

The Soviet Great Famine, 1931–33*

(Prepared for the 2020 Summer Workshop in Economic History and Historical Political Economy of Russia;

Preliminary: Do not cite without the Authors' Permission)

Andrei Markevich,[†] Natalya Naumenko[‡] and Nancy Qian[§]

June 20, 2020

This paper documents new facts about the Soviet Great Famine, 1931–33. First, weather was not an important contributor to the famine. Second, excess mortality was much higher for regions with a higher share of ethnic Ukrainians, even outside of the Soviet Republic of Ukraine. Third, this cannot be explained by differences in historical productivity, urbanization or historical political variables. Instead, we find that Soviet policies were implemented more zealously in these areas, which resulted in more collectivization of agriculture, lower grain productivity and higher procurement during the famine. These and other results in the paper are consistent with the presence of ethnic bias in famine-era Soviet policies.

JEL: N4, P2 **Keywords:** Communism, U.S.S.R., Planned Economies, Economic History, Political Economy, Ethnic Bias

1 Introduction

In the past one hundred years, more people have perished from famine than both World Wars combined (Sen, 1981). In just two years, 1932 and 1933, an estimated 5.5 to 10.8 million individuals died from starvation in the Soviet Great Famine.¹ In terms of total mortality, this was the second deadliest famine in the 20th century.² An inextricable fact from the Soviet Great Famine is the variation in mortality rates across ethnic groups. Ethnic minorities such as Ukrainians and Kazakhs suffered higher mortality than the Russian majority. Ethnic Ukrainians, the largest ethnic minority in the Soviet Union comprising 21% of the total 1926 population, amounted to 30% to 45% of total famine mortality.³ Thus, an analysis of the causes of Ukrainian mortality rate is central to understanding the causes of the Soviet Famine. We discuss other (smaller) ethnic groups after presenting the main results about Ukrainians.

*We express our deepest gratitude to Ekaterina Zhuravskaya, with whom we started this project. We thank Paul Catañeda Dower, Cormac Ó Gráda, Scott Gehlbach, Mikhail Golosov, Oleg Khlevniuk, Joel Mokyr and Stephen Wheatcroft for their insights; and the participants of the Northwestern Economic History lunch and HECER (Helsinki Center of Economic Research) Labor and Public Economics seminar for helpful comments. We thank Pavel Bacherikov, Nikita Karpov and Ivan Korolev for their excellent research assistance. All mistakes are our own.

[†]amarkevich@nes.ru, New Economic School (Moscow, Russia), CEPR

[‡]George Mason University, nnaumenk@gmu.edu

[§]nancy.qian@kellogg.northwestern.edu, NBER, BREAD, CEPR

¹Davies and Wheatcroft (2004) gives an estimate of 5.5 to 6.5 million deaths. Ellman (2005) estimates “about eight and a half million” victims of famine and repression” for the period of 1930 to 33. A leading Russian famine historian, Victor Kondrashin, gives a range between five to seven million victims (Kondrashin, 2008). Russian historical demographers give a range of 7.2 to 10.8 million famine victims (Polyakov and Zhiromskaya, eds, 2000). In 2008, Russian State Duma issued a statement postulating that within the territories of the Volga Region, the Central Black Earth Region, Northern Caucasus, Ural, Crimea, Western Siberia, Kazakhstan, Ukraine and Belarus, the estimated famine death toll is seven million people (State Duma, 2008). This is similar to the estimate from Conquest (1986).

²The Chinese Great Famine (1959–61) experienced higher total mortality, but lower mortality as a share of the population than the Soviet Great Famine.

³For the Ukrainian SSR, Meslé et al. (2013) estimates excess deaths to be 2.6 million, Rudnytskyi et al. (2015) — 3.9 million. It is impossible to tell how many of the famine victims were of Ukrainian ethnicity within and outside of the Ukrainian SSR. In the Ukrainian SSR, 80% of the population were ethnic Ukrainian. Assuming that within Ukraine famine victims were distributed among all ethnic groups proportionally to the group size, and that outside Ukraine none of the famine victims were ethnic Ukrainian (both assumptions this paper

The causes of this famine, and particularly of the disproportionate Ukrainian mortality rates, which is often referred to as *Holodomor*, has been a subject of intense controversy. While there is a consensus that state policy contributed to famine mortality, disagreement exists for the extent of state responsibility versus bad weather and ethnic discrimination in Soviet policies. Existing explanations can be broadly categorized into two groups. The first blames Soviet economic policies, which together with bad weather, reduced grain production in 1931 and 32, and caused over-procurement of food from the peasants. This view implies that Ukrainians suffered disproportionately because they lived in places that were unfortunately struck with weather shocks and/or in places that were more agriculturally productive (and thus strategically important in terms of procurement policy) regions for the Soviet state (e.g., Davies and Wheatcroft, 2004; Kondrashin, 2008).⁴ The second ascribes an intent of Soviet policy to target Ukrainians in order to increase political and ethnic homogenization of the newly established U.S.S.R. (e.g., Graziosi, 2004, Mace, 2004).⁵ Some go so far as to argue that the famine was a genocide committed on the Ukrainian people.⁶

The main challenge in distinguishing between the two views has been the lack of systematic disaggregated data. Existing studies rely on narrative and descriptive evidence using nation-level estimates. Importantly, there is no direct documentary evidence that Stalin “ordered” a famine.⁷ For the sake of space and given that the audience of this workshop is familiar with Russian history, we move a more detailed description of 1932 - 33 famine to the appendix (A).

The primary goal of our study is to make progress on this important question. We combine archival and published data to construct novel province-level and district level panel data on mortality, policy outcomes and natural conditions. To the best of our knowledge, we are the first to construct such comprehensive and granular panels for the entire Soviet Union. These data allow us to conduct rigorous empirical analysis and distinguish between the prevailing hypotheses by controlling for confounding factors.

We begin our analysis with an exploration of the province-level panel, which covers almost the entire inter-war period (1925—1940) and whole European part of the Soviet Union. We examine the effect of Ukrainian population share on famine mortality rates by regressing provincial mortality rates on pre-famine Ukrainian population share and its interaction with a famine period dummy variable. The baseline regression controls for province fixed effects to account for all time invariant differences across provinces (e.g., some provinces always have higher mortality rates), year fixed effects to account for changes over time that affect all provinces similarly (e.g., technological advances that affect mortality), and the urban population share and its interaction with the famine period dummy variable. The latter is to account for the urban-favoritism of Soviet food procurement policies during the famine era. We find that famine mortality rates increase with Ukrainian population share, which is consistent with the claim that ethnic Ukrainians suffered higher famine mortality than the other Soviet ethnicities (mostly ethnic Russians).

Moreover, we find that higher mortality rates in areas with a higher share of ethnic Ukrainians were not due to an unhappy coincidence of the economic and political attributes of Ukrainians and Soviet policy. The baseline result is robust to controlling for a large number of variables that proxy for factors which historians have raised as possible drivers of famine: e.g., contemporaneous weather, historical agricultural production and political variables. Interestingly, we show that there was no relationship between grain predicted by natural conditions and mortality rates. In fact, temperature and rainfall during the famine were similar to other years when there was no famine. These latter results mean that weather did not contribute to famine mortality in Ukrainian-populated regions and cannot explain timing of famine.

To understand the drivers of higher Ukrainian mortality, we examine policy outcomes as dependent variables.

refutes), and taking total famine excess deaths to be 7 million (the most cited figure), ethnic Ukrainian deaths constitute at least 30% (0.8*2.6/7) to 45% (0.8*3.9/7) of the famine victims.

⁴Also see Wheatcroft (2017) and Kotkin (2017).

⁵Also see, for example, Applebaum (2017), Ellman (2007), Rosefielde (2010) and Snyder (2010))

⁶Such interpretation goes back to Lemkin (2014 based on the speech from 1953) and Conquest (1986), which concluded “it certainly appears that a charge of genocide lies against the Soviet Union for its actions in the Ukraine”. In a 2003 letter to Wheatcroft and Davies, Conquest retracted this view. However, the debate continued. In 2006, the Ukrainian Parliament classified *Holodomor* as a genocide of Ukrainians. A number of countries, including the United States, Poland, Mexico, and Latvia supported this legal verdict.

⁷Kotkin (2017) notices that in contrast to the 1933 famine, there is plenty of direct evidence demonstrating Stalin’s intent for other killings such as during the Great Purge.

We find that at the time of the famine, Ukrainian rural population share was positively associated with the extent of agricultural collectivization, negatively associated with grain productivity and food availability (calculated as production minus government procurement).

We further delve in to the policy mechanisms behind excess Ukrainian mortality by estimating heterogeneous treatment effects of the Ukrainian population share and Soviet policies. The results show that famine mortality was higher in Ukrainian regions that were known to have been the focus of Soviet agrarian policies, had a higher Bolshevik vote share in 1917 (the last free elections before the establishment of the communist dictatorship), more rural communists (who procured food), and sent a larger number of local Party Secretaries to the 1930 Party Congress to vote on forcing collectivization (a measure of commitment to Stalin and/or collectivization of the local Party elite). In contrast, the presence of opposing political forces, such as the nationalist vote shares in 1917 reduced famine mortality in areas with more ethnic Ukrainians. These results are consistent with claims that Soviet policies and their zealous enforcement disproportionately targeted Ukrainians.

Interestingly, we find no evidence that the ethnicity of the local Party bosses matter (conditional on Ukrainian population share). This is consistent with the historical evidence of a struggle between the Bolshevik regime and Ukrainians, and a view that there were no other tensions between ethnic Russians and Ukrainians per se.

In addition to the province-level panel, we construct a novel district-level panel. The increased granularity allows us to provide several additional pieces of evidence, even though these data include fewer variables and years. First, we show that the baseline result of excess Ukrainian mortality during the famine is true at the district level, even when we add province-year fixed effects. This is consistent with the conventional wisdom that Soviet policies were centrally planned and implemented top down – e.g., if collectivization or procurement targets were partly based on Ukrainian population share, we should expect to see similar associations across large administrative units and across smaller ones within the large units. Second, we find that the distance to large cities is negatively associated with famine mortality for Ukrainians (relative to other ethnicities). This is consistent with the fact that the state bureaucracy was based out of urban areas. The district-level results support the province-level results that higher famine mortality for Ukrainians is a result of systematic bias in Soviet policies.

Finally, we supplement the main results with two additional exercises. First, we expand our analysis to separately examine famine mortality for other ethnic groups in our data. We find that famine mortality rates for ethnic Kazakhs and Belorussians were also higher than for Russians. The evidence is consistent with ethnic bias, but also suggests that the bias may be driven by different policy mechanisms than for Ukrainians. See Section 4 for a more detailed discussion. Second, for comparison purposes, we examine the famine of 1892, which occurred during the Tsarist regime. We find that the Ukrainian population share was uncorrelated with famine mortality in the Tsarist regime. Thus, the main results are unlikely to be due to ethnic-specific characteristics which causes higher mortality during famines (e.g., differences in genetics or cultural practices).

To the best of our knowledge, our study is the most systematic and rigorous empirical analysis of the causes of the Soviet Great Famine. The results support the view that high Ukrainian famine mortality was a result of systematic bias in Soviet policy, and unlikely to have been the result of unfortunately coincidence.

