimates are all very similar in direction, significance and magnitude to the ones reported in the main text.

¹⁰We thank anonymous reviewer for indicating this possibility and suggesting this test.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|---|--------------------------|----------------------------|------------------------|
| <i>Manually selected instruments</i> | | | |
| Opposition to Red partisans (1941-1944) | 0.45 (0.15)** | 15.74 | 0.45 |
| Anti-Soviet vote (1946-1958) | -0.14 (0.14) | 20.82 | 0.99 |
| Anti-Soviet protests (1987-1991) | 0.80 (0.25)** | 13.52 | 0.99 |
| Anti-Russian vote (2002-2014) | 0.36 (0.06)** | 18.84 | 0.96 |
| Anti-Yanukovich protests (2009-2013) | 0.77 (0.29)** | 15.54 | 0.98 |
| <i>Instruments selected by BMA</i> | | | |
| Opposition to Red partisans (1941-1944) | 0.62 (0.07)** | 18.86 | 0.15 |
| Anti-Soviet vote (1946-1958) | -0.51 (0.13)** | 18.27 | 0.99 |
| Anti-Soviet protests (1987-1991) | 0.45 (0.13)** | 16.06 | 1.00 |
| Anti-Russian vote (2002-2014) | 0.36 (0.05)** | 18.56 | 0.76 |
| Anti-Yanukovich protests (2009-2013) | 0.53 (0.17)** | 16.70 | 0.97 |
| <i>Instruments selected by LASSO</i> | | | |
| Opposition to Red partisans (1941-1944) | 0.69 (0.10)** | 13.26 | 0.01 |
| Anti-Soviet vote (1946-1958) | -0.36 (0.10)** | 10.69 | 0.99 |
| Anti-Soviet protests (1987-1991) | 0.68 (0.18)** | 8.76 | 1.00 |
| Anti-Russian vote (2002-2014) | 0.30 (0.06)** | 10.45 | 0.90 |
| Anti-Yanukovich protests (2009-2013) | 0.67 (0.20)** | 9.97 | 0.97 |

All specifications control for pre-famine rural population, proportion of Ukrainians, proportion of Russians, forestation, industrialization, crop suitability, oblast fixed effects, synthetic spatial covariates, and cubic spline for the distance to Russia. Significance levels (two-tailed): † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 7.7: IV results using alternative instruments.

7.3. Alternative Measures of WWII Partisan Behavior

Table ?? evaluates the robustness of our results to alternative measures of ant-Soviet political behavior during the WWII. The outcome variables here include

- *Opposition to Red partisans.* In the main text, we use 50km bandwidth to calculate the local density of partisan camps. We now use two alternative bandwidths of 25km and 75km.
- *Distance to partisan operations.* Minimum distance (in kilometers) from the centroid of each 1933 rayon to the nearest major Red partisan operation against German forces. Higher values indicate greater behavioral disloyalty.
- *Density of partisan operations.* Kernel density (25 km bandwidth) of major Red partisan operations in each 1933 rayon. Lower values indicate greater behavioral disloyalty.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|---------------------------------|--------------------------|----------------------------|------------------------|
| Opp. to Red partisans (25km bw) | 0.54 (0.09)** | 13.54 | 0.88 |
| Opp. to Red partisans (75km bw) | 0.65 (0.08)** | 11.09 | 0.99 |
| Distance to partisan operations | 0.64 (0.09)** | 15.23 | 1.00 |
| Partisan operations (density) | -0.74 (0.12)** | 13.91 | 0.95 |

The estimates are for logged famine mortality instrumented by weather adversity index. All specifications control for pre-famine rural population, proportion of Ukrainians, proportion of Russians, forestation, industrialization, crop suitability, oblast fixed effects, synthetic spatial covariates, and cubic spline for the distance to Russian border. Significance levels (two-tailed): $\dagger p < 0.1$; $*p < 0.05$; $**p < 0.01$.

Table 7.8: IV RESULTS FOR WWII OUTCOMES USING ALTERNATIVE MEASURES.

The estimates in Table 7.8 are consistent with our theoretical expectations, and the results reported in the paper. First, main partisan operations occurred further away from locations that experienced high famine mortality. Second, rayons that experienced more famine saw fewer local partisan operations during WWII. In Appendix 11, we discuss two additional measures of WWII behaviors.

7.4. Measurement Error in the Instrument Construction

The weather adversity instrument that we use in the baseline specifications relies on weights for each weather shock, estimated through Bayesian model averaging. Due to uncertainty in this instrument construction stage, a potential concern is that our instrument may be measured with error.¹¹ In a sense, similar issues may arise with respect to *any* instruments, including manually selected ones. In most applications, there is at least some uncertainty as to whether the selected instruments are the “correct” ones.

To check in a principled way if our results would change substantively when one takes into account estimation errors in the instrument construction stage, we adopt a multiple imputation approach. Specifically, we sample the instrument value from the model estimated via Bayesian model averaging procedure, and then repeatedly estimate a TSLS regression for each sampled value of the instrument. This procedure allows the uncertainty in instrument construction to pass through into the main regression estimates.

Formally, let $\hat{\gamma}_k^m$ denote the weight for weather shock k in the model $m \in \mathcal{M}$, and let $\hat{\gamma}_m = (\hat{\gamma}_1^m, \dots, \hat{\gamma}_{48}^m)$ be a vector of model m 's weights for all 48 weather shocks. Let $\hat{\Sigma}_m$ denote the covariance matrix of the coefficient estimates in model m . We then iterate the following steps for $b = 1, \dots, B$.

1. Sample a model $m^* \sim \Pr(m|Y)$, where $\Pr(m|Y)$ is the posterior probability of model m .
2. Sample regression coefficients from their posterior distribution in the model m^* : $\gamma^* \sim \mathcal{N}(\hat{\gamma}_{m^*}, \hat{\Sigma}_{m^*})$.
3. Calculate weather adversity index $Weather_i^b = \sum_{k=1}^{48} \gamma_i^* w_{i,k}$.
4. Plug in $Weather_i^b$ as the instrument into the TSLS regression and estimate the coefficient for the treatment effect ($\hat{\theta}_b$), its cluster-corrected standard error ($\hat{\sigma}_b$), and the first-stage F -statistic for the instrument (F_b).

We set $B = 1,000$, and we use standard formulas (Robins and Wang, 2000) to aggregate the coefficients and their standard errors from the imputed datasets. For the TSLS coefficient, we use:

$$\hat{\theta} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b. \quad (8)$$

¹¹We are grateful to an anonymous reviewer for pointing us to this possibility.

We calculate standard errors using:

$$S.E.(\hat{\theta}) = \sqrt{\frac{1}{B} \sum_{b=1}^B \hat{\sigma}_b + \left(1 + \frac{1}{B}\right) \frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b - \hat{\theta})^2}. \quad (9)$$

We also report the first stage F -statistic, averaged across all imputations.

Table 7.9 reports the imputed estimates. The point estimates of the coefficient are nearly identical to the baseline specifications, suggesting that measurement error in the weather adversity instrument is not likely to bias the results. The standard errors are, expectedly, larger than in the baseline specifications, because the imputed estimates account for an additional layer of uncertainty in instrument construction. Three of the five TSLS coefficients remain significant at a 95% confidence level or higher. TSLS coefficients for the two protest variables are significant at a 90% confidence level. Overall, and despite the somewhat lower precision of the latter two estimates, the results are consistent with the baseline estimates.

| Dependent variable | IV coefficient | S.E. | p-value | First-stage F |
|---|----------------|------|---------|-----------------|
| Opposition to Red partisans (1941-1944) | 0.77 | 0.13 | 0.00 | 12.03 |
| Anti-Soviet vote (1946-1958) | -0.45 | 0.21 | 0.03 | 12.51 |
| Anti-Soviet protests (1987-1991) | 0.45 | 0.23 | 0.06 | 10.98 |
| Anti-Russian vote (2002-2014) | 0.36 | 0.12 | 0.00 | 12.65 |
| Anti-Yanukovich protests (2009-2013) | 0.51 | 0.29 | 0.08 | 11.30 |

Imputed estimates of the IV effects accounting for the uncertainty in the instrument measurement (baseline covariates). Standard errors clustered by oblast. Significance levels (two-tailed): † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 7.9: IV RESULTS USING IMPUTED INSTRUMENT.

7.5. Sensitivity Check

The current section evaluates the sensitivity of our results to outliers. To implement this test, we remove rayons one-by-one from the dataset and re-estimate our TSLS regressions, using the same base specification. Figure 7.3 reports point estimates and 95 percent confidence intervals for each regression. The estimates are quite stable, especially for *Red partisans*, *Anti-Soviet vote*, and *Pro-Russian vote*. There is one outlying observation – the city of Kyiv – that affects estimates for the two protest variables. Note, however, that in each case the magnitude of the estimated effect *increases* after we remove Kyiv from the analysis. This increase makes intuitive sense: Kyiv is Ukraine’s capital city, where protests happen for many reasons, so one may reasonably expect the presence of this district in the dataset to obfuscate the relationship between famine and protest behavior.

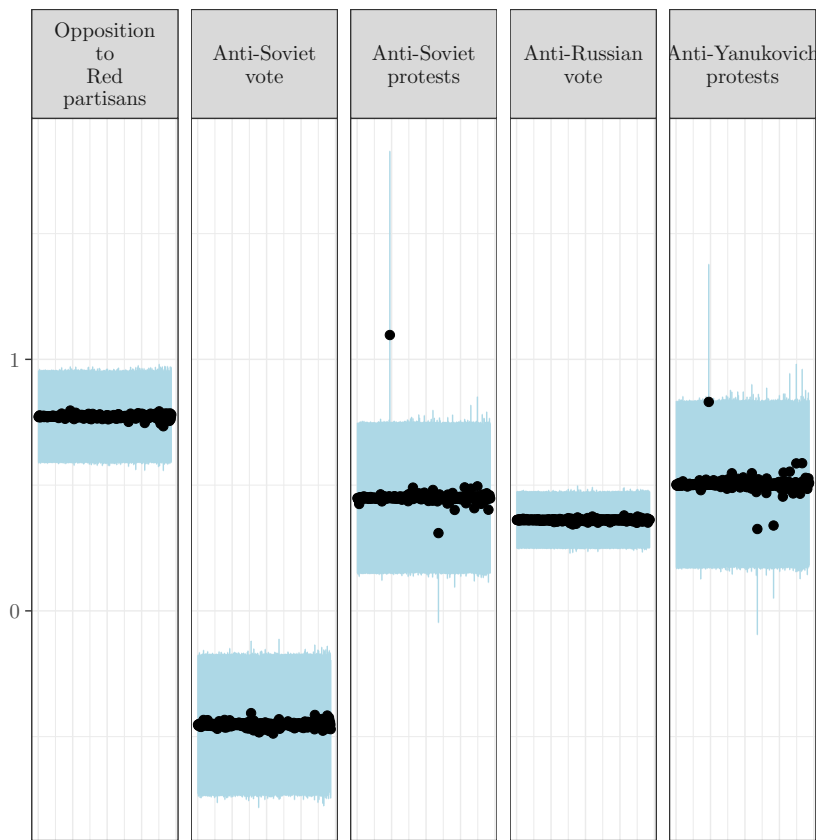


Figure 7.3: Point estimates of the famine effect (instrumented by weather adversity index) with the 95 percent confidence intervals removing rayons one-by-one.

8. A PLACEBO TEST: WESTERN UKRAINE

The weather shocks we use as instruments for famine could have impacted political outcomes through channels other than famine. In this section, we investigate whether such a violation of the exclusion restriction is empirically plausible (as is generally the case, we cannot completely rule out such a possibility on empirical grounds).

Our placebo test exploits the fact that not the entire territory of post-famine Ukraine was under Soviet control during the famine. Until 1939, a sizable proportion of western Ukraine was controlled by Poland, Czechoslovakia and Hungary and did not experience the famine, despite being in a similar climactic zone as Soviet Ukraine (central and eastern Ukraine before 1939). Were the exclusion restriction violated, we would observe a similar reduced-form relationship between weather adversity and political outcomes in western Ukraine as we do in the Soviet Ukraine. In that sense, western Ukraine can serve as a placebo case to evaluate the plausibility of the exclusion restriction.

Using the weights for each weather variable calculated from the Soviet Ukraine data, we construct a weather adversity index for western Ukraine. Even though there was no famine in western Ukraine, the variable *Weather* here represents the accumulation of weather shocks that would have produced famine in a counter-factual scenario of western Ukraine being under the same political and economic system as Soviet Ukraine in 1933. Because the Soviet Union annexed western Ukraine only in 1939 and was under German occupation for much of WWII, the earlier set of rayon boundaries we could obtain for this region were from 1945.

We then estimate the following reduced form regression equations for western and Soviet Ukraine

$$\mathbb{E}(y_i) = \alpha \cdot \text{Weather}_i + \beta' \mathbf{x}_i + \text{Oblast}_{j[i]} + \mathbf{u}_i, \quad (10)$$

where \mathbf{x} is the vector of control variables, *Oblast* is the regional fixed effect and u_i 's are the spatial synthetic covariates. If a violation of the exclusion restriction is biasing our results in the direction reported in the paper, then the coefficient α in the above equation estimated on western Ukraine should be in the same direction and significance as coefficient estimates for Soviet Ukraine. This would mean that adverse weather affected political outcomes through some other channels than weather.

Table 8.10 reports the results. We see that the reduced-form relationships between weather adversity and political outcomes are all significant and in the same direction as the IV estimates (as expected). However, the estimates in the placebo case of western

| | Western Ukraine (N = 226) | Soviet Ukraine (N = 386) |
|---|------------------------------|-----------------------------|
| Opposition to Red partisans (1941-1944) | -0.11 (0.07) [†] | 0.27 (0.03)** |
| Anti-Soviet vote (1946-1958) | 0.15 (0.09) [†] | -0.16 (0.05)** |
| Anti-Soviet protests (1987-1991) | 0.09 (0.08) | 0.16 (0.05)** |
| Anti-Russian vote (2002-2014) | 0.03 (0.11) | 0.13 (0.02)** |
| Anti-Yanukovich protests (2009-2013) | 0.07 (0.08) | 0.18 (0.07)* |

[†]p<0.1; *p<0.05; **p<0.01.

Table 8.10: Reduced form relationships between weather adversity index and political outcomes in western Ukraine (placebo case) and Soviet Ukraine. The coefficients are for weather adversity index after controlling for all first stage covariates and oblast fixed effects as well as synthetic spatial covariates.

Ukraine are either small in magnitude and statistically not significant (for *Anti-Soviet vote*, *Anti-Soviet protests*, and *Anti-Russian protests*) or have the opposite sign compared to the estimates for Soviet Ukraine (*Red partisans* and *Pro-Russian vote*). These results suggest that if there are violations of the exclusion restriction for these two variables, then they are likely to bias our results to the opposite direction than the reported IV estimates, especially for Red partisans and anti-Soviet votes.

We conducted additional placebo tests for instruments selected by manually, by LASSO or Bayesian model selection. The basic idea behind these placebo tests is the same: we want to see whether weather variables in western Ukraine are similarly correlated with post-famine political outcomes, after baseline adjustments. However, here we face a somewhat tricky problem of multiple comparisons: for example, LASSO procedure selects 11 weather variables, and since we have five outcomes, we would need to compare 55 coefficients between regressions for western and central/eastern Ukraine.

We adopt the following rule to decide whether there is a placebo effect that would bias our IV results away from zero: for a given weather variable, if coefficients in western and central-eastern Ukraine are both significant and in the same direction, then the weather variable has a placebo effect. In other words, we should not expect all weather variables in western Ukraine to be insignificant – some of them can be significant by chance, whereas others could be significant, but in the opposite direction to the non-placebo case (which would not indicate a violation of the exclusion restriction that would bias our results away from zero). For significance tests, to make sure that we are not putting too stringent requirement for a placebo test to pass, we set the confidence level of 90 percent; a higher confidence level would generate fewer placebo effects.

Table 8.11 summarizes the results of placebo tests. The first row essentially replicates

| | Opposition to Red partisans | Anti-Soviet vote | Anti-Soviet protests | Anti-Russian vote | Anti- Yanukovich protests |
|------------------|-----------------------------------|---------------------|-------------------------|----------------------|---------------------------------|
| Weather index | 0/1 | 0/1 | 0/1 | 0/1 | 0/1 |
| Manual selection | 1/6 | 0/6 | 1/6 | 0/6 | 0/6 |
| BMA selection | 3/16 | 0/16 | 1/16 | 1/16 | 3/16 |
| LASSO selection | 1/19 | 1/19 | 2/19 | 3/19 | 1/19 |

Table 8.11: Additional placebo tests. Rows represent instrument selection methods, and columns represent outcomes. Each numerator is the number of placebo tests for a given instrument that succeeded, and the denominator is the the number of instruments (placebo tests that can succeed or fail).

the finding reported earlier: when we use weather index, there is only one placebo test for each outcome, and each of those placebo tests fail (show no placebo effect). The second row shows placebo tests for manually selected instruments: there were six instruments selected manually, and in two cases (column 1 and 3), one of those six instruments had a reduced form effect on political outcomes comparable to the respective effect in non-placebo case. Similarly, very few placebo effects were detected for instruments selected by Bayesian model selection (row 3) or by LASSO (row 4).

9. ROBUSTNESS CHECKS FOR SURVEY RESULTS

The current section evaluates the robustness of survey results in the main text to (a) alternative measures of opposition to pro-Russian armed groups, and (b) deviations from linearity assumptions.

9.1. *Alternative Measures of DNR/LNR Exposure*

Our survey results in the main text are based on an index measure of opposition to pro-Russian separatists (as we described in Section 2.2). Here, we check how the results change if we re-estimate our regressions using each of the four individual outcome measures that comprise this index. As we note in the main text, to make comparisons between DNR/LNR controlled areas most plausible, we restrict the sample to respondents in the conflict zone of Donets’k and Luhans’k oblasts.

The results are in Table 9.12. It is clear from the reported number of observations that questions on support for ‘DNR/LNR’ and support for ‘Russian army’ have a very large proportion of missing values. Using these measures alone would therefore be at best a very noisy measure of support for pro-Russian separatists. Accordingly, the coefficient estimates are small and not significant. When we change the outcome variables to support for the ‘Right Sector’ and ‘Ukrainian army,’ response rates are much higher, and coefficient estimates in both cases match closely the results from our index-based measure. These estimates suggest that the index-based results we report in the text are largely driven by the latter two variables, which – reassuringly – have fewer measurement issues than the former two.

| | DNR/LNR | Russian army | Right sector | Ukrainian army | Index |
|--------------------------------|--------------|--------------|---------------|----------------|---------------------------|
| Family losses | 0.08 (0.13) | 0.15 (0.11) | 0.31 (0.13)* | 0.57 (0.13)** | 0.34 (0.12)** |
| Family losses \times DNR/LNR | -0.16 (0.38) | -0.01 (0.34) | -0.48 (0.24)* | -0.59 (0.51) | -0.46 (0.27) [†] |
| DNR/LNR | -0.01 (0.35) | 0.02 (0.34) | -0.17 (0.14) | -0.27 (0.37) | 0.34 (0.38) |
| Observations | 446 | 510 | 773 | 767 | 832 |

OLS coefficients with standard errors in parentheses clustered by 1933 rayons. Significance levels: [†] $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 9.12: Family-level effects of famine losses by coercion threat on each outcome individually. The last column (Index) shows results for the index of opposition to pro-Russian separatism after adjusting for pre-famine covariates and regional fixed effects.

In the last column of Table 9.12, we also report results of an OLS regression that adjust for all pre-famine covariates used in our IV regressions and oblast level fixed effects. The

key estimates are slightly smaller and noisier than the ones reported in the paper, but the overall conclusion remains the same: the point estimate for non-DNR/LNR areas is positive (0.34) and significant, whereas the point estimate for non-DNR/LNR areas is negative ($0.34 - 0.46 = -0.12$) and not significant (the difference between the two coefficients is significant at 95 percent confidence level).

9.2. Violations of the Linearity Assumption

Our main analysis presumed linearity in how exposure to DNR/LNR violence moderates the effect of individual famine losses. This might not be plausible. We now conduct two tests to evaluate the robustness of our results to deviations from the linearity assumption.

First, we change our measure of exposure to DNR/LNR violence from days under DNR/LNR control to a dummy variable equal to one if the number of days under DNR/LNR control was more than x , and zero otherwise. We then estimate an OLS regression with opposition to pro-Russian separatists on the left-hand side and, on the right hand side, individual famine losses, the indicator of DNR/LNR control, and their interaction. Figure 9.4 reports the estimated marginal effects (with 95 percent confidence intervals, adjusted for clustering by rayons), for settlements above and below the cut-off. We iterated the procedure for ten cut-offs – the first ten unique day counts in the data. The marginal effect is always positive and significant in non-DNR/LNR controlled areas for all cut-offs, whereas it is never significant for DNR/LNR controlled areas. Moreover, differences between the two estimates become statistically significant once we set the cut-off at two days or more.

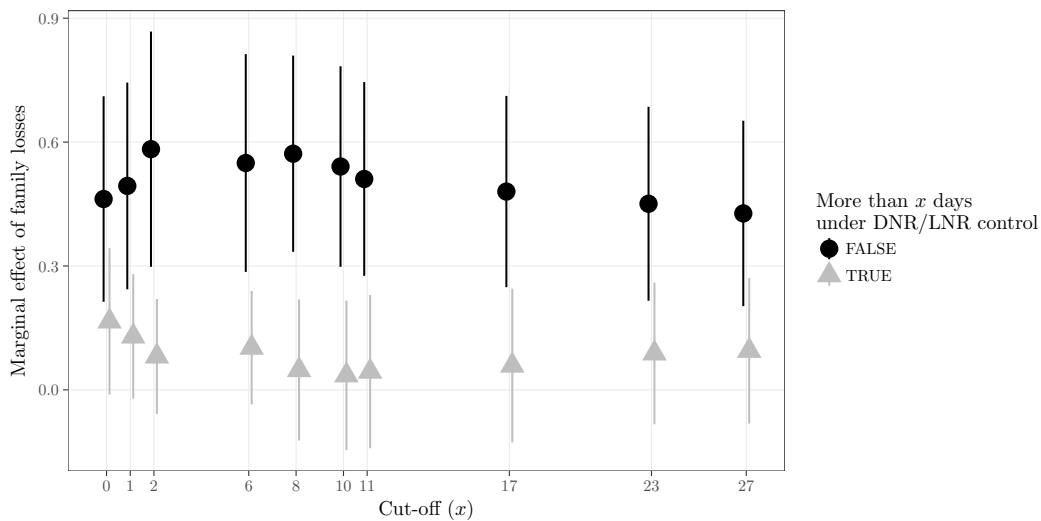


Figure 9.4: Differences in the effects of famine losses by DNR/LNR controlled areas using cut-offs.

Second, we estimate a more flexible model that allows the effect of individual famine losses to vary by exposure to days under DNR/LNR control non-linearly. We estimate this model in the generalized additive modeling framework, using tensor product smooth terms (Wood, 2008). Figure 9.5 shows how the marginal effects of individual famine losses change as a function of days under DNR/LNR control. In places that were not captured by DNR/LNR or were captured for a short time, the marginal effect is positive and significant, but it clearly attenuates as the length of time under DNR/LNR rule increases.

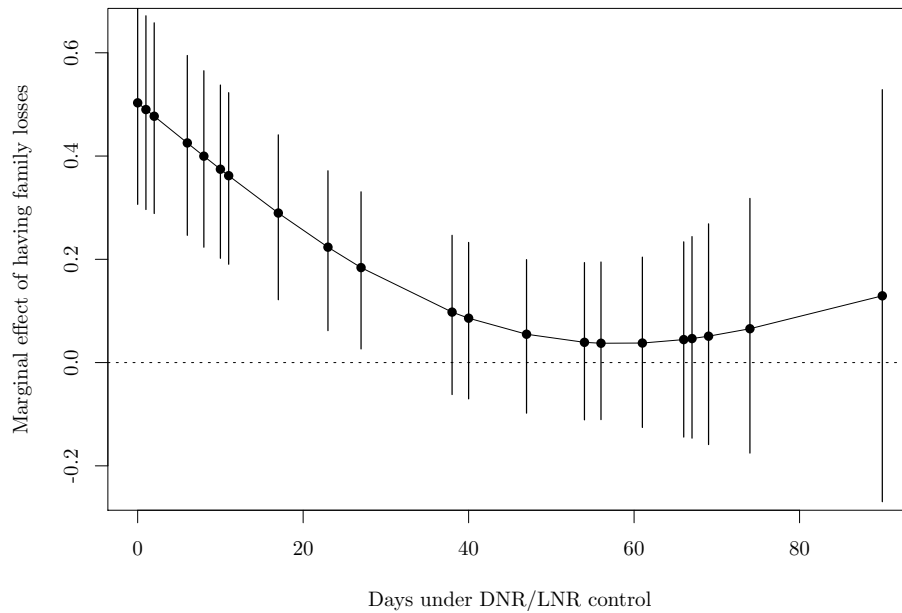


Figure 9.5: The marginal effect of family losses in famine conditional on days settlement was controlled by DNR/LNR separatist forces, estimated using interactive GAM. The 95 percent confidence bounds account for clustering by 1933 rayon.

Overall, the results are quite robust to how we operationalize exposure to DNR/LNR violence: linearly, in cut-offs, or smoothly but non-linearly.

9.3. *Alternative Estimators and Individual Level Demographics*

Because descendants of famine victims may differ from non-descendants in many ways other than famine exposure, we performed a matched analysis of the survey data, in which respondents differed in family history of famine, but were as similar to each other as possible on other observable dimensions. Our quantity of interest here is the Average Treatment Effect on the Treated (ATT), or the difference in potential outcomes associated with family exposure to famine, for respondents whose families were likely to have been exposed to famine,

$$ATT = E(E(Y_i|X_i, T_i = 1) - E(Y_i|X_i, T_i = 0) | T_i = 1)$$

where Y_i is respondent i 's expressed level of opposition to pro-Russian armed groups, $T_i = 1$ if i 's family member(s) died in the famine, and X_i is a matrix of observable pre-treatment covariates for i . Estimation of the ATT relies on the assumption that treatment assignment is unconfounded conditional on X , and that there is overlap in the support of X for the treated and control groups. The goal of matching is to re-weight the data, so that the remaining pairs of treated and control individuals (i.e. descendants and non-descendants) are as similar as possible on all covariates in X .

For consistency, we kept the set of pre-treatment covariates X the same in the matched analysis as in our baseline specification in the main text (i.e. oblast, proportion Ukrainian in 1926 census, proportion rural in 1926, forest cover, dominant crop, industrial production, total population in 1926, distance to Russia). Because individual demographics in 2017 are all technically post-treatment, we exclude them from our core specification. However, we compare our main matched results to estimates with an expanded set of covariates that also includes individual characteristics (sex, age, employment, education, marital status, preferred language, household size).