These findings are related to several literatures in economics in addition to the studies on the Soviet Famine discussed earlier. The first are studies of the causes of famine, which have examined contexts such as China (Becker, 1996; Thaxton, 2008), India (Sen, 1981), and Ireland (Ó Gráda, 1999).⁸ We add to these by documenting a new mechanism, ethnic bias, in a new context, the Soviet Union. Interpreted through the lens of the food availability framework laid out by Sen (1981), ethnicity is simply another dimension for inequitable distribution.

Our findings contribute to our understanding of the consequences of the economic policy of the U.S.S.R., one of the largest political-economic experiment in the 20th Century, if not all of human civilization. We complement macro calibrations of Soviet industrialization policies by Allen (2003) and Cheremukhin et al. (2017) which do not take into account human costs of Stalin's great leap forward. In examining the famine, we are most closely related to Naumenko (2020), which documents a positive association between collectivization and the famine mortality in a cross-section of districts of the Soviet Republic of Ukraine.⁹

⁸See Ó Gráda (2009) and Alfani and Ó Gráda, eds (2017) for an overview.

⁹There are numerous studies in other disciplines about the causes of the Soviet Great Famine. They rely on narrative or descriptive empirical evidence. For some well-known examples, see the studies referenced earlier in the Introduction.

More generally, the results are related to the political economy literature on ethnic conflict (Padró i Miquel, 2007; Chassang and Padró i Miquel, 2009).¹⁰ Our context is most closely related to studies of mass killings. The findings of our paper are directly in line with the theoretical predictions for mass killings and genocide from Esteban et al. (2015). The fact that these events unfolded only a decade after the USSR was established in 1922 and when Stalin was still solidifying power is consistent with the view that ethnic homogenization is complementary to nation building (Alesina and Reich, 2015).¹¹

2 Main Results

2.1 Data

The data used in the province-level panel come from a range of published and archival sources (see the appendix for the full reference list).¹² Data on our major variable of interest, population ethnic composition, come from the 1926 Soviet census, which is commonly viewed as one of the best Soviet census and a reliable source of information (Andreev et al., 1998).

At the province level, we collect data on mortality, grain production and procurement, collectivization, weather, indicators of political development and other local characteristics, and construct a panel of seventeen provinces spanning most of the European part of the Soviet Union and the years of 1925 to 1940 (A.1a).¹³ During this period, there were ongoing administrative changes in the Soviet Union that led to multiple changes in administrative borders. We convert all our variables to 1932 provincial units because the grain procurement figures are reported for these administrative borders. An average region in our panel has 6.7 million people in 1926. The Soviet Republic of Ukraine stands for one province in this panel. According to the 1926 population census, ethnic Ukrainians comprised 23.2 out of 29.2 million citizens of this region, and 5.6 ethnic Ukrainians lived in the other regions of our sample. Altogether, our panel covers 77.3% of the Soviet population, and 76.8% of 1928 Soviet grain was produced in the regions that we study.

Figure 1a plots mortality rates (the total number of deaths divided by total population) during 1900 - 1990. 1b looks at only the 1925-1940 period.¹⁴ The data show a downward trend over time, with a few spikes in 1910, 1932, 1933 and 1947. These correspond to the 1910 cholera epidemic, and the 1932-1933 and 1947 famines (Ellman, 2000; Davis, 2018). The famine that we study during 1932-1933 coincides with the largest jump in mortality. Mortality rates in the Ukraine are almost 60 per 1,000, approximately twice the mortality rate in Russia, which was approximately 30 per 1,000.

Figure A.1a maps excess mortality in 1933, the year with the highest famine mortality, for our sample.¹⁵

Figure 2 plots per capita grain production and procurement for the Soviet Union from 1916 to 1940. Note that henceforth, our data includes only the European part of the Soviet Union. Per capita grain production increases over time, with a few notable downturns. One was in 1921 and was followed by a famous Volga region famine in 1922. The other is in 1931 and 1932, which was followed by the famine we study. The procurement data show a steady increase over time.¹⁶ Consistent with historical evidence, total grain procurement declines during both

¹⁰Also, see Blattman and Miguel (2010) for an overview of the Civil War literature.

¹¹Ou and Xiong (2018) documents how the Chinese government used radio to linguistically homogenize the new Chinese state and promote the Cultural Revolution (1966-76). Cantoni et al. (2017) documents how the Chinese government uses high school textbooks to affect ideology in the late 1990s.

¹²We primarily worked in three archives: State Archive of the Russian Federation (Russian: Gosudarstvennyy Arkhiv Rossiyskoy Federatsii, GARF), Russian State Archive of the Economy (Rossiyskiy Gosudarstvennyy Arkhiv Ekonomiki, RGAE), and Russian State Archive of Socio-Political History (Rossiyskiy Gosudarstvennyy Arkhiv Sotsial'no-Politicheskoy Istorii, RGASPI). We also use a wide range of official statistical publications from the 1920s and 1930s.

¹³The only European regions not covered by our panel are Kalmykia and republics of the North Caucasus. They were small in terms of population share and grain output share. There are no reliable mortality data for the rest of the Soviet Union until the mid-1930s. In particular, note that our data do not include Kazakhstan with total estimated deaths between 1 and 2 million (Kondrashin, 2008).

¹⁴No reliable mortality data are available for the periods of 1916–1923 (World War I, Civil War, and the 1922–23 famine) and 1941–45 (World War II).

¹⁵We calculate 1933 excess mortality as 1933 mortality minus 1928 mortality. Note that we only use this calculation for the maps.

¹⁶We do not have disaggregated procurement data for 1934 and later.

1921 and 1932. However, as historians have pointed out, the reduction in procurement was insufficient for avoiding famine.

It is important to briefly discuss the reliability of the historical data. The raw data used to generate the province-level tabulations were official reports, sent upwards through the different levels of government. With few exceptions, they were meant exclusively for internal use and not known to have ever been systematically manipulated (Harrison, 2011). The government used various cross-checks procedures to improve their quality. For the purposes of our paper, an important exception is aggregate grain production. This measure was viewed as a key and public indicator for Soviet economic health – i.e., the success of the new Bolshevik regime. As such, it was controversial and there exists much debate on the accuracy of the data. Both falsifications from above and below were possible (see e.g., Wheatcroft and Davies, 1994). However, only the aggregate production numbers are believed to have been inflated. Most important for our study is that there have been no claims that that falsifications were correlated with regional ethnic composition. Nevertheless, we will be careful in our analysis later and show that our results are similar regardless of whether we control for reported grain production or predicted grain production (variation in production driven by natural conditions).

We discuss the district-level data later when they become relevant.

2.2 Baseline Estimates

Because food produced in a given year is largely used to feed the population for the next calendar year, we model mortality rates as a function of food produced in the past year. The baseline equation is

$$mortality_{it+1} = \alpha + \beta Ukrainian_i \times Famine_t + \Gamma X_{it} + \gamma_i + \delta_t + \varepsilon_{it}, \quad (1)$$

Where mortality rates for province i in year $t + 1$ is a function of: the interaction of the share of Ukrainians in the rural population as reported by the 1926 Population Census, $Ukrainian_i$, and a dummy variable that equals one if it is the year of the famine, $Famine_t$; province fixed effects, γ_i ; and year fixed effects δ_t . In the baseline, the additional controls, ΓX_{it} , include the share of the urban population and its interaction with the famine dummy variable. This accounts for the fact that the policies related to the famine (agricultural production, food distribution) were very different between urban and rural areas.¹⁷ Our baseline defines the famine dummy to take a value of one if the year is 1932 because that was when the famine became apparent in all regions and because the highest mortality rates were in 1933. We estimate robust standard errors to account for heteroskedasticity.

Table 1 column (1) presents the baseline estimate. The interaction of Ukrainian population share and the famine dummy is 0.051 and statistically significant at the 1% level. Taken literally, this means if a province was comprised of 100% ethnic Ukrainians, famine mortality rates would have been higher than a province with no Ukrainians by 51 per 1,000 individuals. To assess the magnitude of the result, the bottom of the table shows that the standard deviation in 1933 mortality rates in our sample is 0.014 and the standard deviation in Ukrainian population share is 0.228. Thus, during the famine, increasing Ukrainian population share by one standard deviation would result in a 0.827 standard deviation increase in mortality. This is a large effect.

The coefficient for urban population share show that on average, more urbanized provinces have lower mortality rates. The interaction of urban population share and the famine dummy variable is zero. This implies that more urbanized provinces did not suffer systematically different famine mortality rates from rural areas on average.

To observe the timing of differential Ukrainian mortality over time, we estimate an equation very similar to the baseline, except that we interact Ukrainian population share with dummy variables for all years instead of only 1932. Each interaction coefficient reflects the mortality difference between regions with high Ukrainian share and regions with low Ukrainian share relative to the difference in the reference year, 1925. Figure 3a shows a striking pattern. Prior to the famine, 1925-1931, there was no difference in mortality rates across regions. However, regions with a higher share of Ukrainians began to experience more mortality in 1932 (approximate 10 deaths per 1,000 individuals) and peaked in 1933 (approximately 50 deaths per 1,000 individuals). This pattern is consistent with

¹⁷The baseline uses a time-varying measure of urbanization measured at the province and year level. The results are similar if we control for urbanization reported by the 1926 Census interacted with the famine dummy. These results are available upon request.

historical evidence that there was a small famine after the harvest of 1931, which was greatly exacerbated after the harvest of 1932. Afterwards, during 1934-1940, regions with higher share of Ukrainians had slight lower mortality rates than other regions. This could be because Soviet agricultural policies were relaxed after the famine or because of positive selection for survival (e.g., if the weakest had perished during the famine, then the surviving population will have lower mortality rates than otherwise).

Given that there was a moderate rise in mortality already in 1932 and a spike in 1933, Table 1 column (2) re-estimates the baseline equation using an alternative definition of the famine dummy that equals one if it is 1931 or 1932. The coefficient is smaller in magnitude than the baseline, but still positive and statistically significant at the 1% level. Henceforth, we will use the 1932 definition of famine and focus on its effect on mortality in 1933.

The baseline measures Ukrainian population share is the share of the rural population because the famine was driven by agricultural policies targeted at the rural population. Columns (3) and (4) show that our results, in terms of standard deviations, are nearly identical if we alternatively used the urban share or total share of Ukrainians. This is not surprising since the three measures of Ukrainians population share – rural, total, urban – are highly correlated across provinces. Note that the point estimate for urban Ukrainian share in column (4) is larger because Ukrainian population share in the urban areas is much smaller than that in the rural areas or the province total. Thus, the estimated level effect on mortality will be larger. Henceforth, we will use rural Ukrainian population share as the explanatory variable.

In column (5), we control for reported per capita grain production and its interaction with famine. The Ukrainian interaction coefficient is 0.048, nearly identical to the baseline, and statistically significant at the 1% level. This means that even controlling for how much grain was produced in 1932, Ukrainians still died more during the famine. This is consistent with the view that the Soviet government procured a higher share from regions with a higher share of Ukrainians. Later, we will examine procurement directly.

In column (6), we exclude the Soviet Republic of Ukraine (Ukrainian SSR). The interaction coefficient is 0.074 and statistically significant at the 1% level. However, the standardized effect is slightly smaller, 0.717, than the full sample, 0.827. This means that the baseline results are not driven by a comparison of Ukraine SSR with the other Soviet provinces. Instead, they reflect systematic higher mortality rates for regions with a higher share of *ethnic* Ukrainians.

2.3 Weather

Next, we examine the role of natural conditions – i.e., weather – in causing high Ukrainian famine mortality. We use monthly temperature and rainfall data from Matsuura and Willmott (2014) together with province-level grain production for earlier years, 1901 to 1915, to predict weather-driven production during our sample of interest.¹⁸ We control for predicted grain and its interaction with the famine dummy variable in column (7). The estimates for Ukrainians are very robust to these additional controls. Thus, higher mortality in the regions with a larger share of ethnic Ukrainians cannot be explained by bad weather – i.e., mortality in such regions were not higher than others because these regions were hit by abnormal weather in 1932. Moreover, the uninteracted and interacted coefficients for the weather-predicted grain variables are both statistically zero, which means that weather does not explain famine mortality.¹⁹

2.4 Controlling for Political Factors

Table 2 controls for political factors that are widely believed to have contributed to the famine and their interactions with the famine year dummy. Column (1) restates the baseline for comparison. Column (2) controls for per capita grain production in 1928. 1928 was the beginning of the first Five Year Plan and these were a widespread reference point in Soviet planning (e.g. see Wheatcroft, 2001). Columns (3)-(5) examine proxies for the populations commitment/loyalty to the Bolsheviks with votes shares from the 1917 Constituency Assembly election. This was

¹⁸See the Appendix B for more discussion about the prediction.

¹⁹Appendix Figure A.4 plots demeaned average seasonal temperature and rainfall over time and show that weather during the famine was similar to other years, when there were no famines.

a universal election, the first since the toppling of the Tsar and the last one before the seventy-year Bolshevik rule. Approximately 60% of the population turned out. We construct measures for the share of votes for the Bolsheviks and for nationalist parties.^{20 21} We control for them individually in columns (3) and (4) and then together in column (5). Column (6) controls for the number of Communist Party Members (averaged over 1922, 1927 and 1931) per one thousand individuals in each province. Party members were key motivators and enforcers of state policy and we interpret this measure as a proxy for state capacity. Column (7) separately controls for urban Communists and rural Communists. These two groups are both part of the state bureaucracy, but played very different roles during the famine. We will discuss this more in the next section. Column (8) controls for the number of Party secretaries (at the province, district, city and, if the city was very large, the borough level) who attended the 1930 Party Congress to vote formally for the policy of total Collectivization . Since the Congress was manipulated such that all delegates voted in the affirmative, the number of voting delegates can be interpreted as a proxy for loyalty to Stalin or commitment to agricultural collectivization. In column (8), we control for all of the political factors together. Our main interaction coefficient for Ukrainians is always similar to the baseline in column (1).

These results meant that higher mortality in areas with higher Ukrainian population share is not due to a coincidence of the presence of political factors and the Ukrainian population.

We note that some of the policy interactions are statistically significant. Later, we will demonstrate that some of these average effects mask heterogeneity and are misleading if interpreted on their own. Thus, we defer the discussion of these other factors until the section on heterogeneous effects. Moreover, the coefficients for the additional controls in the horse race is not conceptually important for our paper, since we do not aim to distinguish the contributions of the different political proxies. Rather, the point is that the main result for Ukrainian bias is robust regardless of how we proxy for political differences.

2.5 Policy Mechanisms

Table 3 explores the mechanisms behind high famine mortality in the regions populated by ethnic Ukrainians. The historical evidence discussed in the Background Section and the descriptive evidence from Naumenko (2020) point out that the collectivization of agriculture is one of the main contributors to the famine. To investigate whether mortality was higher in regions with a higher Ukrainian share because these territories were collectivized more intensely, we examine collectivization as the dependent variable in the baseline equation. Collectivization is measured as the share of rural households in collective farms.²² Table 3 column (1) shows that the interaction coefficient is 0.135 and statistically significant at the 10% level. This means that in 1932, a higher share of land was collectivized in provinces with a higher share of Ukrainians.

Column (2) examines reported per capita grain production. Consistent with the belief that collectivization reduced agricultural productivity, we find that the interaction coefficient on production is -0.738 and significant at the 5% level. Column (3) examines per capita procurement as share of production. Collectivization dramatically increased state capacity for the procurement of grain from the countryside. We measure procurement as a share of grain production. The interaction coefficient is positive, but imprecise possibly because of a short time period for which data (1925-1933). Column (4) examines food availability, measured as per capita food retention, which we calculate as the difference between production and procurement. The coefficient is -0.441 and is statistically significant at the 1% level. This means that famine food availability was substantially lower in provinces with more ethnic Ukrainians.

Together, these results form a consistent picture. Collectivization was more intense for areas with more ethnic Ukrainians. This led to lower food production and higher procurement share, lower food retention and higher

²⁰We follow Castañeda Dower and Markevich (2020) and use disaggregated district-level data on votes for particular parties from Protasov et al. (2014).

²¹See the Data Appendix for details.

²²During the 1920s, the government advertised and promoted but did not enforce collectivization. The course for rapid large-scale collectivization was adopted in the late 1929. In early 1930, total collectivization became an official goal. So 1930 was the first year when collectivized agriculture was a non-trivial share of the rural economy. We have data for collectivization rates from 1927, when it still was virtually zero, to 1939, when most of the rural households belonged to the collectives and collectivization of agriculture was complete. See the Data Appendix for more details.

mortality. These findings are consistent with the historical narrative evidence discussed in the Background Section.

There were many other elements of collectivization which were likely to have contributed to famine mortality – e.g., the depletion of livestock, de-kulakization. They are excluded from our analysis because these variables are only available for certain years. Since they are outcomes of collectivization, they will not overturn our main finding of ethnic bias against Ukrainians.

2.6 Heterogeneous Effects

Another way of asking whether Ukrainians were differentially treated by Soviet policies is to examine the heterogeneous effects of the political variables we examined in Section 2.4 for Ukrainian population share. Table 4 estimates the fully saturated triple interaction specification of these policy variables on mortality. The double interaction captures the effect of the policy on excess famine mortality rates in a province with no Ukrainians (in rural areas). The interaction of Ukrainian population share and the famine dummy variables is not very meaningful in this table since it captures excess mortality rates for Ukrainians in provinces where the policy variables of interest has a value of zero. In these estimates, we also control for the triple interaction of urbanization, Ukrainian population and famine share to account for the possible correlation between urbanization and the policy variables.

The triple interaction captures the effect of the policy on excess famine mortality rates in a province that is 100% Ukrainian (in rural areas) relative to a province with no Ukrainians. Column (1) shows that regions which were agriculturally productive in 1928 suffered higher mortality rates, but only if there were Ukrainians in the region. This implies that the agricultural policy bias towards productive areas existed only in the regions with Ukrainians. Moreover, the negative and statistically significant coefficient on the interaction of the share of Ukrainians with famine dummy suggests that, in a hypothetical region with no grain production (and accordingly no room for collectivization policy), a higher share of Ukrainians would decrease mortality. This is consistent with the findings in the previous section.

Columns (2)-(4) conduct a similar analysis with Bolshevik and nationalist party vote shares in the 1917 election. Column (2) shows that in places with no Ukrainians, higher Bolshevik vote share is negatively correlated to famine mortality. However, the large positive triple interaction coefficient indicates that Bolshevik vote share is positively associated with mortality for regions with a high Ukrainian population share. Column (3) shows that nationalist vote shares are uncorrelated with famine mortality in regions without Ukrainians, but reduce mortality in regions with many Ukrainians. These results are consistent with the later Stalinist practice of rewarding loyal populations and penalizing disloyal ethnicities, which was openly realized in the late 1930s and mid 1940s (Polyan, 2001). When we control for all four variables in one regression in column (4), only the double and triple interactions of votes for the Bolsheviks remain statistically significant.

Column (5) examines the number of Communist Party members per capita in the years prior to the famine, which reflects the state capacity of the central government in each region. We divide communists into rural and urban communists because of they were responsible for implementing different policies. Rural communists were tasked with procurement, while urban communists were tasked with distributing the procured food to urban populations and mitigating the consequences of famine when they spilled over to urban areas.²³ Rural communists increase mortality everywhere but especially for areas with many Ukrainians. The difference between the two coefficients is statistically significant at the 1% level. Urban communists moderated famine mortality in areas populated by Ukrainians but not in the other regions. The latter may be an artifact of the fact that mortality rates were much lower in the other regions of our sample (and thus required less mitigation).

Column (6) examines the effect of the presence of Party Secretaries, who participated in the 1930 Party Convention and voted for collectivization, in a province. The triple coefficient is positive and statistically significant at the 1% level, and the double coefficient is positive but not significant. The estimates imply that political commitment of the bureaucratic leadership increased famine mortality, but only in regions with Ukrainians.

Next, we investigate the notion that Stalin exploited ethnic differences of his lieutenants to implement famine-era policies. We are able to identify the ethnicity of Party Secretaries who were sent to the 1930 Party Congress to

²³For example, many famine refugees went to cities to beg for food. Urban government also set up orphanages to abandoned children (Davies and Wheatcroft, 2004; Kondrashin, 2008).

vote for collectivization. Upon arriving to the Congress, each delegate had to fill a registration form which had a question on ethnicity, and these forms are available in the former Soviet archives (RGASPI 58/1 files 1-16). This allows us to distinguish the effects of having delegates who were themselves Ukrainian versus having delegates who were of other ethnicities. In column (8), we add the triple interactions of each of the two new controls with Ukrainian population share and the famine dummy variable. Both triple interactions are positive and statistically significant, while the double interactions are statistically zero. Moreover, the triple interactions are similar in magnitude. This implies that while the presence of voting delegates only increased famine in Ukrainian regions, the ethnicity of the delegate did not matter. Ethnically Ukrainian delegates were just as zealous as delegates of other ethnicities.

As in our earlier analysis, we can examine collectivization, grain production, procurement share and retention as dependent variables for the heterogenous treatment estimates. The estimates are presented in Table A.4 in the appendix. The precision varies, but the signs and relative magnitudes provide a consistent explanation: Soviet policies and zealous Party leaders targeted Ukrainians, leading to higher collectivization, lower grain production, higher procurement, lower food retention and, ultimately, higher mortality.

The timing of the effects of these policies are very striking. When we estimate the triple and double interactions dummy variables for all years, we find that the effect manifests during the famine. This goes against concerns that the estimates are driven by spurious correlations. See Appendix Figures A.5, A.6, A.7, A.8 and A.9.

3 District-Level Analysis

The district-level panel comprises of two years: 1928 and 1933. See the Data Appendix for more details.

There is substantial variation in famine mortality across districts, even those within the same province. Figure shows a map of excess mortality in 1933 for each district. We define excess mortality as the difference between 1933 and 1928 mortality rates.

These data allow us to examine the claim that there was a strong border effect in that the famine was notably more severe on the Ukrainian side of the border than the Russian side.²⁴ Figure A.10a plots 1933 excess mortality against the distance to the border between Russian and Ukrainian SSR. It shows that there is a jump downwards at the border into Russia. However, this jump disappears once we control for urbanization and the rural population share of Ukrainians. This can be seen in Figure A.10b, which plots the residual mortality against distance to the border. These results are consistent with our interpretation that the Soviet policies which led to the famine targeted ethnic Ukrainians rather than the Ukrainian SSR.