Table 9.13 reports ATTs from an exact matching estimator, which has the strictest standard for covariate imbalance, and requires that all matched pairs have identical values on all covariates X . We ran this estimator on the full survey dataset ('Overall'), and on subsets of respondents who lived on territories DNR/LNR once controlled ('DNR/LNR') and never controlled ('no DNR/LNR'). These results are consistent with those reported in the main text: family famine exposure has a positive effect on opposition to pro-Russian groups, but this effect is significant only in areas the DNR/LNR never controlled.

Estimates that use an expanded set of covariates are similar in direction and significance to those that use only our baseline specification of pre-famine covariates. Not surprisingly, however, the matched sample size is considerably smaller for the expanded

Table 9.13: MATCHED ANALYSIS: EFFECT OF FAMINE EXPOSURE ON OPPOSITION TO PRO-RUSSIAN GROUPS. Exact matching estimator. ATT: average treatment effect on the treated. SE: standard error. Imbalance: mean standardized difference between covariates in the treated and control units, before and after matching.

| Subgroup | ATT | SE | Obs | Matched | Imbalance_Pre | Imbalance_Post |
|--|-------|----------|------|---------|---------------|----------------|
| <i>Pre-famine covariates</i> | | | | | | |
| No DNR/LNR | 0.187 | (0.097)' | 608 | 146 | 0.10 | 0.00 |
| DNR/LNR | 0.135 | (0.113) | 574 | 91 | 0.07 | 0.00 |
| Overall | 0.161 | (0.073)* | 1182 | 237 | 0.15 | 0.00 |
| <i>Pre-famine covariates + individual demographics</i> | | | | | | |
| No DNR/LNR | 0.133 | (0.055)* | 602 | 5 | 0.12 | 0.00 |
| DNR/LNR | 0.119 | (0.089) | 570 | 14 | 0.09 | 0.00 |
| Overall | 0.09 | (0.05)' | 1172 | 20 | 0.14 | 0.00 |

*'p < .10, *p < .05, **p < .01, ***p < .001*

set of covariates – due to the increased number of variables on which to match exactly, and the continuous nature of many of the new variables.

While the exact matching routine has the advantage of minimizing covariate imbalance, it also assigns a weight of zero to a large number of non-matched observations. To ensure that our results are not unique to a single matching procedure, we estimated the ATTs across several alternative matching routines:

- *Propensity score matching* minimizes imbalance on a scalar, representing the probability of assignment to treatment. Formally, the propensity score is $Pr(T_i = 1|X_i) = E(T_i|X_i)$, where T_i is treatment assignment and X_i is the matrix of covariates.
- *Mahalanobis distance matching* minimizes the multivariate distance between column vectors, defined as

$$md(X_i, X_j) = \sqrt{(X_i - X_j)' S^{-1} (X_i - X_j)}$$

where S is the variance-covariance matrix of X .

- *Genetic matching* is a generalization of the Mahalanobis metric,

$$d(X_i, X_j) = \sqrt{(X_i - X_j)' S^{-1} W S^{-1/2} (X_i - X_j)}$$

where W is a $k \times k$ weight matrix and $S^{1/2}$ is the Cholesky decomposition of variance-covariance matrix S . A genetic search algorithm optimizes the weight each covariate receives in W , to find a set of matches that minimize the discrepancy between the distribution of potential confounders in treated and control groups (Sekhon, 2011).

Table 9.14: ALTERNATIVE MATCHING ESTIMATORS: EFFECT OF FAMINE EXPOSURE ON OPPOSITION TO PRO-RUSSIAN GROUPS. ATT: average treatment effect on the treated. SE: standard error. Imbalance: mean standardized difference between covariates in the treated and control units, before and after matching.

| Subgroup | ATT | SE | Obs | Matched | Imbalance_Pre | Imbalance_Post |
|-----------------------------|-------|----------|------|---------|---------------|----------------|
| <i>Propensity score</i> | | | | | | |
| No DNR/LNR | 0.187 | (0.096)* | 608 | 146 | 0.10 | 0.01 |
| DNR/LNR | 0.135 | (0.11) | 574 | 91 | 0.07 | 0.00 |
| Overall | 0.174 | (0.073)* | 1182 | 237 | 0.15 | 0.01 |
| <i>Mahalanobis distance</i> | | | | | | |
| No DNR/LNR | 0.187 | (0.096)' | 608 | 146 | 0.10 | 0.00 |
| DNR/LNR | 0.135 | (0.11) | 574 | 91 | 0.07 | 0.00 |
| Overall | 0.161 | (0.072)* | 1182 | 237 | 0.15 | 0.00 |
| <i>Genetic matching</i> | | | | | | |
| No DNR/LNR | 0.187 | (0.095)* | 608 | 146 | 0.10 | 0.00 |
| DNR/LNR | 0.135 | (0.107) | 574 | 91 | 0.07 | 0.00 |
| Overall | 0.161 | (0.071)* | 1182 | 237 | 0.15 | 0.00 |

['] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 9.14 reports ATTs from these additional tests. The results are quite stable across the estimators, and very similar to the exact matched estimates in Table 9.14.

9.4. Sample Attrition

There is a reasonable possibility that non-ignorable self-selection into the sample affects our survey results. For example, people with less pro-Russian preferences may have migrated away from towns the DNR/LNR had controlled. If so, then differences in the effects of family losses inside and outside DNR/LNR areas might be due to non-random sample attrition in these areas. To understand how exactly this kind of attrition might bias our results, we first formalize the problem.

Let $y_i^* \in \{0, 1\}$ denote a variable indicating respondent's i political preference, where $y_i^* = 0$ denotes support for pro-Russian armed groups and $y_i^* = 1$ denotes support for anti-Russian armed groups. Let $\mathbb{E}(y_i^* | x_i, z_i)$ denote the expected political preference of respondent i , given the binary indicator of him having family members who died in famine ($x_i = 1$) or not ($x_i = 0$) and given that he resides in formerly DNR/LNR controlled areas ($z_i = 1$) or not ($z_i = 0$). In a random sample without attrition, we would observe the variable y_i^* and would then estimate the quantity of interest, the difference between those who have and don't have family losses (inside DNR/LNR areas):

$$\Delta^* = \mathbb{E}(y_i^* | x_i = 1, z_i = 1) - \mathbb{E}(y_i^* | x_i = 0, z_i = 1), \quad (11)$$

However, if anti-Russian respondents are more likely to leave DNR/LNR controlled areas, then we do not have a random sample. Let y_i denote a respondent's i political preference in a sample contaminated by non-random attrition. Formally, let

$$\Pr(y_i = 1 | y_i^* = 0, x_i, z_i = 1) = 0, \quad (12)$$

$$\Pr(y_i = 1 | y_i^* = 1, x_i, z_i = 1) = 1 - a, \quad (13)$$

so that people with pro-Russian preferences ($y_i^* = 0$) remain in DNR/LNR controlled areas whereas a proportion of people with anti-Russian preferences ($y_i^* = 1$) leave those areas.¹² Given this potentially biased sample, our estimates in the paper are given by

$$\Delta = \mathbb{E}(y_i | x_i = 1, z_i = 1) - \mathbb{E}(y_i | x_i = 0, z_i = 1), \quad (14)$$

and so it is clear that $\Delta \neq \Delta^*$ unless $a = 0$ (no self-selected attrition). To see the relationship between the two estimands, use the law of iterated expectations so that:

$$\begin{aligned} \mathbb{E}(y_i = 1 | x_i, z_i = 1) &= \Pr(y_i = 1 | x_i, z_i = 1), \\ &= \Pr(y_i = 1 | y_i^* = 0, x_i, z_i = 1) \Pr(y_i^* = 0 | x_i, z_i = 1) \\ &\quad + \Pr(y_i = 1 | y_i^* = 1, x_i, z_i = 1) \Pr(y_i^* = 1 | x_i, z_i = 1), \\ &= (1 - a) \Pr(y_i^* = 1 | x_i, z_i = 1), \\ &= (1 - a) \mathbb{E}(y_i^* | x_i, z_i = 1), \end{aligned}$$

which implies that

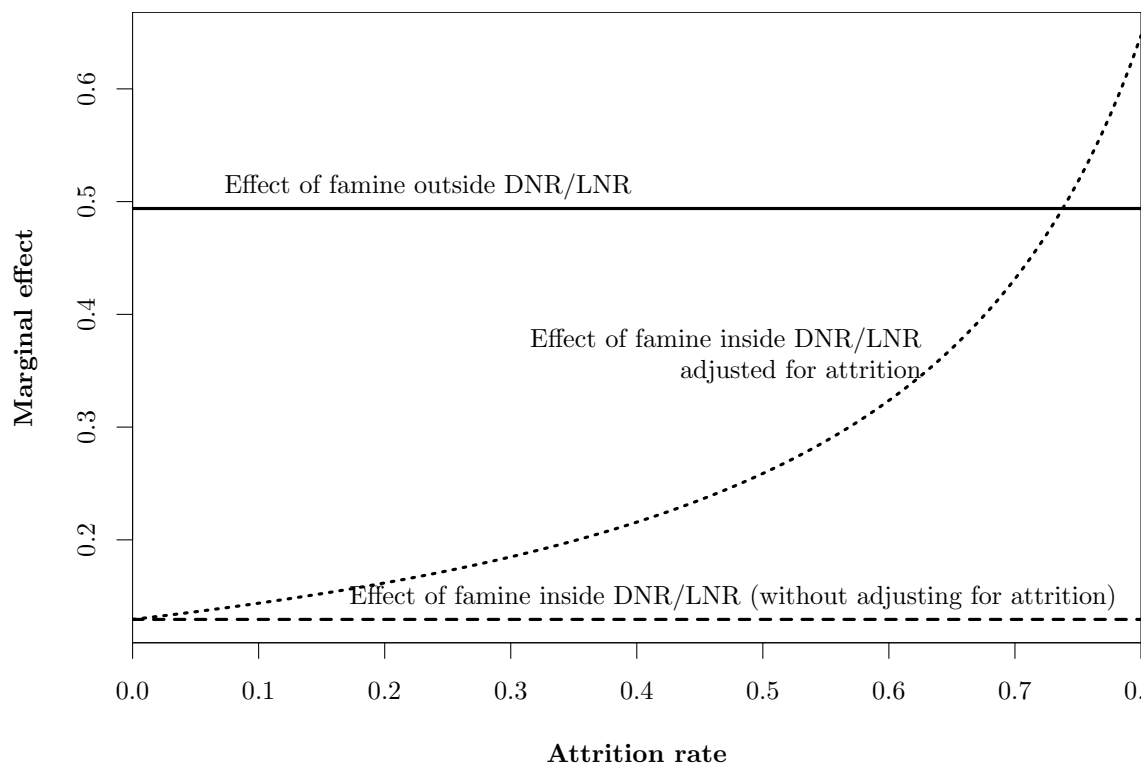
$$\Delta^* = \frac{1}{1 - a} \Delta. \quad (15)$$

So our estimate in the self-selected sample, Δ has an attenuation bias equal to $-\frac{a}{1-a} \Delta$.

Given the above result, in those cases where our coefficient estimates for DNR/LNR areas are negative, the attenuation bias is positive, which means that sample attrition *weakens* heterogeneity. For example, in the regression results reported in the paper (Table 4 column 3), the effect for family losses in DNR/LNR areas is $0.46 - 0.60 = -0.16$. If, say, 20 percent of people with anti-Russian preferences have left DNR/LNR controlled areas, then the implied coefficient without attrition would be $-0.16 / (1 - 0.25) = 0.2$, and we would observe a larger deterrent effect of famine losses on expressed attitudes in

¹²We could assume that some proportion of people with pro-Russian preference also migrate, but the only thing that matters is the difference in the probabilities of attribution between pro-Russian and anti-Russian respondents, so this would not change the results.

DNR/LNR controlled areas than we do in the sample with attrition.



The two horizontal lines display the reported point estimates of having family members died in famine on anti-Russian attitude index. The dotted curve represents the adjusted estimate that accounts for attrition rate of anti-Russian respondents from DNR/LNR controlled areas.

Figure 9.6: Sensitivity of estimates to attrition of anti-Russian respondents from DNR/LNR controlled areas.

However, as additional robustness checks indicate, sometimes the coefficient estimates for family losses in DNR/LNR controlled areas are positive (but never significant). In that case, the differences in estimates between inside and outside DNR/LNR areas could be attributed to sample attrition. For example, consider our specification where we use day cut-offs to identify DNR/LNR areas (Section 9.2). If we use a one day cut-off, then the estimate of famine losses is 0.51 outside DNR/LNR areas (significant) and 0.13 inside DNR/LNR areas (not significant).