Table 5 column (1) first replicates the baseline specification from the province-level estimate.²⁵ The interaction between Ukrainian population share and the famine dummy is 0.048 and statistically significant at the 1% level. In terms of magnitude, a one-standard deviation increase of the local share of ethnic Ukrainians results in a 0.543 standard deviation increase in mortality. This is a sizable effect.

Column (2) includes province-year effects. The results are very robust. This is interesting given the numerous historical works which claim that the regional famine was due to differences in the political preferences of provincial leaders. While our results do not refute the potential importance of individual persons, the exhibition of similar patterns at different levels of bureaucracy is consistent with the presence of a systematic and centrally planned policy.

Next, we exploit the granularity of the data to investigate the relationship between distance from the railroad and cities. Railways were the major method of transportation across long distances. Railways reduced the cost of procuring grain. They also reduced the cost of going to urban areas to escape the famine. Similarly, distance from cities had opposing effects in principle. Soviet policy aimed to secure the grain supply for industrial populations. This meant that cities absorbed food from nearby rural regions. It also meant that cities could provide safe harbors from starvation. We will estimate the net of the opposing forces for each variable.

²⁴The government introduced a ban on migration from the Republic of Ukraine and the North Caucasus region in January 1933 (Danilov et al., eds, 1999-2006 Vol. 3) that could cause a discontinuity effect at the Ukrainian border.

²⁵Note that we used urbanization from 1926 and 1933 because urbanization is not available from 1928.

We gradually introduce the triple interactions of distance to the nearest railway and distance to the nearest large city (e.g., population of 20,000 or more) with Ukrainian population share and the famine dummy variable in columns (3) and (5) and (4) and (6). In columns (7) and (8), we control for both in the same specification. Note that we always estimate the fully saturated model with all of the lower order interaction terms, as well as the triple interaction of urbanization, Ukrainian share and the famine dummy.

For brevity, we focus on the final column. It shows that distance to the railroad has no effect on famine mortality anywhere. However, distance from a large city reduces famine mortality. This is true everywhere, but the reduction is almost an order of a magnitude larger in regions with a large Ukrainian population share. These results suggest that the advantage of facing less procurement outweighs the disadvantage of needing to travel further to escape a famine. That the estimate is so much larger for districts with many Ukrainians is consistent with the fact that Ukrainian districts were the ones targeted by procurement agents.

In addition to the results discussed in this section, we are able to replicate all of the province-level estimates with the district-level panel (to the extent that we have data at the district level). They are presented in the Appendix Table A.5.

4 Additional Results

4.1 Other Ethnic Minorities

In this section, we return to using the province-level panel. Appendix Table A.1 lists the ethnicities most often mentioned in the famine literature. Other than Ukrainians (21.2% of the Soviet Population in 1926), Belorussians (3.2%), Kazakhs (2.7%), Germans (0.8%) and Poles (0.5%) are often thought to have higher mortality rates than Russians. In contrast, ethnic Jews (1.8%), are often thought to have suffered lower mortality rates.

Table 6 column (1) shows the baseline equation for with all ethnic groups as well as a residual group of all other ethnicities (each interacted with the famine year dummy) in the same regression. When we do this, we see that regions with higher shares of Ukrainians, Belorussians and Kazakhs suffered higher famine mortality than areas populated mostly by ethnic Russians. The magnitude is the largest for the interaction with ethnic Kazakh share. The areas with other ethnic groups suffered similar mortality as territories with ethnic Russians.

Column (2) controls for weather-induced grain. Column (3) controls for political variables interacted with famine. Column (4) include the same political variables except that it separates urban and rural communists. The estimates for Ukrainians, Belorussians and Kazakhs change little with these controls, which means that the excess mortality rates for these minorities cannot be explained by bad luck – living in areas hit by bad weather – or differences in the political factors prior to the famine. In other words, famine mortality is biased towards these ethnic groups. We note that the coefficient for Poles become large and statistically significant as we add controls. This is hard to interpret because there are very few poles in the population and they almost entirely reside in two provinces.

In the Appendix, we examine mortality rates over time for each group, as well as the mechanisms driving mortality rates in regions with a high share of Kazakh and Belorussian. The estimates suggest that the policy mechanisms may be quite different for these groups than for Ukrainians. One significant limitation for studying the Kazakh famine experience is that we do not have data on Kazakhstan, where 94% of Kazakhs resided at the time.²⁶ It is beyond the scope of this paper to delve deeper into the Belorussian famine experience. This is an important topic for future study. See Appendix Section C.

4.2 The 1892 Famine

To understand whether ethnic bias in Ukrainian famine mortality is an outcome of Soviet policy, or whether some populations always suffer higher mortality during famines (e.g., differences in social capital, networks or culture), we examine ethnic-specific mortality rates during the Tsarist famine of 1892.²⁷ Table 7 replicates the baseline as

²⁶There are no mortality data for Kazakhstan until 1935.

²⁷Volha Charnysh kindly shared 1885–1896 mortality data with us, Charnysh and McElroy (2020).

closely as possible with the earlier data. For comparison purposes, we present results in pairs for the Tsarist famine (columns 1 to 4) and the Soviet famine (columns 5 to 8). Note that the sample for the former famine is larger because Russian Empire provinces (gubernias) were smaller than the subsequently created Soviet administrative units.

We examine all of the ethnic groups that we have data for (Ukrainians, Belorussians, Kazakhs, Germans, Poles, Jews and all the rest except Russians) by including a dummy variable for each ethnic group population share interacted with the famine dummy. All specifications control for urbanization and its interaction with the famine dummy variable. To account for access to food during and outside famine periods, we add grain output per capita and its interaction with the famine dummy in columns (2) and (5). Alternatively, we control for grain yield (grain harvest per hectare per million people) and its interaction with the famine year in columns (3), (4), (7) and (8). These additional controls allow us to account for potential variation in labor inputs per unit of land since the degree of physical exertion can influence mortality. In columns (4) and (8), we also include distance to the national capital interacted with the famine period. The central government decided on the amount of relief for starving areas, and this amount could depend on available information about the scale of the disaster that was likely correlated with distance.

The main Ukrainian population share interaction is statistically zero in all specifications for the Tsarist famine. In contrast, they are positive, large in magnitude and statistically significant at the 1% for the Soviet Famine. This implies that our main results is unlikely to be driven by time-invariant characteristics of ethnic Ukrainians. Instead, the results are consistent with the belief that higher mortality in Ukrainian regions during the Soviet famine was a result of Soviet-specific policies, and the earlier evidence that Ukrainians experienced policies which led to famine, such as collectivization, more intensely.

5 Conclusion

The Soviet Great Famine was one of the largest and most controversial economic disasters in recent history. Within two years, between 5.5 and 10.8 millions had died throughout the nation and the ethnic Ukrainian population, the second largest ethnic group, had been decimated. Without systematic data to rule out confounding factors or direct documentary evidence on the intents of the government, this tragedy has unsurprisingly become a subject of heated debate. Our study contributes to this debate by constructing the largest and most comprehensive dataset on mortality, economic policy and natural conditions. The data allow us to conduct a rigorous empirical analysis that accounts for multiple factors. The results unambiguously show that Ukrainian mortality during the famine was due to systematic *ethnic* bias in Soviet policy. It was not a coincidence of where Ukrainians happened to live and abnormal weather conditions (or other exogenous factors).

These sorrowful findings suggest several avenues of future research. First, more research is needed to reveal motives of ethnic bias in Soviet policy. Second, we need to better understand the causes of the excessively high famine mortality rates of other ethnic minorities, such as the Belorussians and Kazakhs. Our precursory examination suggests that they also suffered from ethnically biased Soviet policies, but the relevant policies and how they transpired into famine are likely to be different than for the Ukrainians. Finally, the results naturally raise the question about the political and economic consequences of the famine for Ukraine and the Soviet Union or post-Soviet Eastern Europe, more generally. These are only a few possibilities within a vast number of interesting and important inquiries. The recent availability of Soviet archival data will undoubtedly lead to a great improvement in the breadth and depth of fascinating context.

References

Alesina, Alberto and Bryony Reich, “Nation Building,” 2015. NBER Working Paper 18839.

Alfani, Guido and Cormac Ó Gráda, eds, *Famine in European History*, Cambridge, UK: Cambridge University Press, 2017.

- Allen, Robert**, *Farm to Factory: A Reinterpretation of the Soviet Industrial Revolution*, Princeton: Princeton University Press, 2003.
- Andreev, E.M., L.E. Darskij, and T.L.Kharkova**, *Demographicheskaya Istoriya Rossii: 1927-1959*, Moscow: Informatika, 1998.
- Applebaum, Anne**, *Red Famine: Stalin's War on Ukraine*, New York: Penguin Press, 2017.
- Becker, Jasper**, *Hungry Ghosts: Mao's Secret Famine*, New York: The Free Press, 1996.
- Blattman, Christopher and Edward Miguel**, "Civil War," *Journal of Economic Literature*, March 2010, 48 (1), 3–57.
- Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang**, "Curriculum and Ideology," *Journal of Political Economy*, 2017, 125 (2), 338–392.
- Charnysh, Volha and Brendan McElroy**, "Explaining outgroup bias in weak states: religion and legibility in the 1891–92 Russian famine," 2020. project in progress.
- Chassang, Sylvain and Gerard Padró i Miquel**, "Economic Shocks and Civil War," *Quarterly Journal of Political Science*, 2009, 4 (3), 211–228.
- Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski**, "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model," *The Review of Economic Studies*, 2017, 84 (2), 613–649.
- Conquest, Robert**, *The Harvest of Sorrow: Soviet Collectivization and the Terror–Famine*, Oxford, UK: Oxford University Press, 1986.
- Danilov, Victor, Robert Manning, and Lynne Viola, eds**, *Tragediya Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenty i materialy v 5 tomakh, 1927-1939*, Rosspen, 1999-2006.
- Davies, Robert W. and Stephen G. Wheatcroft**, *The Years of Hunger: Soviet Agriculture, 1931–1933*, Vol. 5 of *The Industrialisation of Soviet Russia*, Basingstoke, UK: Palgrave Macmillan, 2004.
- Davis, John P.**, *Russia in the Time of Cholera: Disease under Romanovs and Soviets*, I.B. Tauris, 2018.
- Ellman, Michael**, "The 1947 Soviet famine and the entitlement approach to famines," *Cambridge Journal of Economics*, 09 2000, 24 (5), 603–630.
- , "The Role of Leadership Perceptions and of Intent in the Soviet Famine of 1931–1934," *Europe-Asia Studies*, 09 2005, 57 (6), 823–841.
- , "Stalin and the Soviet famine of 1932–33 Revisited," *Europe-Asia Studies*, 2007, 59 (4), 663–693.
- Esteban, Joan, Massimo Morelli, and Dominic Rohner**, "Strategic Mass Killings," *Journal of Political Economy*, 2015, 123(5), 1087–1132.
- Graziosi, Andrea**, "The Soviet 1931–1933 Famines and the Ukrainian Holodomor: Is a New Interpretation Possible, and What Would Its Consequences Be?," *Harvard Ukrainian Studies*, 2004, 27 (1/4), 97–115.
- Gregory, Paul**, *Terror by Quota State Security from Lenin to Stalin (an Archival Study)*, Yale University Press, 2009.
- Harrison, Mark**, "Forging Success: Soviet Managers and Accounting Fraud, 1943 to 1962," *Journal of Comparative Economics*, 2011, 39(1), 43–64.

- i Miquel, Gerard Padró**, “The Control of Politicians in Divided Societies: The Politics of Fear,” *The Review of Economic Studies*, 10 2007, 74 (4), 1259–1274.
- Kondrashin, Viktor**, *Golod 1932-1933 godov: Tragediya Sovetskoy derevni*, Moscow: Rosspen, 2008.
- Kotkin, Stephen**, *Stalin. Waiting for Hitler, 1929–1941*, New York: Penguin Press, 2017.
- Lemkin, Raphael**, *Soviet genocide in the Ukraine*, Kingston, Ontario: Kashtan Press, 2014 based on the speech from 1953.
- Mace, James E.**, “Soviet Man-Made Famine in Ukraine,” in William S. Parsons Totten Samuel and Israel W. Charny, eds., *Century of Genocide: Critical Essays and Eyewitness Accounts.*, Rouledge, 2004.
- Matsuura, Kenji and Cort Willmott**, “Terrestrial Air Temperature and Precipitation: 1900–2014 Gridded Monthly Time Series (Version 4.01),” *National Oceanic and Atmospheric Administration*, 2014.
- Meslé, France, Jacques Vallin, and Evgeny Andreev**, “Demographic Consequences of the Great Famine: Then and Now,” in Andrea Graziosi, Lubomyr A. Hajda, and Halyna Hryn, eds., *After the Holodomor: The Enduring Impact of the Great Famine on Ukraine*, Harvard University Press, 2013.
- neda Dower, Paul Casta and Andrei Markevich**, “Democratic support for the Bolshevik Revolution: An empirical investigation of 1917 Constituent Assembly elections,” *Working paper*, 2020.
- Ó Gráda, Cormac**, *Black '47 and Beyond: The Great Irish Famine in History, Economy, and Memory*, Princeton, NJ: Princeton University Press, 1999.
- , *Famine: A Short History*, Princeton, NJ: Princeton University Press, 2009.
- Ou, Susan and Heyu Xiong**, “Mass Persuasion and the Ideological Origins of the Chinese Cultural Revolution,” 2018. Mimeo.
- Polyakov, Yu.A. and V.B. Zhiromskaya, eds**, *Naselenie Rossii v XX veke: istoricheskie ochery. Vol. 1 1900-1939*, Rosspen, 2000.
- Polyan, Pavel**, *Ne po svoej vole... Istoriya i geografiya prinuditelnykh migratsij v SSSR*, Moscow: O.G.I, 2001.
- Protasov, L.G., A.V. Repnikov V.V. Zhuravlev, and V.V. Shelokhaev**, *Vserossijskoe uchreditelnoe sobranie. Entsiklopedia. (All-Russian Constitutional Assembly. Encyclopedia)*, Moscow: Rosspen, 2014.
- Rosefelde, Steven**, *Red Holocaust*, London and New York: Routledge, 2010.
- Rudnytskyi, Omelian, Nataliia Levchuk, Oleh Wolowyna, Pavlo Shevchuk, and Alla Kovbasiuk**, “Demography of a man-made human catastrophe: The case of massive famine in Ukraine 1932–1933,” *Canadian Studies in Population*, 2015, 42 (1–2), 53–80.
- Sen, Amartya**, *Poverty and Famines. An Essay on Entitlement and Deprivation*, Oxford, UK: Oxford University Press, 1981.
- Snyder, Timothy**, *Bloodlands: Europe Between Hitler and Stalin*, New York City, NY: Basic Books, 2010.
- State Duma**, “Russian: Postanovleniye GD FS RF ot 02.04.2008 N 262-5 GD O Zayavlenii Gosudarstvennoy Dumy Federal'nogo Sobraniya Rossiyskoy Federatsii “Pamyati zhertv goloda 30-kh godov na territorii SSSR”,” 2008.
- Thaxton, Ralph A.**, *Catastrophe and Contention in Rural China: Mao's Great Leap Forward Famine and the Origins of Righteous Resistance in Da Fo Village*, Cambridge, UK: Cambridge University Press, 2008.

Viola, Lynne, *Peasant Rebels Under Stalin: Collectivization and the Culture of Peasant Resistance*, Oxford, UK: Oxford University Press, 1996.

—, *The Unknown Gulag: The Lost World of Stalin's Special Settlements*, Oxford: Oxford University Press, 2007.

Wheatcroft, S.G. and R.W. Davies, "Agriculture," in R.W. Davies, Mark Harrison, and S.G. Wheatcroft, eds., *The Economic Transformation of the Soviet Union, 1913-1945*, Cambridge University Press, 1994.

Wheatcroft, Stephen, "Eastern Europe (Russia and the USSR)," in Guido Alfani and Cormac Ó Gráda, eds., *Famine in European History*, Cambridge University Press, 2017, pp. 212–239.

Wheatcroft, Stephen G., "O zernovikh balansakh i otsenkakh urozhajnosti v SSSR v 1931-1933," in Victor Danilov, Robert Manning, and Lynne Viola, eds., *Tragediya Sovetskoy Derevni. Kollektivizatsiya i raskulachivanie. Dokumenty i materialy v 5 tomakh, 1927-1939. Vol. 3. Konetz 1930 - 1933*, Rosspen, 2001.

Table 1: Ethnic Ukrainians and Famine Mortality

	Dependent Variable: Mortality in year t+1						
	Baseline (1)	Famine = 1931, 32 (2)	Total Ukrainians (3)	Urban Ukrainians (4)	Reported Grain (5)	Omit Ukraine SSR (6)	Predicted Grain (7)
Ukrainians x Famine	0.051*** [0.007]	0.031*** [0.010]	0.055*** [0.008]	0.092*** [0.013]	0.048*** [0.005]	0.074*** [0.007]	0.051*** [0.009]
<i>Normalized</i>	0.827	0.503	0.825	0.804	0.782	0.717	0.835
Urbanization	-0.009* [0.005]	-0.009* [0.005]	-0.009* [0.005]	-0.009* [0.005]	-0.007 [0.005]	-0.008 [0.005]	-0.009* [0.005]
Urbanization x Famine	0.003 [0.009]	0.010 [0.007]	0.004 [0.009]	0.002 [0.010]	0.011 [0.016]	0.001 [0.009]	0.003 [0.010]
Grain, x Famine					Y		Y
Observations	255	255	255	255	255	240	255
R-squared	0.799	0.744	0.798	0.789	0.822	0.781	0.800
Mortality 1933							
Mean	0.031	0.031	0.031	0.031	0.031	0.029	0.031
Std. Dev.	0.014	0.014	0.014	0.014	0.014	0.012	0.014
Ukrainians							
Mean	0.107	0.107	0.098	0.059	0.107	0.060	0.107
Std. Dev.	0.228	0.228	0.208	0.122	0.228	0.120	0.228

Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Ukrainians is the share of ethnic Ukrainians in the rural population (Columns 1-2, 5-7), total population (Column 2), or urban population (Column 3) according to the 1926 Population Census. Famine is an indicator that equals one in 1932 (Columns 1, 3-7) or in 1931 and 1932 (Column 2), and zero otherwise. Estimates in Column 5 control for reported grain and reported grain x Famine. Estimates in Column 7 control for predicted grain and predicted grain x Famine. Reported and predicted grain are measured in kilograms per person per day. All estimates control for province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Normalized coefficient expresses the effect of increasing Ukrainian population share by one standard deviation on 1933 mortality in terms of standard deviations. Appendix C shows the exact source of every variable.

Table 2: Ethnic Ukrainians and Famine Mortality, Controlling for Political Factors

	Dependent variable: Mortality in year t+1							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	0.051*** [0.007]	0.049*** [0.005]	0.048*** [0.007]	0.051*** [0.008]	0.052*** [0.007]	0.047*** [0.005]	0.049*** [0.006]	0.057*** [0.005]
Grain 1928 x Famine		0.009** [0.005]						0.011*** [0.004]
Bolshevik votes 1917 x Famine			-0.013 [0.009]		-0.019* [0.011]			-0.009 [0.009]
Nationalist votes 1917 x Famine				-0.001 [0.007]	-0.009 [0.009]			-0.020** [0.009]
Urban communists x Famine						-0.001 [0.001]		0.000 [0.001]
Rural communists x Famine						0.003* [0.002]		0.001 [0.002]
Voting delegates 1930 x Famine							0.048* [0.027]	0.022 [0.015]
Observations	255	255	255	255	255	255	255	255
R-squared	0.799	0.828	0.804	0.799	0.807	0.823	0.815	0.851

Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Population Census. Famine is an indicator that equals one in 1932 and zero otherwise. Grain 1928 is 1928 grain harvest measured in kilograms per person per day. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. Nationalist votes 1917 is the share of votes for nationalistic parties in the 1917 elections. Urban (rural) communists is the average over 1922, 1927, and 1931 of the number of urban (rural) Communist Party members and candidates per 1,000 people. Voting delegates 1930 is the number of province-, county-, city-, and borough-level Party secretaries that participated and had a right to vote in the 1930 Party Congress per 100,000 people. All estimates control for urbanization, urbanization x famine, province, and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table 3: Ethnic Ukrainians and Collectivization, Grain Production, Procurement, and Retention

	Dependent variable:			
	Collectivization	Grain	Procurement Share	Retention
	(1)	(2)	(3)	(4)
Ukrainians x Famine	0.135* [0.077]	-0.738** [0.302]	0.007 [0.032]	-0.431*** [0.150]
Observations	221	272	149	149
R-squared	0.966	0.795	0.904	0.790

Notes: Observations are at the province and year level. Collectivization is the share of rural households in collective farms. Grain is harvest measured in kilograms per person per day. Procurement share is the share of harvest procured by the government. Retention is grain minus procurement measured in kilograms per person per day. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Population Census. Famine is an indicator that equals one in 1932 and zero otherwise. All estimates control for urbanization, urbanization x famine, province, and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table 4: Heterogeneous Effects of Politics on Mortality