How large would the sample attrition need to be to produce such a difference in coefficients? In Figure 9.6 we show the imputed estimate of Δ^* for different values of a , and compare them with the estimates in our sample. We see that the effect of famine on holding anti-Russian attitudes is smaller inside DNR/LNR areas than outside DNR/LNR areas for a very wide range of a values. To account for difference in point estimates this large, 70 percent or more of potential respondents with anti-Russian attitudes would need

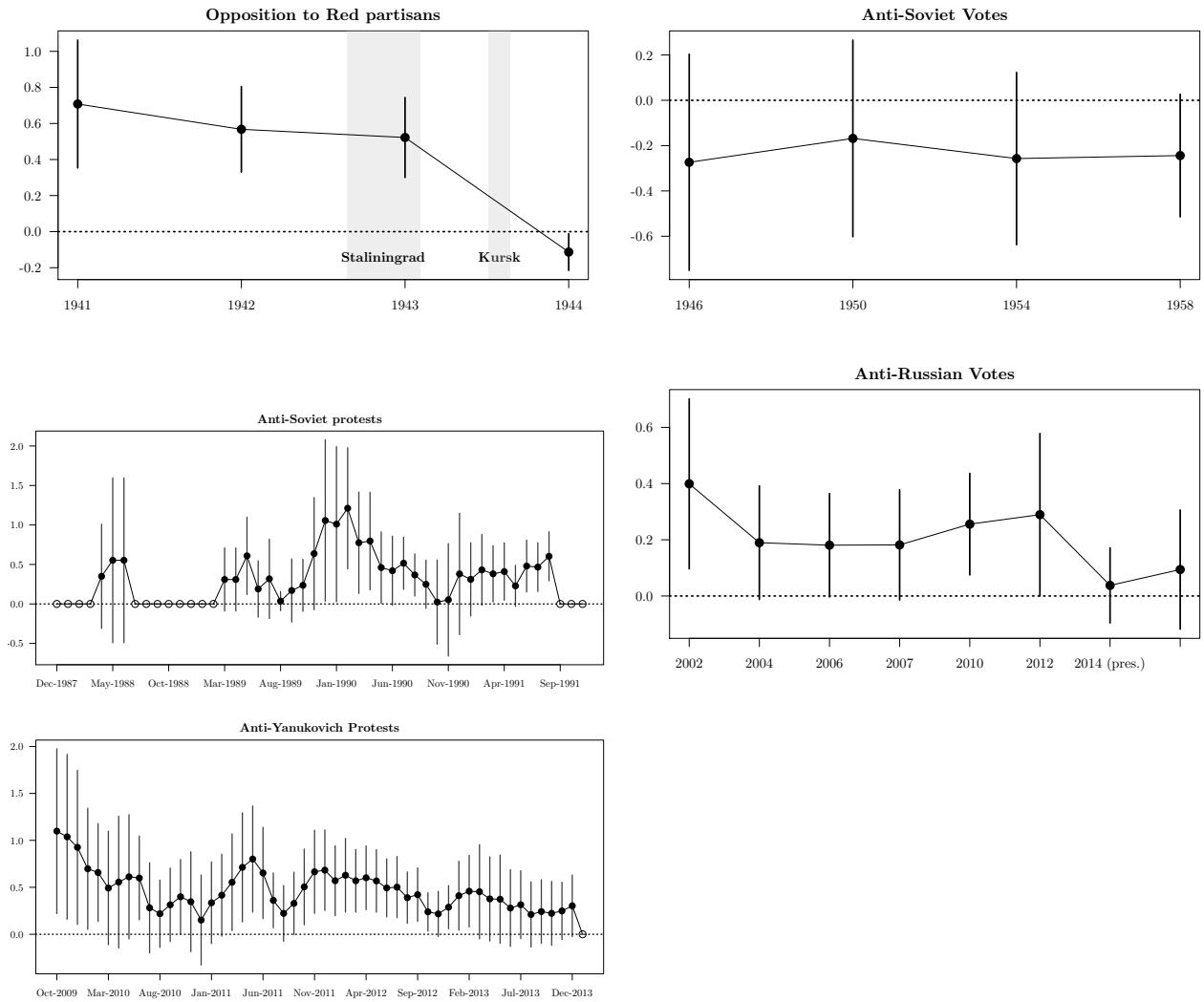
move out from formerly DNR/LNR controlled areas. The degree of attrition would need to be even larger to explain differences in estimates when longer day cut-offs are used.

A quick back-of-the-envelope calculation can demonstrate that 70 percent is an implausibly large attrition rate for our sample. Estimates of total Ukrainian internally displaced persons in 2017, who fled their homes due to either the conflict in the Donbas or Russia's annexation of Crimea range from 800,000 to 1,800,000.¹³ If we assume, for the sake of argument, that all of these IDPs came from the Donbas rather than Crimea, this number would still represent only 12 to 27 percent of the pre-war population of Donetsk and Luhans'k oblasts (6.6 million). Even if we assume that individuals with anti-Russian views were equally likely to have fled from territories still controlled by DNR/LNR as from the government-liberated territories in our sample, this attrition rate is still well below the 70 percent needed to explain the heterogeneity of famine effects.

¹³<http://www.internal-displacement.org/database/displacement-data>,
<http://www.unhcr.org/ua/wp-content/uploads/sites/38/2018/06/2018-03-UNHCR-UKRAINE-Operational-Update-FINAL-UA.pdf>

10. DISAGGREGATED SECOND-STAGE RESULTS

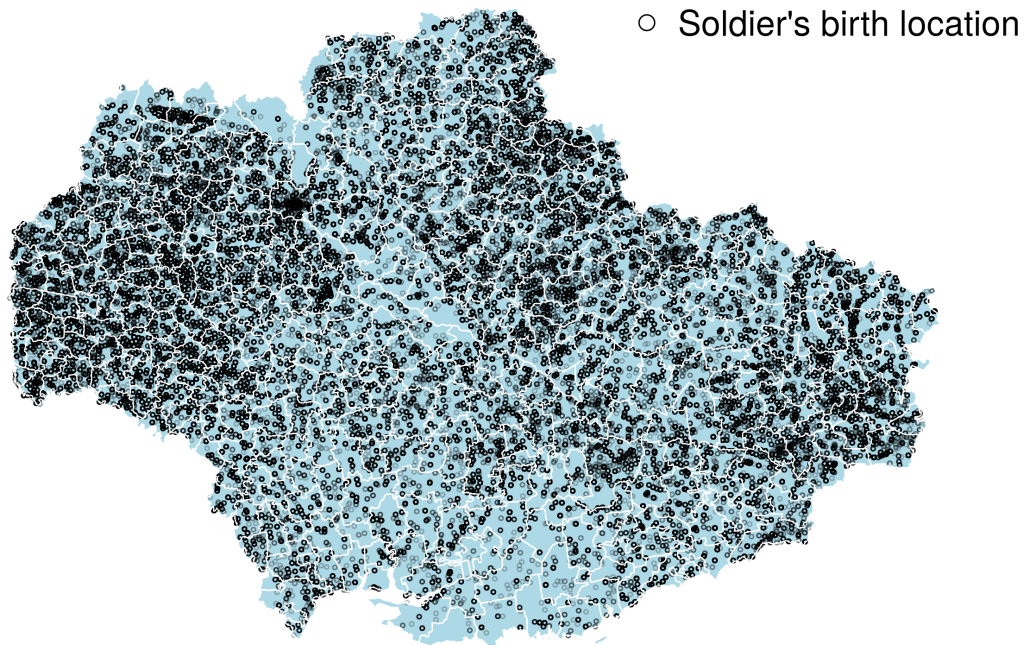
Figure 10.7: The effects of famine on behavioral opposition to Moscow, disaggregated by time. Estimates are second stage coefficients with 95% confidence intervals from either our baseline IV specification with outcome data for individual years (partisans, elections), or from the baseline model fit on a three month window rolled forward by 30 days (protests). Hollow circles are for windows with insufficient data. The outcome variables are standardized within each election.



11. EVIDENCE FROM BATTLEFIELD BEHAVIOR IN WWII

To evaluate the relationship between famine and the battlefield performance in WWII, we collected the personnel records of 1,048,986 Red Army soldiers and officers born in Soviet Ukraine. We obtained these records through the *Pamyat' Naroda* database maintained by the Central Archives of the Ministry of Defense of the Russian Federation (TsAMO RF, <https://pamyat-naroda.ru>). We obtained geographic coordinates for the birth addresses of these individuals, using Yandex's geo-coding API. Figure 11.8 shows the spatial distribution of these birth locations, against 1933 rayon borders.

Figure 11.8: Geographic distribution of Ukrainian Red Army personnel in WWII, by location of birth. Data from <https://pamyat-naroda.ru>.



After matching individuals' birth locations to 1933-era administrative units, we calculated *Battlefield disloyalty* scores for each rayon, defined as the proportion of soldiers born in the rayon who either deserted, defected, committed treason, were missing in action, or became prisoners of war. We also calculated rayon-level *Battlefield sacrifice* scores, defined as the share of soldiers who were either killed or wounded in action. We then used these two measures as outcome variables in our core TSLS model. The first-stage statistics, IV coefficients, and second-stage spatial residual correlation statistics are shown in Table 11.15.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|------------------------|---------------------------|----------------------------|------------------------|
| Battlefield disloyalty | 0.41 (0.21) [†] | 12.04 | 0.48 |
| Battlefield sacrifice | -0.51 (0.26) [†] | 11.89 | 0.49 |

The estimates are for logged famine mortality instrumented by weather adversity index. All specifications control for pre-famine rural population, proportion of Ukrainians, proportion of Russians, forestation, industrialization, crop suitability, oblast fixed effects, synthetic spatial covariates, and cubic spline for the distance to Russian border. Significance levels (two-tailed): [†] $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 11.15: IV RESULTS FOR WWII BATTLEFIELD BEHAVIORS.

The results shown in the paper not only alleviate the concern that our earlier WWII results were driven by German propaganda, but they also alleviate potential concerns that our earlier measures of political disloyalty during the WWII are not sufficiently refined to capture the phenomenon.

12. ALTERNATIVE EXPLANATIONS FOR SOVIET ELECTIONS

12.1. Propaganda

Another possible explanation for the results concerning Soviet elections is that the Communist Party and state apparatus used more propaganda and vote-mobilization in famine-ridden areas, which could explain why fewer anti-Soviet votes were cast in those areas. Although our reading of the historical literature has not indicated this possibility,¹⁴ this hypothesis makes theoretical sense and should be investigated.

To measure state propaganda efforts during Soviet times, we exploit the fact that the Soviets relied heavily on iconography and symbolism (Kenez, 1985; Bonnell, 1999). The most focal measurable artifact of state propaganda were Lenin’s statues. We assembled a database of Lenin’s statues built on the territory of Ukraine (except Western Ukraine) during the post-war period. We started building this dataset using an online crowdsourced database that collects information on the demolitions of Soviet statues (www.leninopad.ru), which we checked against and complemented with information from online news sources and a compendium of Soviet monuments collected by the Ukrainian government as part of its implementation of the “de-communization” law.

We identified 2,098 statues of Lenin built after 1945 for which we could credibly determine a location. We then calculated the total sum of Lenin’s statues per 1933 rayon, and used this variable to measure the intensity of Soviet propaganda efforts after WWII. For robustness, we also consider the proportion of Lenin’s statues relative to pre-famine (1926) population as an alternative measure. We then replicate our baseline IV regressions using all covariates, oblast fixed effects, and cubic regression splines for distance to Russia, but use Lenin’s statues as a dependent variable. If the Soviets did use more propaganda in famine-ridden areas, we should expect to see a positive association between the famine deaths and the number (or proportion) of Lenin’s statues.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran’s I (p-value) |
|------------------------------|--------------------------|----------------------------|------------------------|
| Lenin’s statues (count) | 0.04 (0.38) | 12.20 | 1.00 |
| Lenin’s statues (proportion) | −0.38 (0.19)* | 11.47 | 0.83 |

Significance levels (two-tailed): † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 12.16: Famine mortality and post-war Lenin statues – IV estimates.

¹⁴It is very likely that local party and state officials did not possess highly sensitive figures on famine deaths to match their propaganda efforts to these figures.

The results are shown in Table 12.16. The coefficient for the count of Lenin’s statues is essentially zero, whereas the coefficient for the proportion of statues is negative. Thus, if anything, the evidence suggests *fewer* propaganda efforts in famine-ridden areas. Clearly, monuments for Lenin constituted only one component of complex and extensive Soviet propaganda efforts, but the evidence here does not suggest Soviet propaganda to be the factor that explains why famine-ridden areas cast *fewer* anti-Soviet votes in Soviet times.

12.2. Electoral Mobilization

Another possibility is that the Soviet authorities (especially the party) engaged in greater mobilization efforts in famine-ridden areas. It is impossible to have a comprehensive measure of electoral mobilization (especially in an authoritarian context), but it is plausible to think that one important way the party would attempt to mobilize voters is by nominating higher-quality candidates. We used the same archival documents from which we extracted election data (referenced in the paper) to collect information about the personal characteristics of each Supreme Soviet candidate.

We focused on three qualities: First is whether the candidate is a member of the Communist party, because it is plausible that such a person would have an easier time accessing party resources for vote-mobilization than a non-party candidate. Second is whether the candidate is of Ukrainian nationality, because the candidates from the titular nation would likely have an easier time attracting local support. Third is whether the candidate is a decorated war veteran, which was (and remains) an important appealing quality after the devastating war. We also constructed an omnibus quality measure that takes a value of one if the candidate is a party member, Ukrainian, and decorated veteran. We then calculated rayon-level candidate quality averaged across each of the four variables, and over all four post-Soviet elections. We then regressed these variables on famine deaths, instrumented by weather, using all baseline covariates and fixed oblast effects.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran’s I (p-value) |
|--------------------|--------------------------|----------------------------|------------------------|
| Party member | −0.34 (0.31) | 12.18 | 0.38 |
| Decorated veteran | −0.14 (0.34) | 12.03 | 0.46 |
| Ukrainian | −0.37 (0.26) | 12.04 | 0.99 |
| Overall quality | −0.28 (0.31) | 12.04 | 0.85 |

Significance levels (two-tailed): † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 12.17: Famine mortality and quality of candidates in the post-war elections – IV estimates.