	Dependent variable: Mortality in year t+1						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ukrainians x Famine	-0.119***	0.010	0.296*	0.022	-0.244	0.083***	0.025
	[0.028]	[0.032]	[0.162]	[0.137]	[0.189]	[0.018]	[0.046]
Grain 1928 x Famine	0.000						
	[0.003]						
Ukrainians x Grain 1928 x Famine	0.270***						
	[0.021]						
Bolshevik votes 1917 x Famine		-0.039***		-0.042***			
		[0.014]		[0.014]			
Ukrainians x Bolshevik votes 1917 x Famine		1.429***		1.325*			
		[0.357]		[0.738]			
Nationalist votes 1917 x Famine			0.004	-0.006			
			[0.007]	[0.005]			
Ukrainians x Nationalist votes 1917 x Famine			-0.165**	-0.007			
			[0.075]	[0.078]			
Urban communists x Famine					0.001		
					[0.000]		
Ukrainians x Urban communists x Famine					-0.242***		
					[0.075]		
Rural communists x Famine					0.003***		
					[0.001]		
Ukrainians x Rural communists x Famine					0.093***		
					[0.009]		
Voting delegates 1930 x Famine						0.015	
						[0.011]	
Ukrainians x Voting delegates 1930 x Famine						1.872***	
						[0.086]	
Voting Ukrainian delegates 1930 x Famine							-0.280
							[0.200]
Ukrainians x Voting Ukrainian delegates 1930 x Famine							2.931***
							[0.787]
Voting non-Ukrainian delegates 1930 x Famine							0.051
							[0.037]
Ukrainians x Voting non-Ukrainian delegates 1930 x Famine							2.176***
							[0.198]
Observations	255	255	255	255	255	255	255
R-squared	0.881	0.867	0.844	0.868	0.884	0.883	0.885

Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Population Census. Famine is an indicator that equals one in 1932 and zero otherwise. Grain 1928 is 1928 grain harvest measured in kilograms per person per day. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. Nationalist votes 1917 is the share of votes for nationalistic parties in the 1917 elections. Urban (rural) communists is the average over 1922, 1927, and 1931 of the number of urban (rural) Communist Party members and candidates per 1,000 people. Voting delegates 1930 is the number of province-, county-, city-, and borough-level Party secretaries that participated and had a right to vote in the 1930 Party Congress per 100,000 people. Voting Ukrainian delegates 1930 is the number of Voting delegates of Ukrainian ethnicity per 100,000 people. Voting non-Ukrainian delegates 1930 is the number of Voting delegates of non-Ukrainian ethnicity per 100,000 people. All estimates control for predicted grain, predicted grain x famine, Ukrainians x predicted grain x famine, urbanization, urbanization x famine, Ukrainians x urbanization x famine, province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table 5: Ethnic Ukrainians and Famine Mortality, District-level Estimates

	Dependent variable: Mortality				
	(1)	(2)	(3)	(4)	(5)
Ukrainians x Famine	0.048*** [0.002]	0.042*** [0.005]	0.048*** [0.006]	0.083*** [0.010]	0.079*** [0.012]
Log distance to a railroad + 1 x Famine			0.002 [0.003]		0.003 [0.003]
Ukrainians x Log distance to a railroad + 1 x Famine			-0.001 [0.017]		0.001 [0.017]
Log distance to a 20K city x Famine				-0.001 [0.001]	-0.001* [0.001]
Ukrainians x Log distance to a 20K city x Famine				-0.009*** [0.002]	-0.008*** [0.003]
District FE	Y	Y	Y	Y	Y
Year FE	Y				
Province-Year FE		Y	Y	Y	Y
Observations	3155	3155	2901	3155	2901
R-squared	0.732	0.764	0.784	0.777	0.788
Provinces	16	16	16	16	16

Notes: Observations are at the district and year level. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Population Census. Famine is an indicator that equals one in 1933 and zero in 1928. Log distance to a railroad + 1 is log of the distance from the district's administrative center to a nearest railroad station. Log distance to a 20K city is log of the distance from the district's centroid to a nearest city with at least 20,000 inhabitants according to the 1926 Population Census. All estimates control for urbanization and urbanization x famine. Estimates in Columns 3--5 also control for Ukrainians x urbanization x famine. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table 6: Other Ethnic Groups and Famine Mortality

	Dependent variable: Mortality in year t+1		
	(1)	(2)	(3)
Ukrainians x Famine	0.071*** [0.005]	0.088*** [0.007]	0.088*** [0.008]
Belorussians x Famine	0.024*** [0.006]	0.038*** [0.007]	0.050*** [0.009]
Jews x Famine	-0.757 [1.151]	-0.091 [1.118]	1.166 [1.444]
Germans x Famine	0.030 [0.102]	-0.003 [0.096]	-0.238* [0.128]
Poles x Famine	-0.700 [1.139]	-1.913* [1.139]	-3.518** [1.657]
Kazakhs x Famine	2.267*** [0.831]	2.445*** [0.784]	4.586*** [1.039]
Other x Famine	0.000 [0.004]	0.005 [0.003]	0.009 [0.012]
Controls:			
Predicted Grain, x Famine		Y	
Political Variables x Famine			Y
Observations	255	255	255
R-squared	0.888	0.891	0.893

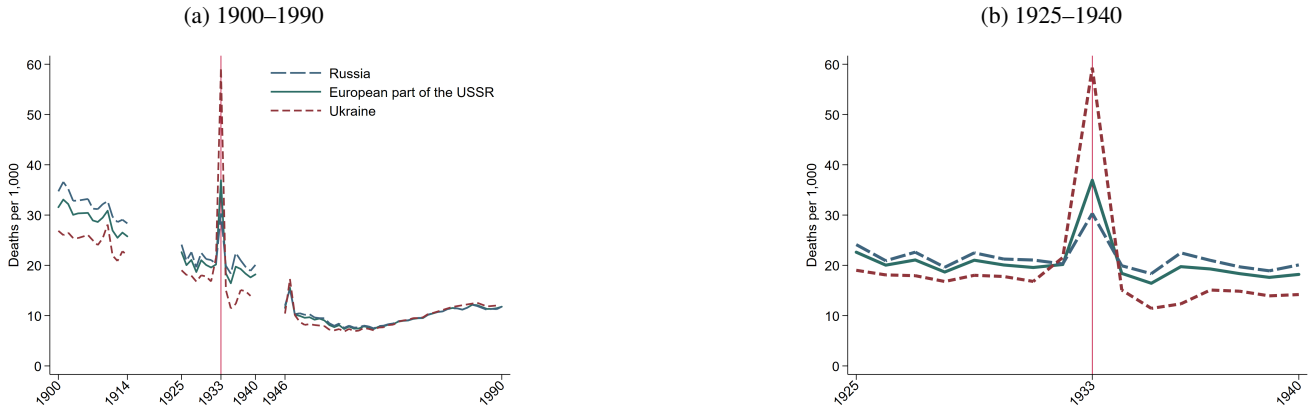
Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Ukrainians, Belorussians, Jews, Germans, Poles, and Kazakhs are the share of each ethnic group in the rural population according to the 1926 Population Census. Other is the share of non-Ukrainians, non-Belorussians, non-Jews, non-Germans, non-Poles, non-Kazakhs, and non-Russians in the rural population according to the 1926 Population Census — omitted category is Russians. Famine is an indicator that equals one in 1932 and zero otherwise. Estimates in Column (2) control for predicted grain and predicted grain x famine. Estimates in Column (3) control for Political variables x famine: Grain 1928 x Famine, Bolshevik votes 1917 x Famine, Nationalist votes 1917 x Famine, Urban Communists x Famine, Rural Communists x Famine, Voting delegates 1930 x Famine. Predicted grain and Grain 1928 are measured in kilograms per person per day. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. Nationalist votes 1917 is the share of votes for nationalistic parties in the 1917 elections. Urban (rural) communists is the average over 1922, 1927, and 1931 of the number of urban (rural) Communist Party members and candidates per 1,000 people. Voting delegates 1930 is the number of province-, county-, city-, and borough-level Party secretaries that participated and had a right to vote in the 1930 Party Congress per 100,000 people. Appendix C shows the exact source of every variable.

Table 7: Ethnic Composition and the 1892 Famine Mortality

	Dependent variable: Mortality in year t+1							
	1892				1933			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	0.002	0.002	-0.001	-0.003	0.071***	0.069***	0.072***	0.068***
	[0.003]	[0.003]	[0.003]	[0.003]	[0.005]	[0.005]	[0.006]	[0.007]
Belorussians x Famine	0.001	0.002	-0.001	-0.001	0.024***	0.021***	0.024***	0.020***
	[0.004]	[0.005]	[0.004]	[0.003]	[0.006]	[0.006]	[0.007]	[0.007]
Jews x Famine	-0.081**	-0.097*	-0.069*	-0.081**	-0.757	-1.232	-0.812	-0.907
	[0.040]	[0.058]	[0.036]	[0.037]	[1.151]	[1.142]	[1.210]	[1.103]
Germans x Famine	0.104*	0.104*	0.075*	0.051	0.030	0.052	0.036	0.044
	[0.053]	[0.059]	[0.045]	[0.048]	[0.102]	[0.098]	[0.110]	[0.097]
Poles x Famine	-0.005	0.014	-0.011	0.021	-0.700	-0.148	-0.680	-0.472
	[0.039]	[0.066]	[0.033]	[0.035]	[1.139]	[1.166]	[1.207]	[1.137]
Kazakhs x Famine	0.040***	0.046***	0.032***	0.020***	2.267***	1.933**	2.224**	2.051**
	[0.005]	[0.014]	[0.005]	[0.008]	[0.831]	[0.819]	[0.894]	[0.834]
Other x Famine	0.000	0.000	0.003	0.002	0.000	-0.001	0.000	-0.004
	[0.004]	[0.004]	[0.004]	[0.003]	[0.004]	[0.004]	[0.006]	[0.007]
Grain, x Famine		Y				Y		
Yield pc, x Famine			Y	Y			Y	Y
Log distance to the capital x Famine				0.005***				0.002
				[0.002]				[0.001]
Observations	1297	1297	1297	1297	255	255	223	223
R-squared	0.87	0.87	0.875	0.877	0.888	0.891	0.89	0.891
Provinces	50	50	50	50	17	17	17	17

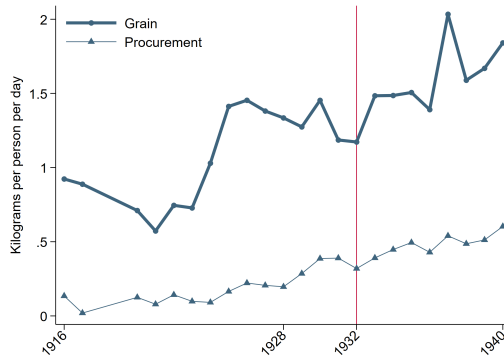
Notes: Observations are at the province and year level (In Columns 1–4, 50 European provinces of the Russian Empire 1885–1913; in Columns 5–8, 17 European provinces of the Soviet Union 1925–1940). Mortality is the number of deaths divided by population. Ukrainians, Belorussians, Jews, Germans, Poles, and Kazakhs are the share of each ethnic group in the rural population according to the 1897 Census (Columns 1–4) or 1926 Census (Columns 5–8). Other is the share of non-Ukrainians, non-Belorussians, non-Jews, non-Germans, non-Poles, non-Kazakhs, and non-Russians in the rural population according to 1897 (1926) Census — the omitted category is Russians. Famine is an indicator that equals one in 1891 (Columns 1–4), in 1932 (Columns 5–8), and zero otherwise. Grain is grain harvest measured in kilograms per person per day. Yield is grain yield measured in centners per hectare per million people. All estimates control for Urbanization, Urbanization x Famine, province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Figure 1: Mortality in the European Part of the Soviet Union, Russia, and Ukraine



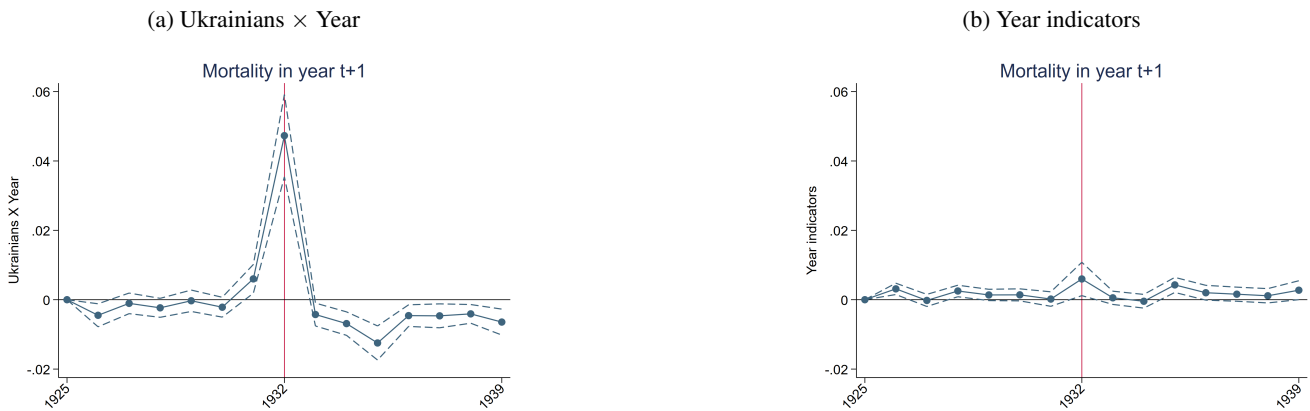
Source: see Appendix D

Figure 2: Grain Production and Procurement per Capita in the Soviet Union



Source: see Data Appendix, Section D.3

Figure 3: Ethnic Ukrainians and Famine Mortality



Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t + 1$ on the rural share of ethnic Ukrainians interacted with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year indicators (fixed effects), and province fixed effects. The 1925 year indicator omitted for comparison. Two sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

Appendix

A Historical Background

A.1 The Chronology of the Famine

The first news of famine began to circulate during the harvest of 1931. According to the official estimates, production was 17% lower than the previous year.²⁸ News of famine traveled to Moscow, but instead of relaxing the policies that were believed to have caused famine, the government asked to intensify them – increasing grain targets by 20% from 22,1 million tons of procured grain in 1930 to 26,6 million for the following year (Wheatcroft, 2001). In the meantime, seed stock was often consumed by starving peasants. The lack of seed stock and weakened labor force contributed to lowering production in 1932, when procurement remained high initially, but was lowered after the severity of the famine became apparent. Mortality from starvation began to increase at the end of the year in the winter, with the peak of mortality being spring of 1933 and continuing for at least the rest of the year. National mortality rates returned to trend in 1934, although in some places, it may have taken longer to recover. Thus, the famine is typically defined to be 1932-33. The Soviet government did not acknowledge the occurrence of the famine until Gorbachev.

A.2 Basic Facts

Officially reported total per capita grain output in 1931 and 1932 was 1.2 kilograms per person per day, which was slightly below the output in non-famine years. The famine affected most of the U.S.S.R., but mortality rates were notably higher in some regions than in others, and higher in rural areas than urban areas.

The Soviet government aimed to centralize food production and distribution. Food produced in rural areas were procured to be distributed to urban industrial populations and for export. In principle, this meant setting production and procurement targets for each region, leaving peasants with enough subsistence. In practice, food was procured even if retention fell below subsistence levels. There are many documents indicating Stalin's advocacy of over-procurement as a method to discipline the peasants, whom he believed lied about their ability to produce and starved as an act of war against the state (Danilov et al., eds, 1999-2006; Davies and Wheatcroft, 2004).

The key to the regimes was the collectivization of agriculture, which began in late 1929. The goal was to remove private property and move peasants to collective farms where they were given food for subsistence. This necessarily undermined individual incentives. As peasants slaughtered, ate or simply neglected collectivized property, between 1929 and 1932, the number of horses declined by 42%, cattle by 40% (Viola, 1996, p. 70). Those who were persecuted as Kulaks were usually peasants who had acquired land in earlier land reforms and become successful.

Collective farmers worked in teams in the field. Food was produced and stored by the collective, and later delivered to state procurement officers. In principle, production that was surplus to procurement was to be allocated between team members proportional to their effort. However, in practice, payments per labor day were often close to zero. The state aimed to remove all private holdings in 1930, including small personal plots for potatoes, the traditional buffer against famine.

Those who resisted joining collective farms or giving up their assets were punished. As a part of dekulakization campaign, about two million peasants were exiled to Siberia and other remote areas of the Soviet Union, and about a half a million perished Viola (2007). Collectivization could have contributed to famine through several channels, such as reducing grain production because of poor incentives, facilitating higher procurement because of more state control over the production, and by reducing the traditional buffer savings of food. We will study the first two channels in our paper. It may have also deteriorated social networks by breaking traditional family/village units by forcing people to work in relatively artificial work teams and by removing family and friends who resisted collectivization. Unfortunately, we do not have the data to study this latter channel or the one of buffer savings.

²⁸Davies and Wheatcroft (2004), Table 1 reports official 1930 harvest estimate to be 83.5 million tons, and official 1931 harvest estimates to be 69.5 million tons.

A.3 Ethnic Bias and the Ukraine

The accusation that the Soviet government targeted Ukrainians above and beyond Russians is driven by several sources. First, aggregate mortality rates in the Ukraine were much higher than in Russia. Second, the Ukraine, being one of the most agriculturally productive regions, was amongst the most resistant to collectivization. Third, Ukrainian nationalism had been a two-edged sword for the Soviet government. On the one hand, many nationalists sided with the revolutionaries to overthrow the Tsar in 1917, and many were moreover socialists. On the other hand, any nationalism undermines the Bolshevik ideal of building socialism in one country which had been an official goal since 1925. Concerns on national opposition to the regime became stronger in the time of troubles.²⁹

There is little explicit discussion of ethnic bias in early Soviet policies by historians. But there are examples of ethnic tensions and allegiances, as well as evidence that Stalin was well-aware of these and may have utilized them for central governance. For example, he was known to recruit members of ethnic minorities into his secret policy widely (Gregory, 2009).

A good example in the context of the famine in Ukraine is the protest from Ukrainian Party members and Stalin's response. In 1931, party members began to report food shortages to Stalin. As the famine increased, they began to ask Stalin to reduce procurement. In 1932, de-classified secret police reports to Stalin indicated the resignations of ethnic Ukrainian Party leaders at all levels with the explicit explanation being that they could not cooperate with the killing of so many Ukrainians. In response, Stalin sent special commissions headed by his closest deputies Vyacheslav Molotov and Lazar Kaganovich who were not Ukrainian to implement his policies in the republic of Ukraine and the North Caucasus, two key grain producing regions where also the bulk of Ukrainians lived. We will explore the extent that such ethnic allegiances played a role in the famine in the analysis.

B Predicted Grain

Grain production function estimates the relationship between 1901–1915 log grain and log province area, FAO GAEZ grain suitability index, their interaction, fall (October, November, December), winter (January, February, March), spring (April, May, June), summer (July, August, September) temperature and rainfall, their pairwise interactions and square terms (without a constant). We then use the estimated grain production function to predict 1925–1940 grain.

We regress 1901-1915 log grain on log province area, FAO GAEZ grain suitability index, their interaction, temperature and rainfall figures for four seasons, their pairwise interactions and square terms (without a constant), and use the estimated grain production function to predict 1925–1940 grain. The predicted grain and actual grain are closely correlated. The only exception is Karelia region that is most likely a result of potential errors in our matching procedure. In-sample R-squared is 0.99. Out of sample R-squared is 0.86 (Figure A.3 in the appendix). This is consistent with the lack of major technological changes in the Soviet agriculture before the 1930s portrayed by historians (Allen, 2003).

C Other Ethnicities

Appendix Figure A.10 shows the interactions of year dummy variables and the population share for each ethnic group. In Appendix Table A.6, we estimate the heterogeneous treatment effects with the rural share of Kazakhs and Belorussians as explanatory variables. The estimates for Kazakhs typically have the same sign as for Ukrainians, but are usually imprecise. The investigation of Belorussians show that the drivers of mortality for this groups is

²⁹A classified decree of the Central Committee of the Communist party and the Soviet government from December 14, 1932, that is, issued in the middle of the famine, illustrates this point: "... frivolous, not arising from the cultural interests of the population, not Bolshevik 'Ukrainization' of almost half of the regions of the North Caucasus, in the complete absence of control over the Ukrainization of schools and the press by the regional authorities, gave legal form to the enemies of the Soviet government to organize resistance to the activities and tasks of the Soviet government by the kulaks, officers, re-emigrant-Cossacks, members of the Kuban Rada, etc." (Danilov et al., eds, 1999-2006Vol. 3, Document 226).

likely to be quite different from that for Ukrainian. The triple interaction with grain productivity in 1928 is negative and statistically significant. The presence of urban communists increased mortality for Belorussians.

D Data Appendix

D.1 Panel data (seventeen provinces spanning European part of the Soviet Union, 1925–1940)

D.1.1 Total and urban population

- 1897: census
- 1898–1903: interpolated between 1897 and 1904
- 1904–1913: Yezhegodnik Rossii 1904–1916
- 1925: Statisticheskiy yezhegodnik 1924 god (vypusk pervyy). Trudy TsSU. Tablitsa 1.B Naseleniye po guberniyam. Naseleniye po ischisleniyu na 1 yanvaryu 1925 goda.
- 1926: interpolated between 1925 and 1927
- 1927: used data from December 1926 Population Census
- 1928–1932: interpolated between 1927 and 1933
- 1933: RGAE 1562/329/19 p. 1–12
- 1934–1936: interpolated between 1933 and 1937
- 1937: census, Zhiromskaya, V.B. and Kiselev, I.N. and Polyakov, Yu.A., “Polveka pod grifom “sekretno”: Vsesoyuznaya perepis naseleniya 1937 goda [Classified for half a century: All-Union population census of 1937]”, 1996
- 1938: interpolated between 1937 and 1939
- 1939: census corrected for centralized additions (pripiski)
- 1940: used 1939 value
- 1946: RGAE 1562/20/626
- 1947: RGAE 1562/20/684
- 1948: RGAE 1562/329/3802
- 1949–1950: RGAE 1562/329/4464, 4465
- 1951–1958: interpolated between 1950 and 1959
- 1959: census
- 1960: interpolated between 1959 and 1961
- 1961–1966: Narkhoz RSFSR, Narkhoz SSSR
- 1967: interpolated between 1966 and 1968
- 1968: Narkhoz RSFSR, Narkhoz SSSR
- 1969: interpolated between 1968 and 1970

- 1970: census
- 1971–1976: Narkhoz RSFSR, Narkhoz SSSR
- 1977: interpolated between 1976 and 1978
- 1978: Narkhoz RSFSR, Narkhoz SSSR
- 1979: census
- 1980–1986: Narkhoz RSFSR, Narkhoz SSSR
- 1987–1988: interpolated between 1986 and 1989
- 1989: census
- 1990: used 1989 value

Except 1933, we calculated population data in administrative borders corresponding to our provinces using hand-created ArcGIS maps (each year is reported using a different administrative division). This operation is legitimate because reported data are more disaggregated than our provinces. 1933 used as reported. 1900–1914 and 1946–1990 data are only used for Figure 1a.