The results are reported in Table 12.17. There is no statistically significant relationship between any measures of the candidate quality, and in each case the estimated coefficients are negative. This suggests that, if anything, lower quality candidates were nominated in famine-ridden areas. In sum, we do not observe evidence consistent with propaganda or differential mobilization as an alternative explanation of our results.

12.3. Election Fraud

It is also possible that our results could be driven by increased election fraud in famine-ridden communities. While it cannot be ruled out conclusively, this explanation is not consistent with the existing evidence. Stalin saw elections as a “the trial court of the electors for the Communist party” (Carson, 1955, p. 93), designed to provide the central government with information on the performance of local officials (Getty, 1991), and uncover “covert opponents of the regime” (Zaslavsky and Brym, 1978, p. 369). Elections would not achieve any of these functions had local authorities been allowed to manipulate vote counts, and historical evidence of vote-falsification is scarce.¹⁵

To check empirically whether our results could be driven by election fraud in the Soviet elections, we conducted the last-digit test on these data (Beber and Scacco, 2012). The idea behind this test is that manipulated numbers often leave fingerprints (i.e., last digits contain excess zero’s or five’s). In non-manipulated elections, last digits should be distributed approximately uniformly.

We examined three variables: the total number of ballots cast, the number of ballots against the candidate, the number of ballots for the candidate. The results are in Figure 12.9. The p-values are for the test that the distribution of last digits is uniform. In each case, we cannot reject the null hypothesis the last digit follows uniform distribution.

While this test cannot rule out the possibility that election results were manipulated in some other way that cannot be detected by this test (no test can do that), it is reassuring that, at least as far as last digits go, we do not observe evidence of fraud.

12.4. Famine-Relief and Post-Famine Economic Investments

Another possibility is that the differences in political behavior are due to differential economic investments conditional on famine losses. Overall, the economic explanation is difficult to reconcile with the dynamics of the famine effects that we observe in the data.

¹⁵Vote-falsification was a criminal act under Soviet law (Friedgut, 2014, p. 116). At least in the early stages, “there appears to be little outright falsification of returns” Karklins (1986). “Partial falsification” had become “widespread and well known” in later periods of Soviet rule, after the 1946-1958 window (Zaslavsky and Brym, 1978, p. 369).

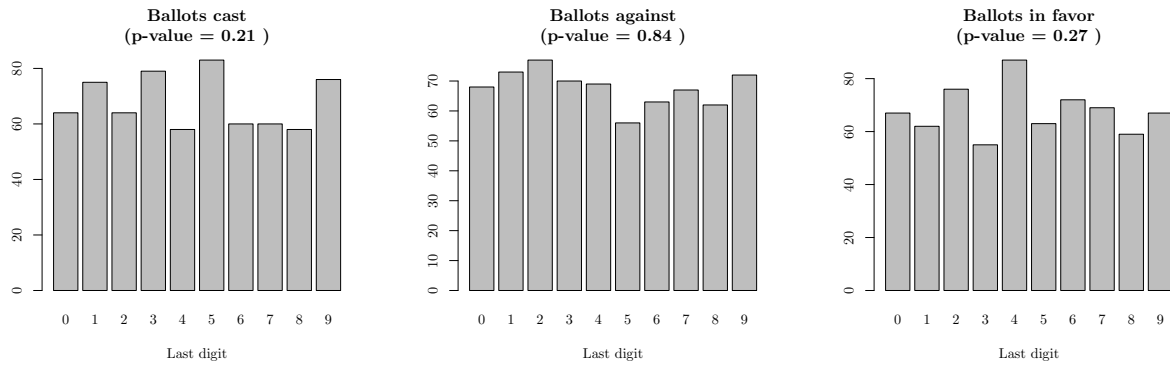


Figure 12.9: Distribution of last digits in 1946-1958 elections.

If the Soviet government invested more in famine-ridden areas right after the famine, then why did those areas behaved more anti-Soviet at the start of the war? If the Soviet government invested less in famine-ridden areas, then why did those areas behaved more pro-Soviet after the war? The very fact that the effects of famine changed in time poses a challenge to the economic explanation.

The only archival or historical evidence that we could find about the government's economic reactions to famine were often belated and short-lived relief efforts in the Spring of 1933. These were not humanitarian relief efforts, because the "main objective was not to prevent starvation but to provide badly needed aid in order to save the next sowing season" (Wolowyna et al., 2016, p. 193). The human resources were so depleted by famine that, in the absence of aid, there were too few hands available to work in agriculture, and the food aid program was to provide food mainly to members of the collective farms who were willing and able to work and the independent farmers willing to join collective farms. The food assistance program targeted more strategically and economically important areas like Odessa and Dnipropetrovsk oblasts (Boriak, 2012). While the death rates were lower in the areas that received food assistance, they were still "extremely high" (p. Wolowyna et al., 2016)

We believe that the food relief efforts cannot account for our results for several reasons. First, the food relief efforts were prioritized on the basis of oblasts. But our estimations use oblast fixed effects, which would partial out the effects of policies implemented at the oblast level. Second, it is not at all clear how the food relief efforts would explain changes in the famine's effect through time. Finally, because we are using an exogenous instrument for famine deaths, even if food efforts were politically motivated, it should not have an effect on our inferences.

What about the governments economic activities after famine? While we could not

obtain granular data on Soviet economic investments after the famine to rule out the economic explanation quantitatively, the circumstantial evidence does not support this case for two main reasons – the policy-makers lacked accurate information about the effects of famine and the centrally-dictated economic priorities, especially in the after-war era, overshadowed any potential effects of famine.

As argued in [Maskudov \(1989, p.272\)](#), “in 1930-1937, the [Soviet] government and even demographers did not know the actual consequences of the implemented social policies. They couldn’t help but mention that famine and mass mortality changed the demographic situation. But how and to what extent - they didn’t know.” The Soviet population census of 1937 could have helped gain this information, but there were many intervening factors that would make interpretation of the census results difficult: the ban on abortions, industrialization in Donbas, and rapid urbanization. The Council of the People’s Commissars of the USSR recognized the results of the census invalid on September 25, 1937, and demographers who conducted it were repressed ([Kulchytskiy, 2003, p.42](#)). Even if the government wanted to direct the pre-war or post-war rebuilding efforts to the regions that suffered most from the famine, it would be difficult to implement such policies due to the lack of information about the famine’s consequences at the local level.

Furthermore, the overall economic priorities of the Soviet authorities were dictated by macro-level economic and geopolitical ambitions of Moscow, and it is very unlikely that famine played a role in those plans, especially after the war. After regaining control of Ukraine in 1944, the government invested in the coal industry of Donbas, the Krivorozhsky iron ore basin, the southern metallurgical district, and the power plants all over Ukraine. Those economic priorities had little to do with the agricultural sector – the most relevant one in famine-ridden areas. After the war, the Soviet economic plans were focused on building large industrial enterprises, coal-metallurgical and military-industrial complexes, and the main priorities for economic development were decided by logistic-economic considerations (access to natural resources, water basins, roads, etc.) ([Kovalskyi, 2004](#)).

13. ALTERNATIVE EXPLANATIONS FOR POST-SOVIET PERIOD

13.1. *Propaganda and Mobilization*

Since Ukraine's independence, the debate over Holodomor focused on its 'genocidal' status. In September 1993, the first President of Ukraine Leonid Kravchuk stated: "I fully agree with the fact that it was a planned action, that it was a genocide against their own people. But I would not put a point here. So, against their own people, but according to a directive from another center. Obviously, this is the way to treat this horrible page of our history" ([International science conference, 1995](#), p. 10). Throughout the 1990s the debates were limited to academic conferences and Holodomor was not widely discussed politically.

In 2002, the government commemorated Holodomor and Political Repressions Memorial Day for the first time. On November 24, 2002 the second President of Ukraine Leonid Kuchma recognized Holodomor as a genocide, called for the construction of a Memorial to commemorate its victims in Kyiv, and asked the international community to recognize the Holodomor as a genocide of the Ukrainian people and a crime against humanity ([Ivanchenko, November 28, 2009](#)). Accordingly, four days later Verkhovna Rada of Ukraine condemned "the policy of genocide held at the state level by the leaders of the totalitarian Soviet regime against the citizens of Ukraine, the national spirit, mentality, and genetic fund of the Ukrainian people." On May 15, 2003, the Verkhovna Rada of Ukraine adopted a resolution announcing the famine of 1932-1933 as an act of genocide deliberately organized by the Soviet authorities against the Ukrainian people ([Verkhovna Rada Ukrainy, November 28, 2002](#)). Despite these events, the public political discussions of the Holodomor issue were sparse.

After the Orange Revolution, the salience of Holodomor increased very significantly, again predominantly due to the third President of Ukraine Viktor Yushchenko's position on this issue. From July 2005 to October 2006 he adopted three resolutions, aimed at Holodomor recognition and commemoration ([Ivanchenko, November 28, 2009](#)). On November 28, 2006, the Verkhovna Rada of Ukraine adopted a Law on the Holodomor of 1932-1933 in Ukraine, which recognized the Holodomor as a genocide of the Ukrainian people. The law was supported by two MPs from the Party of Regions, 118 from the Yulia Tymoshenko Bloc, 79 from Our Ukraine (Nasha Ukraina), 30 MPs of the Socialist Party, 4 non-factional MPs, and 0 communists ([Verkhovna Rada Ukrainy, November 28, 2006a](#)). The law was part of President Viktor Yushchenko's Holodomor recognition campaign, which also included construction of memorials and holding special events. The transcript of parliamentary debates suggest that supporting parties (especially Nasha Ukraina, the

party of Viktor Yushchenko) framed the discussion in the context of crimes against humanity and solidarity with foreign parliaments that recognized Holodomor as a genocide. The President's Representative, Igor Yukhivskiy, associated Holodomor with Stalin's repressions and named it the primary cause of current demographic problems (*Verkhovna Rada Ukrainy, November 28, 2006b*).

In their turn, the Communist Party of Ukraine disseminated to the Parliament and the general public a brochure titled "The myth of Holodomor is the invention of mind manipulators," which claimed the Holodomor was an invention of Harvard speculators and American spies. The brochure attributed food shortages to the mass-scale reform of agriculture in 1932-1933 and unfamiliarity with new agricultural technologies (*Tkachenko, November 20, 2008*). This was an indication of the Holodomor becoming a wedge issue.

After Viktor Yanukovich became prime-minister in 2007 and then president in 2010, the government's position on the Holodomor issue changed again. The Holodomor section was removed from the President's website, and *The Book of Memory* (a document listing Holodomor victims by name) became unavailable online. The President's speaker, Anna Herman, announced that the information on Holodomor will not be updated, since this topic is well-researched (*Viche, November, 2010*). On April 27, 2010, while speaking in Strasbourg to the PACE, Yanukovich stated that "it would be wrong and unfair to recognize the Holodomor as a fact of genocide against one or another people." According to him, it was a common tragedy of the states that were part of the USSR (*Pravda, April 27, 2010*). At that time, the minister of education, Dmytro Tabachnyk, also supported Yanukovich's position by claiming that Holodomor should be removed as a topic from history textbooks as "hyperbolic nonsense." Notably, Tabachnyk's views on Holodomor were inconsistent: in 2003 when speaking to the Ukrainian Parliament he called Holodomor the "Ukrainian Holocaust" (*Viche, November, 2010*).

For current president Petro Poroshenko, Holodomor is a highly salient issue. He regularly participates in commemoration events, followed by reports on his Facebook page. According to Poroshenko, Holodomor is the historic illustration of Russian aggression against Ukraine. Poroshenko also actively promotes awareness of Holodomor abroad, especially in the US, where he demanded official recognition from Congress (*Pravda, June 22, 2017*). That said, Holodomor is not mentioned in the party platform documents of Petro Poroshenko's Bloc, similarly to all other major Ukrainian political parties.

The only party that continues to actively mention Holodomor in its election program is the right-wing All-Ukrainian Union 'Svoboda.' The party's goal is the recognition of Holodomor as genocide against Ukrainians by the UN and the EU Parliament. Svoboda devotes a paragraph of its platform to the necessity of criminal prosecution of crimes

committed during Holodomor and after (including its denial) (Svoboda, August 12, 2009).

All of the above indicates that Holodomor is clearly a politicized issue in Ukraine. The question is whether the *dynamics* of its politicization can explain our findings. Put more concretely, is it the case that the effects of famine become more pronounced when the issue becomes more politicized?

To answer this question systemically, we need to quantify the salience of the famine in Ukrainian politics, and see if it can account for variation in coefficient estimates over time. We measure the salience of Holodomor in Ukraine in two ways.

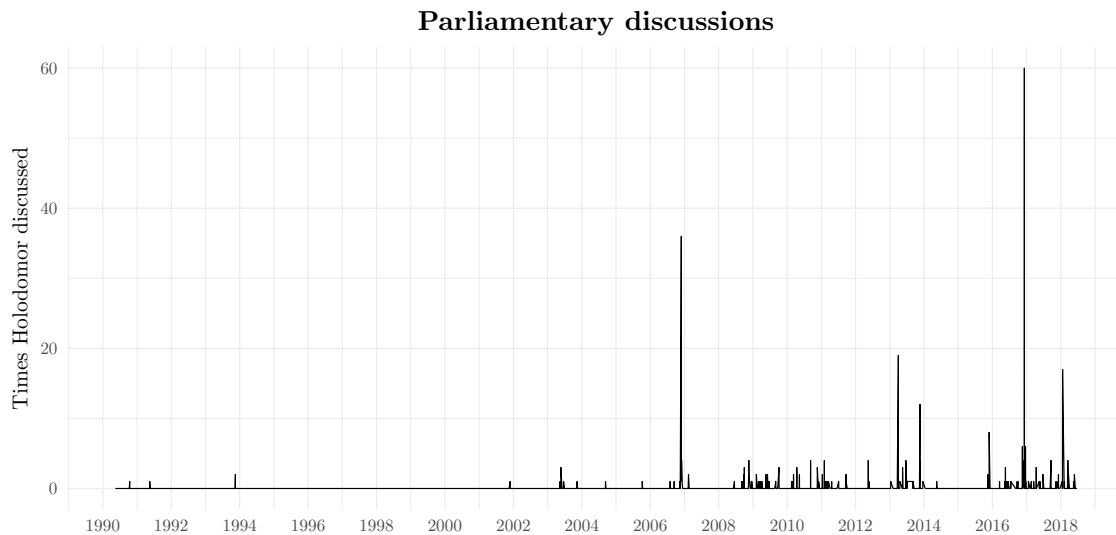


Figure 13.10: Salience of famine in the parliamentary debates.

First, we consider how frequently the famine is mentioned during floor debates in Rada, the Ukrainian parliament. For that purpose, we collected transcripts of floor debates from 1990 to 2017, and calculated how many times the term ‘Holodomor’ appears in the parliamentary record. The results are shown in Figure 13.10. We barely see any real extensive discussions of the famine during the 1990’s and even the early 2000’s. The issue becomes politicized in earnest only around 2006-2007 and again after 2009.

Our second measure of salience looks at how frequently Ukrainian media report on the topic of famine. Ideally, we would have liked to measure how much this topic is covered on Ukrainian television – the most widely consumed medium. Unfortunately, we were able to obtain television data only going back to 2011, which is insufficient for our test. We instead obtained transcripts from four Ukrainian newspapers: the daily *Den*, a pro-western Kyiv-based daily, *Fakty i Kommentarii*, biggest Russian language tabloid, *Ukrayinska Pravda*, a pro-western Ukrainian-speaking internet news site, and *Dzerkalo Tyzhnia*, an influential analytical weekly. Figure 13.11 shows media mentions of Holodomor from

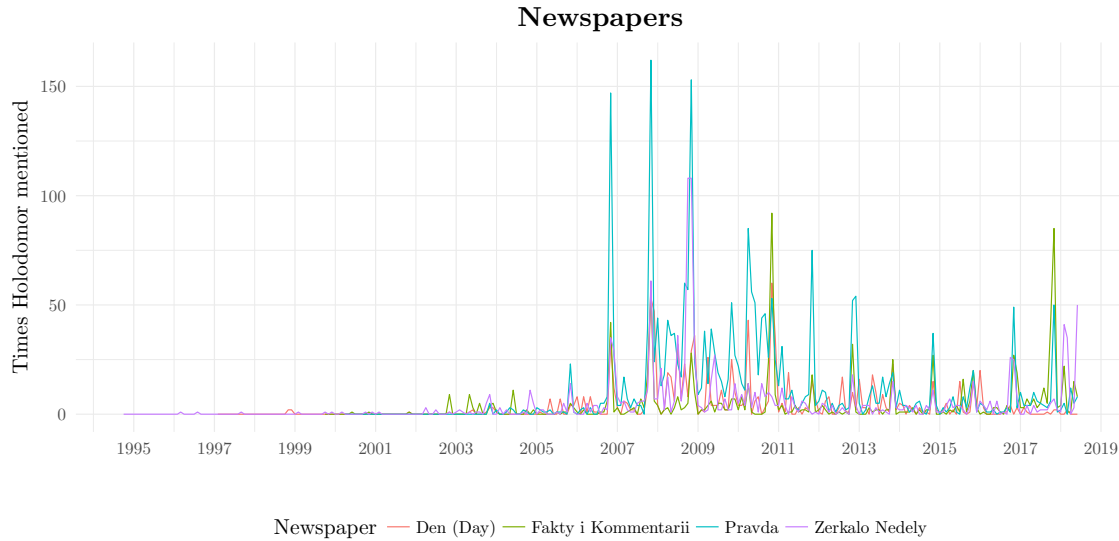


Figure 13.11: Mentioning of the famine in Ukrainian newspapers.

1995 to 2018. The famine really started being discussed on Ukrainian media (both Ukrainian-speaking and Russian-speaking) in earnest only from 2007 onwards, similarly to its discussion on the parliament floor.

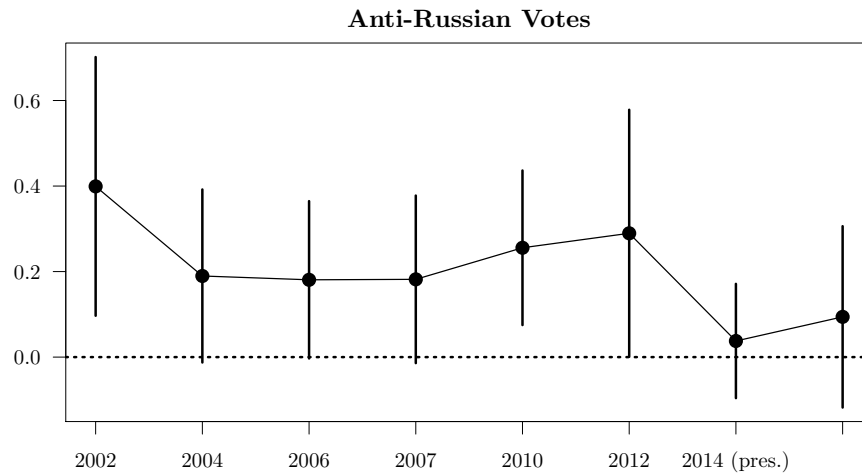


Figure 13.12: The effects of famine on pro-Russian votes in each 2002-2014 election separately. The outcome variables are standardized within each election.

If the inflammatory effects of famine in independent Ukraine are due to endogenous salience of this issue in Ukrainian politics, then these effects should be more pronounced in time periods when the Holodomor topic was most salient: in 2007-2009, and after 2013. To see whether that is the case, we estimated effects of famine (using our baseline specifi-

cations) on *Anti-Russian* votes in 2002-2014 by each election separately. The outcomes are normalized as standard deviations from that year's mean so that we can meaningfully compare the magnitudes of effects across different elections. These coefficient estimates and their 95 percent confidence intervals are shown in Figure 13.12.

Variation in the political salience of Holodomor, as measured in Figures 13.10 and 13.11, cannot explain why the effects of famine vary as they do in Figure 13.12. We see that famine had an effect on voting behavior already in 2002 and 2004, before Holodomor became a widely discussed topic in Ukrainian politics and media (in fact, the largest point-estimate is from 2002). In turn, the weakest effects are in the two 2014 elections, when the topic of Holodomor was highly salient in media and politics. We cannot tell what the effect of the famine would be had it been completely ignored in Ukrainian politics, but variation in the salience of famine in politics does not appear to explain variation in its effect.

13.2. *Post-Famine Ethnic Composition*

We now consider whether the famine's effects on local ethnic composition might explain our results, in the short and long run. For the short-term effect, we examine the effect of famine on the percent of individuals who self-identify as being of Ukrainian or Russian nationality in the 1939 Soviet census (Central Statistical Directorate of USSR, 1939). Because Soviet authorities systematically manipulated the 1939 census to hide the famine's extent (Zhiromskaia, 1990), we use adjusted census figures that remove these biases (Rudnytskyi et al., 2015).

For the long-term effect, we examine famine's effect on the percent of self-identified Ukrainian and Russian speakers in the 2001 Ukrainian census.¹⁶ We do the same with 2001 census data on self-identified nationality.¹⁷

In Table 13.18, we show the estimated causal effects of famine on post-famine ethnic composition, both short-term and long-term. Consider the first two rows of the table: the estimated coefficients of the famine effect are very small and not significant, indicating that famine did not have a short-term effect on the ethnic composition. The average local proportions of Ukrainians and Russians in 1939 did not change relative to 1926 as a result

¹⁶The census provides information on this variable at the level of settlement (village, town, or city). We geocoded these settlements, spatially matched them to 1933 rayons, and calculated rayon-level averages of Ukrainian and Russian-speakers. Unfortunately, we were not able to obtain ethnic composition figures at an equally high geographic resolution from the intervening Soviet censuses of 1959, 1970, and 1989.

¹⁷This variable is publicly reported only at the level of present-day rayons and cities (but not smaller towns or villages). Because the borders of 1933 and 2001 rayons do not perfectly align, we matched the nationality data to 1933 rayons using the same areal weighting procedure we used for other polygon data (e.g. rainfall), as we described in Section 1.2 and Figure 1.1b.

of famine. Therefore, it is extremely unlikely that, for example, the Red partisans had more success operating outside famine-ridden areas at the start of the war (and less success at the end of the war) because of the effects of famine on the local ethnic composition.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|-----------------------------------|--------------------------|----------------------------|------------------------|
| Percent Ukrainians (1939) | -0.01 (0.07) | 11.87 | 0.66 |
| Percent Russians (1939) | -0.02 (0.09) | 11.93 | 0.86 |
| Percent Ukrainian-speakers (2001) | 0.45 (0.11)** | 11.68 | 0.87 |
| Percent Russian-speakers (2001) | -0.48 (0.09)** | 11.60 | 0.87 |
| Percent Ukrainians (2001) | 0.23 (0.11)* | 12.01 | 0.46 |
| Percent Russians (2001) | -0.33 (0.12)** | 11.43 | 0.61 |

Significance levels (two-tailed): † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 13.18: Famine mortality and post-famine ethnic composition – IV estimates.

Yet, as results in rows 3-6 of Table 13.18 indicate, famine did have an effect on ethnic composition in the long run. In 2001, there were more people who self-identified as Ukrainian speaker and Ukrainians by nationality in famine-ridden areas. To the extent that ethnic and linguistic self-identification is an individual choice that is shaped by the political context (Balcells, 2012), these results are compatible with the overall argument on the political effects of repression on political behavior. If famine did create anti-Soviet preferences that could be expressed only when the opportunity structure became more permissive in post-Soviet times, then it is conceivable that individuals living in famine-ridden areas would be more apprehensive to self-identify as Russians and Russian-speakers in a census. However, it is important to check whether these long-term effects of famine on the ethnic composition would actually explain why we observe more protests and fewer pro-Russian votes in famine-ridden areas.

To check whether ethnic composition is the mediator through which famine effects unfolded in the long run, we re-run the TSLS analyses on the three long-term outcomes by adjusting for the proportion of ethnic Ukrainians and Ukrainian-speakers in 2001. We report these results in Table 13.19: we see that the point estimates and the standard errors remain very similar to our baseline estimates.¹⁸ If indeed it were the case that ethnic composition explains the long-term effects of the famine, then the famine effect would be washed away by these control variables, but this is not what we observe.

While such post-treatment adjustments for potential mediators are commonly used to

¹⁸The results are nearly identical if we control for Russian ethnics or Russian speakers in 2001, because these variables are highly (negatively) correlated with the proportion of Ukrainian ethnics and speakers.

| Dependent variable | IV coefficient (S.E.) | First-stage F-statistic | Moran's I (p-value) |
|--|--------------------------|----------------------------|------------------------|
| <i>Adjusting for Ukrainian language in 2001</i> | | | |
| Anti-Soviet protests (1987-1991) | 0.54 (0.18)** | 9.77 | 0.99 |
| Anti-Russian vote (2002-2014) | 0.37 (0.07)** | 11.43 | 0.48 |
| Anti-Yanukovich protests (2009-2013) | 0.60 (0.19)** | 9.72 | 0.99 |
| <i>Adjusting for Ukrainian nationality in 2001</i> | | | |
| Anti-Soviet protests (1987-1991) | 0.41 (0.14)** | 10.72 | 0.99 |
| Anti-Russian vote (2002-2014) | 0.38 (0.05)** | 12.86 | 0.72 |
| Anti-Yanukovich protests (2009-2013) | 0.44 (0.17)* | 10.87 | 0.94 |

Significance levels (two-tailed): † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 13.19: IV estimates after adjusting for ethnic distribution in 2001.

evaluate causal mechanisms, they can result in biased inferences. We therefore investigate the issue further using the sequential g -estimation method, which corrects inferences for potential post-treatment biases (Acharya, Blackwell and Sen, 2016).

Using this method, we estimate the average causal direct effects of famine (instrumented by weather adversity) and compare them with our TSLS estimates. The average causal direct effects (ACDE's) estimate the effect of famine on political outcomes that cannot be explained by post-famine ethnicity and language. If the ACDE's attenuate significantly relative to the TSLS estimates, it would indicate that ethnic/linguistic composition is a plausible explanation of the famine's long-term effects.

Figure 13.13 displays TSLS and ACDE estimates with 95 percent confidence intervals for three political outcomes in modern Ukraine (columns) when the potential mediator is Ukrainian speakers in 2001 (first row) or ethnic Ukrainians in 2001 (second row). In each case, TSLS and ACDE estimates are nearly identical to each other. This suggests that even though famine did have an effect on local ethnic and linguistic composition in the long run, these ethnolinguistic changes cannot explain why famine-ridden areas held more anti-Soviet and anti-Russian protests, and why they voted less for pro-Russian parties.

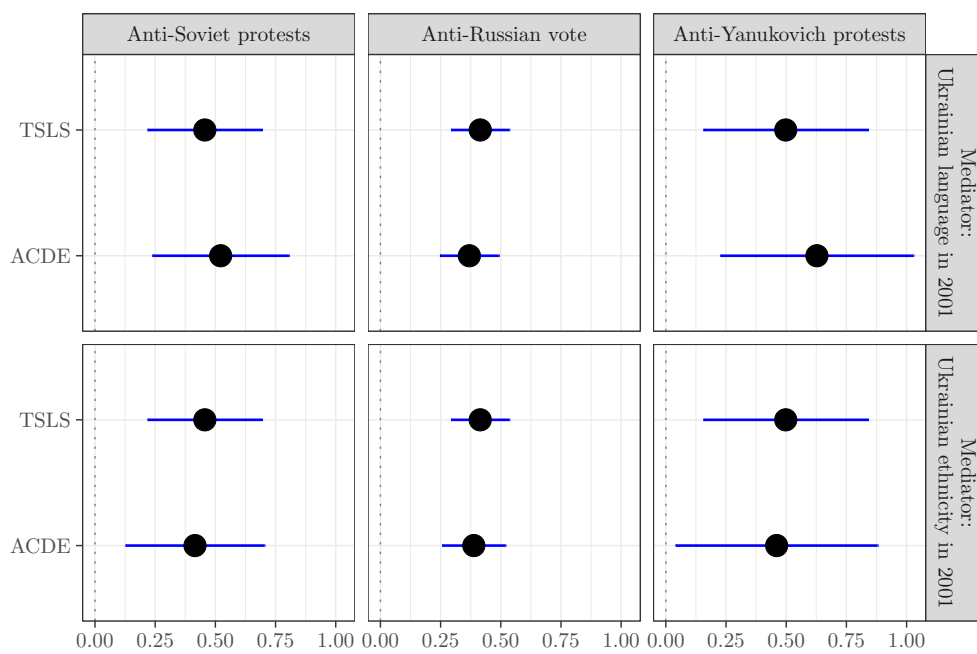


Figure 13.13: TSLS coefficients compared to average causal direct effects (ACDE) of famine, estimated using sequential g -estimator.

14. MECHANISMS OF PERSISTENCE

The current section provides additional empirical results in support of the discussion on persistence mechanisms in the main text. Table 14.20 reports regression results from five specifications. In first two columns, we regress the index of opposition to pro-Russian armed groups on rayon-level famine mortality (logged, as in our baseline models).

| | (1) | (2) | (3) | (4) | (5) |
|---------------------------------|---------------|--------------------------|---------------|--------------------------|--------------------------|
| Famine (rayon-level) | 0.54 (0.08)** | 0.30 (0.17) [†] | 0.53 (0.08)** | 0.31 (0.17) [†] | 0.31 (0.17) [†] |
| Family losses | | | 0.25 (0.07)** | 0.18 (0.07)** | 0.21 (0.12) [†] |
| Interaction | | | | | -0.02 (0.07) |
| Rayon-level covariates and FE's | | ✓ | | ✓ | ✓ |
| Observations | 1182 | 1182 | 1182 | 1182 | 1182 |

OLS coefficients with standard errors in parentheses clustered by 1933 rayons. The dependent variable is the index measure of opposition to pro-Russian separatism. Significance levels: [†] $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Table 14.20: Opposition to Russian separatism as a function of rayon-level and family-level famine losses.

The results in columns 1 and 2 indicate that respondents who reside in famine-ridden communities are more opposed to pro-Russian separatism, whether we control for pre-famine covariates or not. In columns 3 and 4, we replicate the specifications in columns 1 and 2, adding an indicator for family-level famine losses. If the only channel through which the effect of famine persists is family-level exposure, then the rayon-level famine losses should not matter after we include this variable. However, this is not at all the case: coefficients for *Famine (rayon-level)* are nearly identical across comparable specifications (1 vs 3 and 2 vs 4). In other words, we cannot explain the increased anti-separatist sentiment among respondents residing in high-famine mortality areas by their personal family victimization. In column 5, we interact family-level and rayon-level famine losses, but the coefficient is very close to zero, which indicates that losing family in Holodomor does not weaken or strengthen the effect of aggregate, rayon-level famine losses.

14.1. Bias Due to Misreporting of Family Exposure

A potential concern is that the above results are biased because some individuals may have mis-stated their family's level of victimization. If, for example, individuals who are already more "anti-Russian" falsely claim that their families suffered in the famine, this

Table 14.21: Conditional probabilities of self-reporting family exposure to Holodomor.

| Conditional probability of x^* | No misreporting | Over-reporting | Under-reporting |
|---|-----------------|----------------|-----------------|
| $p_{111} = P(x_i^* = 1 x_i = 1, c_i = 1)$ | = 1 | = 1 | < 1 |
| $p_{011} = P(x_i^* = 0 x_i = 1, c_i = 1)$ | = 0 | = 0 | > 0 |
| $p_{110} = P(x_i^* = 1 x_i = 1, c_i = 0)$ | = 1 | = 1 | < 1 |
| $p_{010} = P(x_i^* = 0 x_i = 1, c_i = 0)$ | = 0 | = 0 | > 0 |
| $p_{101} = P(x_i^* = 1 x_i = 0, c_i = 1)$ | = 0 | > 0 | = 0 |
| $p_{001} = P(x_i^* = 0 x_i = 0, c_i = 1)$ | = 1 | < 1 | = 1 |
| $p_{100} = P(x_i^* = 1 x_i = 0, c_i = 0)$ | = 0 | > 0 | = 0 |
| $p_{000} = P(x_i^* = 0 x_i = 0, c_i = 0)$ | = 1 | < 1 | = 1 |

measurement bias is likely to inflate differences in observed political attitudes between descendants and non-descendants (i.e. make descendants appear more “anti-Russian” and non-descendants less “anti-Russian”). Alternatively, if some “true descendants” under-report their family’s exposure, then family-level effects should diminish (i.e. make descendants appear less “anti-Russian” and non-descendants more “anti-Russian”).

We now consider the implications of misreporting bias for both family-level and community-level effects. We show that misreporting is more likely to bias estimates of family-level effects, but the direction of this bias depends on the communities in which respondents live: positive in high-exposure communities, negative in low-exposure communities. For community-level effects, however, misreporting bias will almost always *attenuate* rather than inflate our estimates (i.e. making famine appear to have less of an effect on anti-Russian attitudes). In other words, the results in Table 14.20 are more likely to be under-estimates rather than over-estimates of community-level effects.

Let $y_i \in \{0, 1\}$ denote respondent i ’s political attitudes, where $y_i = 1$ denotes opposition to pro-Russian separatism. Let $E[y_i | x_i, c_i]$ denote the expected attitudes of respondent i , given the “true” level of family famine exposure ($x_i = 1$ if family was exposed, 0 otherwise) and community of residence ($c_i = 1$ if high-famine, 0 otherwise).

Let us further define $x_i^* \in \{0, 1\}$ as i ’s *self-reported* level of famine exposure ($x_i^* = 1$ if i claims that their family was exposed, 0 otherwise). If there is no misreporting, $x^* = x$. Over-reporting would imply $x^* > x$, while $x^* < x$ indicates under-reporting. The probabilities of each type of self-reporting, conditional on one’s true level of family exposure and communities of residence, are in Table 14.21.

For *family-level effects*, our quantity of interest is the difference in political attitudes

among those whose families were exposed to famine:

$$\Delta_{c=1} = E[y_i|x_i = 1, c_i = 1] - E[y_i|x_i = 0, c_i = 1] \quad (16)$$

$$\Delta_{c=0} = E[y_i|x_i = 1, c_i = 0] - E[y_i|x_i = 0, c_i = 0] \quad (17)$$

where $\Delta_{c=1}$ is the difference in high-exposure communities (rayons) and $\Delta_{c=0}$ in low-exposure communities.

Focusing for now on high-exposure communities, we can use the law of total probability to expand expression (16) as follows:

$$\begin{aligned} \Delta_{c=1} &= (P(y_i|x_i^* = 1, x_i = 1, c_i = 1) p_{111} + P(y_i|x_i^* = 0, x_i = 1, c_i = 1) p_{011}) \\ &\quad - (P(y_i|x_i^* = 1, x_i = 0, c_i = 1) p_{101} + P(y_i|x_i^* = 0, x_i = 0, c_i = 1) p_{001}) \end{aligned}$$

Plugging in the values from Table 14.21, the family effect without misreporting is:

$$\Delta_{c=1}^* = P(y_i|x_i^* = 1, x_i = 1, c_i = 1) - P(y_i|x_i^* = 0, x_i = 0, c_i = 1)$$

The family effect with over-reporting is:

$$\begin{aligned} \Delta_{c=1}^{*(over)} &= P(y_i|x_i^* = 1, x_i = 1, c_i = 1) \\ &\quad - P(y_i|x_i^* = 1, x_i = 0, c_i = 1) p_{101} - P(y_i|x_i^* = 0, x_i = 0, c_i = 1) p_{001} \end{aligned}$$

The family effect with under-reporting is:

$$\begin{aligned} \Delta_{c=1}^{*(under)} &= P(y_i|x_i^* = 1, x_i = 1, c_i = 1) p_{111} + P(y_i|x_i^* = 0, x_i = 1, c_i = 1) p_{011} \\ &\quad - P(y_i|x_i^* = 0, x_i = 0, c_i = 1) \end{aligned}$$

The bias due to misreporting then becomes:

$$\begin{aligned} \epsilon_{c=1}^{over} &= \Delta_{c=1}^* - \Delta_{c=1}^{*(over)} = p_{101} (P(y_i|x_i^* = 1, x_i = 0, c_i = 1) - P(y_i|x_i^* = 0, x_i = 0, c_i = 1)) \\ \epsilon_{c=1}^{under} &= \Delta_{c=1}^* - \Delta_{c=1}^{*(under)} = p_{011} (P(y_i|x_i^* = 1, x_i = 1, c_i = 1) - P(y_i|x_i^* = 0, x_i = 1, c_i = 1)) \\ \epsilon_{c=0}^{over} &= \Delta_{c=0}^* - \Delta_{c=0}^{*(over)} = p_{100} (P(y_i|x_i^* = 1, x_i = 0, c_i = 0) - P(y_i|x_i^* = 0, x_i = 0, c_i = 0)) \\ \epsilon_{c=0}^{under} &= \Delta_{c=0}^* - \Delta_{c=0}^{*(under)} = p_{010} (P(y_i|x_i^* = 1, x_i = 1, c_i = 0) - P(y_i|x_i^* = 0, x_i = 1, c_i = 0)) \end{aligned}$$

If political attitudes are orthogonal to misreporting, then all four of these terms will reduce to zero. Yet if we assume, more realistically, that respondents who over-report family exposure are more anti-Russian on average than those who report correctly ($P(y|x^* =$

$1, x = 0, c) > P(y|x^* = 0, x = 0, c)$), then the direction of over-reporting bias is positive ($\epsilon^{over} > 0$). If we further assume that the average descendant who under-reports family exposure is less anti-Russian than one who reports correctly ($P(y|x^* = 0, x = 1, c) < P(y|x^* = 1, x = 1, c)$), then under-reporting bias will have a negative sign ($\epsilon^{under} < 0$).

The relative magnitude of each type of bias depends on how likely respondents are to misreport their family exposure in the two types of communities ($p_{101}, p_{011}, p_{100}, p_{010}$). While the true values of these probabilities are unobservable, there are good reasons to expect that social desirability and recall error are more likely to produce over-reporting in high-exposure communities ($p_{101} > p_{100}$) and under-reporting in low-exposure communities ($p_{010} > p_{011}$). For a variety of reasons (e.g. greater public awareness of famine, collective memories of victimization, a desire to “fit in”) respondents in high-exposure areas may be more likely to mis-identify themselves as descendants of famine victims. Likewise, true descendants living in high-exposure areas may be more aware of their families’ difficult past, and therefore less likely to under-report than descendants in low-exposure communities.

This analysis suggests that misreporting is likely to inflate family-level effect estimates in high-exposure communities, and attenuate them in low-exposure communities.

We now consider how misreporting might bias *community-level effect* estimates:

$$\Delta_{x=1} = E[y_i|x_i = 1, c_i = 1] - E[y_i|x_i = 1, c_i = 0] \quad (18)$$

$$\Delta_{x=0} = E[y_i|x_i = 0, c_i = 1] - E[y_i|x_i = 0, c_i = 0] \quad (19)$$

Repeating the same analytical steps as before for (18)-(19) gives us the following expressions of misreporting bias for community-level effects:

$$\begin{aligned} \epsilon_{x=1}^{over} &= \Delta_{x=1}^* - \Delta_{x=1}^{*(over)} = 0 \\ \epsilon_{x=1}^{under} &= \Delta_{x=1}^* - \Delta_{x=1}^{*(under)} = p_{011}(P(y_i|x_i^* = 1, x_i = 1, c_i = 1) - P(y_i|x_i^* = 0, x_i = 1, c_i = 1)) \\ &\quad - p_{010}(P(y_i|x_i^* = 1, x_i = 1, c_i = 0) - P(y_i|x_i^* = 0, x_i = 1, c_i = 0)) \\ \epsilon_{x=0}^{over} &= \Delta_{x=0}^* - \Delta_{x=0}^{*(over)} = p_{100}(P(y_i|x_i^* = 1, x_i = 0, c_i = 0) - P(y_i|x_i^* = 0, x_i = 0, c_i = 0)) \\ &\quad - p_{101}(P(y_i|x_i^* = 1, x_i = 0, c_i = 1) - P(y_i|x_i^* = 0, x_i = 0, c_i = 1)) \\ \epsilon_{x=0}^{under} &= \Delta_{x=0}^* - \Delta_{x=0}^{*(under)} = 0 \end{aligned}$$

Among descendants ($x = 1$), under-reporting bias is the greater concern for community-level effect estimates. If we assume that descendants who under-report are less anti-Russian than those who correctly report ($P(y|x^* = 1, x = 1, c) > P(y|x^* = 0, x = 1, c)$), and that under-reporting is more likely in low-exposure communities ($p_{010} > p_{011}$), then

the direction of this bias will be negative. For non-descendants ($x = 0$), the direction of over-reporting bias will similarly be negative, if we assume that over-reporters are more anti-Russian ($P(y|x^* = 1, x = 0, c) > P(y|x^* = 0, x = 0, c)$) and that over-reporting is more likely in high-exposure communities ($p_{010} > p_{011}$).

Taken together, misreporting may inflate estimates of family-level effects (but only in high-exposure communities), but is more likely to *attenuate* estimates of community-level effects. Even if individual survey responses are badly contaminated by imperfect recall – or even willful exaggeration of family victimization by respondents with anti-Russian views – the fact that we still find evidence of positive community-level famine effects should assuage concerns than inaccurate self-reporting drives our survey results.

REFERENCES

- Acharya, Avidit, Matthew Blackwell and Maya Sen. 2016. “Explaining causal findings without bias: Detecting and assessing direct effects.” *American Political Science Review* 110(3):512–529.
- Balcells, Laia. 2012. “The consequences of victimization on political identities: Evidence from Spain.” *Politics & Society* 40(3):311–347.
- Beber, Bernd and Alexandra Scacco. 2012. “What the numbers say: A digit-based test for election fraud.” *Political Analysis* 20(2):211–234.
- Belloni, Alexandre, Daniel Chen, Victor Chernozhukov and Christian Hansen. 2012. “Sparse models and methods for optimal instruments with an application to eminent domain.” *Econometrica* 80(6):2369–2429.
- Belloni, Alexandre, Victor Chernozhukov and Christian Hansen. 2011. “Lasso methods for gaussian instrumental variables models.”
URL: <https://arxiv.org/pdf/1012.1297.pdf>
- Berelovich, A, ed. 2005. *Sovetskaya derevnya glazami OGPU-NKVD (Russian) [Soviet village in the eyes of OGPU-NKVD]*. Vol. 3 (1930-1934) Moscow: ROSSPEN.
- Bivand, Roger, Jan Hauke and Tomasz Kossowski. 2013. “Computing the Jacobian in Gaussian spatial autoregressive models: An illustrated comparison of available methods.” *Geographical Analysis* 45(2):150–179.
URL: <http://www.jstatsoft.org/v63/i18/>

- Bonnell, Victoria E. 1999. *Iconography of power: Soviet political posters under Lenin and Stalin*. Vol. 27 Univ of California Press.
- Boriak, T. 2012. Prodovol'cha dopomoha Kremlia iak instrument Holodomoru v Ukraïni [Kremlin food aid as an instrument of the Holodomor in Ukraine]. In *Zlochyny totalitarnykh rezhymiv v Ukraïni: Naukovyi ta osvithnii pohliad [Crimes of Totalitarian Regimes in Ukraine: A Scholarly and Educational Perspective]*, ed. K.C. Berkhoff, A. Podol's'kyi and Tiahlyi M. Centre for Holocaust and Genocide Studies of the Royal Netherlands Academy of Arts and Sciences.
- Burkov, A. November 2, 1932. "Stenogramma vystupleniya direktora Souzzernotresta Ukrainy i Kryma A. Burkova na soveshchaniï pri TsK VKPB direktorov po voprosam planov khlebosdachi. 2 noyabria 1932 g.. (Russian) [Transcript of the speech by A. Burkov, Director of the Soviet Union State Trust of Ukraine and Crimea, at the meeting of the Central Committee of the CPSU (B) directors on the issues of the grain delivery plan. November 2, 1932]." <http://istmat.info/node/26760>. Online; accessed 21 July 2017.
- Carson, George Barr. 1955. *Electoral Practices in the U.S.S.R.* New York: Frederick A. Praeger.
- Central Statistical Directorate of USSR. 1939. *Vsesoyuznaya perepis' naseleniya 1939 goda*. Moscow: Central Statistical Directorate of USSR.
- Clyde, Merlise and Edward I George. 2004. "Model uncertainty." *Statistical science* pp. 81–94.
- Davies, R and S Wheatcroft. 2009. *The Industrialisation of Soviet Russia. Volume 5: The Years of Hunger: Soviet Agriculture 1931-1933*. New York: Palgrave Macmillan.
- Dolot, Miron. 1985. *Execution by Hunger: The Hidden Holocaust*. W. W. Norton & Company.
- Friedgut, Theodore H. 2014. *Political Participation in the USSR*. Princeton University Press.
- Gaure, Simen. 2018. *lfe: Linear Group Fixed Effects*.
URL: <https://CRAN.R-project.org/package=lfe>
- Getty, J Arch. 1991. "State and Society under Stalin: Constitutions and Elections in the 1930s." *Slavic Review* 50(1):18–35.
- Golodomor 1932-1933. N.d. "Golodomor 1932-1933 rokiv (Ukrainian) [Holodomor in years 1932-1933]." <https://tinyurl.com/ycx6gv8a>.

- International science conference. 1995. *Holodomor 1932-1933 rr. v Ukraini: prychny i naslidky. (Ukrainian) [Holodomor 1932-1933 in Ukraine: causes and consequences]*. Kyiv: International science conference, 9-19 September, 1993.
- Ivanchenko, G. November 28, 2009. "Khronologiya Vyznannia Holodomoru Genotsydom Ukrain's'koyi Natsii. (Ukrainian) [Chronology of Recognition of Holodomor as a Genocide of Ukrainian Nation]." <http://www.webcitation.org/6I7fG5jyD>. Online; accessed 20 July 2017.
- Karklins, Rasma. 1986. "Soviet Elections Revisited: Voter Abstention in Noncompetitive Voting." *American Political Science Review* 80(2):449–470.
- Katchanovski, Ivan. 2008. "The Orange Evolution? The "Orange Revolution" and Political Changes in Ukraine." *Post-Soviet Affairs* 24(4):351–382.
- Kenez, Peter. 1985. *The birth of the propaganda state: Soviet methods of mass mobilization, 1917-1929*. Cambridge University Press.
- Kondrashin, V. V., ed. 2012. *Golod v SSSR 1929–1934 (Russian) [Famine in USSR 1929–1934]*. Vol. 2 Moscow: MFD.
- Kovalskyi, B. 2004. *History of Ukraine: Social and Political Aspects*. Kiev: Kiev Politechnical Institute.
- Kulchytskiy, S. 2003. *Demographic Results of the Holodomor 1933 in Ukraine*. Kiev: Institute of History of Ukraine at National Academy of Sciences of Ukraine.
- Kuzio, Taras. 2005. "From Kuchma to Yushchenko Ukraine's 2004 presidential elections and the Orange revolution." *Problems of post-communism* 52(2):29–44.
- Leonov, F. October 22, 1932. "Shifrotelegramma pervogo sekretarya vostochno-sibirskogo kraykoma VKP(b) F. G. Leonova I.V. Stalinu o plane khlebozagotovok. 22 oktyabrya 1932 g.. (Russian) [Ciphered telegram of the first secretary of the East Siberian Regional Committee of the CPSU (b) F. G. Leonova I. V. Stalin on the plan of grain procurement. October 22, 1932]." <http://istmat.info/node/28865>. Online; accessed 21 July 2017.
- Maskudov, S. 1989. *The Loss of Population in the USSR*. Michigan University: Chalidze Publications.
- Natsional'nyy tsentr narodnoyi kul'tury. 2018. "Memoirs of Holodomor Eyewitnesses." <https://tinyurl.com/ybfdumea>.

- Naumenko, Natalya. 2017. "The Political Economy of Famine: the Ukrainian Famine of 1933."
- Pravda. April 27, 2010. "Yanukovitch Zdav Rosii shche i Holodomor. (Ukrainian) [Yanukovitch Gave Holodomor to Russia]." <http://www.pravda.com.ua/news/2010/04/27/4978955/>. Online; accessed 20 July 2017.
- Pravda. June 22, 2017. "Congress: SShA Pryskoryt' Vyznannya Holodomoru Genotsydom - Poroshenko. (Ukrainian) [USA Congress Will Speed Up Holodomor Recognition As Genpcode]." <http://www.pravda.com.ua/news/2017/06/22/7147593/>. Online; accessed 20 July 2017.
- Pyrig, Ruslan, ed. 2007. *Holodomor 1932–1933 rokiv v Ukraini: Dokumenty i materialy (in Ukrainian) [Holodomor 1932–1933 in Ukraine: Documents and Sources]*. Kyiv: Institute of History of Ukraine.
- Raftery, Adrian E, David Madigan and Jennifer A Hoeting. 1997. "Bayesian model averaging for linear regression models." *Journal of the American Statistical Association* 92(437):179–191.
- Raftery, Adrian, Jennifer Hoeting, Chris Volinsky, Ian Painter and Ka Yee Yeung. 2018. *BMA: Bayesian Model Averaging*. R package version 3.18.8.
URL: <https://CRAN.R-project.org/package=BMA>
- Robins, James M and Naisyin Wang. 2000. "Inference for imputation estimators." *Biometrika* 87(1):113–124.
- Rudnytskyi, Omelian, Nataliia Levchuk, Oleh Wolowyna, Pavlo Shevchuk and Alla Kovbasiuk. 2015. "Demography of a man-made human catastrophe: The case of massive famine in Ukraine 1932-1933." *Canadian Studies in Population* 42(1-2):53–80.
- Sekhon, Jasjeet S. 2011. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* 42(7):1–52.
- Sheboldaev, B. August 20, 1932. "Dokladnaya Zapiska Pervogo Sekretaria severokavkazskogo kraykoma VKP(b) B. P. Sheboldaeva I. V. Stalinu o neobkhodimosti sokrashcheniya plana khlebozagotovok, 20 avgusta 1932 g.. (Russian) [Memorandum of the First Secretary of the North Caucasus Regional Committee of the CPSU (B) B. P. Sheboldaev, to Stalin on the need to reduce the plan for grain procurement August 20, 1932]." <http://istmat.info/node/25468>.

- Simon, Noah, Jerome Friedman, Trevor Hastie and Rob Tibshirani. 2011. "Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent." *Journal of Statistical Software* 39(5):1–13.
URL: <http://www.jstatsoft.org/v39/i05/>
- Svoboda. August 12, 2009. "Programa VO Svoboda. (Ukrainian) [Program of All-Ukrainian Union Svoboda]." <http://svoboda.org.ua/party/program/>. Online; accessed 20 July 2017.
- Tauger, M. B. 2001. "Natural disaster and human actions in the Soviet famine of 1931–1933." *Carl Beck papers in Russian and East European studies* 1506:1–65.
- Tkachenko, G. November 20, 2008. "Mif o Golodomore Izobreteniyе Manipulatorov Soznaniyem. (Russian) [The Myth of Holodomor is the Invention of Mind Manipulators]." <https://kprf.ru/international/61342.html>. Online; accessed 20 July 2017.
- Vareikis, I. November 28, 1932. "Dokladnaya zapiska pervogo sekretarya Tsentral'no-Chernozomnogo obkoma VKP(b) I. M, Vareykisa I. V. Stalinu o khode khlebozagotovok, osennem seve i polozhenii v derevne. 28 noyabrya 1932 g.. (Russian) [Memorandum of the First Secretary of the Central-Chernozem Regional Committee of the CPSU (b) I.M, Vareikis I. V. Stalin on the progress of grain procurement, autumn sowing and the situation in the village. November 28, 1932]." <http://istmat.info/node/27287>. Online; accessed 21 July 2017.
- Verkhovna Rada Ukrainy. November 28, 2002. "Postanova Verkhovnoyi Rady, 789-15. (Ukrainian) [Resolution of Verkhovna Rada, 789-15]." <http://zakon3.rada.gov.ua/laws/show/789-15>. Online; accessed 20 July 2017.
- Verkhovna Rada Ukrainy. November 28, 2006a. "Proekt Zakonu pro Holodomor, 376-V. (Ukrainian) [Draft of Holodomor Law, 376-V]." http://w1.c1.rada.gov.ua/pls/zweb2/webproc4_1?pf3511=28790. Online; accessed 20 July 2017.
- Verkhovna Rada Ukrainy. November 28, 2006b. "Zasidannia 36 Verkhovnoyi Rady. (Ukrainian) [Meeting 36 of Verkhovna Rada]." http://static.rada.gov.ua/zakon/sk15/2session/STENO0G/28110602_36.htm. Online; accessed 20 July 2017.
- Viche. November, 2010. "Potashniy Y. Holodomor i Polityky. (Ukrainian) [Holodomor and Politicians]." <http://www.viche.info/journal/2272/>. Online; accessed 20 July 2017.

- Wolowyna, Oleh, Serhii Plokyh, Nataliia Levchuk, Omelian Rudnytskyi, Pavlo Shevchuk and Alla Kovbasiuk. 2016. "Regional variations of 1932–34 famine losses in Ukraine." *Canadian Studies in Population* 43(3–4):175–202.
- Wood, Simon N. 2008. "Fast stable direct fitting and smoothness selection for generalized additive models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 70(3):495–518.
- Zaslavsky, Victor and Robert J Brym. 1978. "The Functions of Elections in the USSR." *Europe-Asia Studies* 30(3):362–371.
- Zhiromskaia, V. 1990. "Vsesoiuznye perepisi naseleniia 1926, 1937, 1939 godov: Istoriia podgotovki i provedeniia [All-Union Censuses in 1926, 1937, 1939: Planning and implementation]." *Instoriia S.S.S.R.* 3:84–103.