D.1.2 Births and deaths, rural and total

- 1899–1914 (total deaths and births only, no rural): Yezhegodnik Rossii 1904–1916
- 1925 (total deaths and births only, no rural): Yestestvennoye dvizheniye naseleniya Soyuzs SSR 1923–1925, Izdaniye TsSU SSSR (1928), Table 1
- 1926 (total deaths and births only, no rural): Yestestvennoye dvizheniye naseleniya Soyuzs SSR v 1926 g, Izdaniye TsSU SSSR (1929), Table 1
- 1927–1932: Belarus, Ukraine – RGAE 1562/329/256; Russia – Demoscope
- 1933–1940: Demoscope
- 1946–1990: Demoscope, Narkhoz UkrSSR, Narkhos BelSSR (1986 not reported, appears to be still classified, probably because of Chernobyl)

Except 1933, we calculated births and deaths data in administrative borders corresponding to our provinces using hand-created ArcGIS maps (each year is reported using a different administrative division). This operation is legitimate because reported data are more disaggregated than our provinces. 1933 used as reported. 1900–1914 and 1946–1990 data are only used for Figure 1a.

D.1.3 Ethnic composition (total, urban, rural)

The 1926 population census; data calculated in our province borders using hand-created district-level 1926 map.

D.1.4 Grain harvest, sown area, and yield

- 1901–1914: Obukhov Dvizheniye urozhayev zernovykh kul'tur v Yevropeyskoy Rossii v period 1883-1915 g.g. and Yezhegodnik Rossii 1904–1916 (used to estimate grain production function)
- 1925–1927: Sel'skoye khozyaystvo SSSR 1925–1928 (Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii). Statisticheskoye izdatel'stvo TSSU SSSR, Moskva, (1929)
- 1928: RGAE 1562/329/1409
- 1929–1930 (grain harvest only, no sown area or yield): XX
- 1931: Sel'skoye khozyaystvo SSSR. Yezhegodnik 1935. Gosudarstvennoye izdatel'stvo kolkhoznoy i sovkhoznoy literatury "Sel'khozgiz", Moskva, (1936), p. 270
- 1932–1940: RGAE 1562/329/1409

We calculated grain data in administrative borders corresponding to our provinces using hand-created ArcGIS maps (each year is reported using a different administrative division). This operation is legitimate because reported data are more disaggregated than our provinces.

D.1.5 Procurement

- 1925: Yezhegodnik khlebnoy trgovli N1, Table 14
- 1926: Yezhegodnik khlebnoy trgovli N1, Table 22
- 1927: Sel'skoye khozyaystvo SSSR 1925–1928 (Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii). Statisticheskoye izdatel'stvo TSSU SSSR, Moskva, (1929)
- 1928: Calculated from grain harvest and procurement as a share of harvest from RGAE 4372/30/871 p. 30
- 1929: Yezhegodnik khleboroba N4, Table 3 and Table 10
- 1930: Yezhegodnik khleboroba N4, Table 29 and Table 36
- 1931: Yezhegodnik khlebooborota za 1931–32, 1932–33 i predvaritel'nyye itogi zagotovok 1933 g. Table 21
- 1932: Yezhegodnik khlebooborota za 1931–32, 1932–33 i predvaritel'nyye itogi zagotovok 1933 g. Table 33
- 1933: Yezhegodnik khlebooborota za 1931–32, 1932–33 i predvaritel'nyye itogi zagotovok 1933 g. Table 53

We calculated 1925–1927 procurement data in administrative borders corresponding to our provinces using hand-created ArcGIS maps (each year is reported using a different administrative division). This operation is legitimate because reported data are more disaggregated than our provinces. 1928–1933 data is used as reported.

D.1.6 Retention

Retention is grain production minus procurement.

D.1.7 Collectivization

- 1927: Kollektivizatsiya sovetskoy derevni (Predvaritel'nyye itogi sploshnykh obsledovaniy 1928 i 1929 gg) II izdaniye
- 1928: RGAE 1562/82/271
- 1929: Kolkhozy v 1929 godu
- 1930: Kolkhozy v 1930 godu
- 1931: Kolkhoznoye stroitel'stvo v SSSR M. (1931) p. 15 and Davies & Wheatcroft The Years of Hunger, Table 27
- 1932: RGAE 1562/82/271
- 1933: Plan Magazine 2-1933
- 1934–1936: RGAE 1562/82/271
- 1937: interpolated between 1936 and 1938
- 1938: Sel'skoye khozyaystvo Soyuza SSR 1939 (Statisticheskiy spravochnik). Razdel 4
- 1939: used 1938 figures

Collectivization is the share of rural households in collective farms.

D.1.8 1917 Elections

Bolshevik votes and Nationalist votes are from Protasov et al. (2014). Data calculated in our province borders using district-level 1917 map from Castañeda Dower and Markevich (2020).

D.1.9 Communists (total, urban, rural)

Communists (total, urban, rural) is the average number of Communist Party members and candidates for members over 1922, 1927, and 1931.

- 1922: Vserossiyskaya perepis' chlenov RKP 1922 goda, vyp 3 (po uyezdnam)
- 1927: Vsesoyuznaya partiynaya perepis' 1927 goda, 1yy vypusk, Chislennyy sostav VKP(b.) na 10 yanvaryaya 1927 g.
- 1931: Sostav VKP(b) v tsifrah. Vypusk XI 1931

We calculated 1922 and 1927 communists data in administrative borders corresponding to our provinces using hand-created ArcGIS maps (each year is reported using a different administrative division). This operation is legitimate because reported data are more disaggregated than our provinces. 1931 data is used as reported.

D.1.10 Voting delegates 1930

We collected location and ethnicity of all 1930 Party Congress delegates that served as province-, district-, city-, or borough-level Party secretary from Rossiyskiy Gosudarstvennyy Arkhiv Sotsial'no-Politicheskoy Istorii (Russian State Archive of Socio-Political History, RGASPI), Fund 58, Register 1, Files 1-16.

D.2 Data for the 1892 famine

For the placebo we use data from 50 European provinces of the Russian Empire.

D.2.1 Population (urban, rural, total)

- 1885–1896: kindly shared by Volha Charnysh (Charnysh and McElroy, 2020)
- 1897: census
- 1898: interpolated between 1897 and 1899
- 1899–1914: Yezhegodnik Rossii 1904–1916

D.2.2 Deaths, births (total only, no rural)

- 1885–1896: kindly shared by Volha Charnysh (Charnysh and McElroy, 2020)
- 1899–1914: Yezhegodnik Rossii 1904–1916

D.2.3 Ethnic composition

1897 Population Census

D.2.4 Grain, sown area, yield

- 1885–1914: Obukhov Dvizheniye urozhayev zernovykh kul'tur v Yevropeyskoy Rossii v period 1883-1915 g.g.

D.3 Sources of Figure 2

D.3.1 Total Grain Production in the Soviet Union

- 1916–1917, 1920–1922: Sbornik statisticheskikh svedeniy po Soyuzu SSR 1918–1923. Za pyat' let raboty Tsentral'nogo Statisticheskogo Upravleniya. Tsentralnoye Statisticheskoye Upravleniye (1924)
- 1923: Statisticheskiy yezhegodnik 1922–1923 g. (Vypusk pervyy). Trudy Tsentralnogo Statisticheskogo Upravleniya Tom VIII Vypusk 5. Tsentralnoye Statisticheskoye Upravleniye (1924)
- 1924–1917: Yezhegodnik khlebnoy trgovli N1
- 1928: RGAE 1562/329/1409
- 1929–1931: Sel'skoye khozyaystvo SSSR. Yezhegodnik 1935. Gosudarstvennoye izdatel'stvo kolkhoznoy i sovkhoznoy literatury "Sel'khozgiz", Moskva, (1936)
- 1932–1940: RGAE 1562/329/1409; 1940 figure excludes western regions occupied by the Soviet Union in 1939.

D.3.2 Total Grain procurement in the Soviet Union

- 1916–1940: RGAE 8040/8/360

D.3.3 Total Population of the Soviet Union

- 1916: used 1914 figure from Statisticheskiy yezhegodnik 1918–1920 gg. (Vypusk pervyy). Trudy Tsentralnogo Statisticheskogo Upravleniya Tom VIII Vypusk 1. Tsentralnoye Statisticheskoye Upravleniye, Moskva (1921)
- 1920: Statisticheskiy yezhegodnik 1918–1920 gg. (Vypusk pervyy). Trudy Tsentralnogo Statisticheskogo Upravleniya Tom VIII Vypusk 1. Tsentralnoye Statisticheskoye Upravleniye, Moskva (1921)
- 1921–1922: interpolated between 1920 and 1923
- 1923: Sbornik statisticheskikh svedeniy po Soyuzu S.S.R. 1918–1923. Za pyat let raboty Tsentralnogo Statisticheskogo Upravleniya. Tsentralnoye Statisticheskoye Upravleniye, Moskva (1924)
- 1924–1926: interpolated between 1923 and 1927
- 1927: used December 1926 Population Census
- 1928–1933: RGAE 4372/30/107
- 1934–1936: interpolated between 1933 and 1937
- 1937: census, Zhiromskaya, V.B. and Kiselev, I.N. and Polyakov, Yu.A., “Polveka pod grifom “sekretno”: Vsesoyuznaya perepis naseleniya 1937 goda [Classified for half a century: All-Union population census of 1937]”, 1996
- 1938: interpolated between 1937 and 1939
- 1939: census corrected for centralized additions (pripiski)
- 1940: used 1939 value

Table A.1: Main Ethnic Groups in the Soviet Union

Ethnicity	1926 census		1939 census	
	(1) Population, millions	(2) % U.S.S.R.	(3) Population, millions	(4) % U.S.S.R.
Russians	77.8	52.9	99.6	58.4
Ukrainians	31.2	21.2	28.1	16.5
Belorussians	4.7	3.2	5.3	3.1
Kazakhs	4.0	2.7	3.1	1.8
Jews	2.6	1.8	3.0	1.8
Germans	1.2	0.8	1.4	0.8
Poles	0.8	0.5	0.6	0.4

Table A.2: 20 Largest Ethnic Groups in Belarus, Russia, and Ukraine

(1) Ethnicity	1926 census		(4) Ethnicity	1939 census	
	(2) Population, millions	(3) % U.S.S.R.		(5) Population, millions	(6) % U.S.S.R.
Russians	75.3	59.3	Russians	94.8	65.0
Ukrainians	30.1	23.7	Ukrainians	27.1	18.6
Belorussians	4.7	3.7	Belorussians	5.2	3.6
Tatars	3.0	2.3	Tatars	4.0	2.7
Jews	2.5	2.0	Jews	2.9	2.0
Mordvins	1.3	1.0	Mordvins	1.4	1.0
Germans	1.1	0.9	Chuvashs	1.4	0.9
Chuvashs	1.1	0.9	Germans	1.3	0.9
Poles	0.7	0.6	Bashkirs	0.8	0.6
Bashkirs	0.7	0.6	Votyaks (Udmurts)	0.6	0.4
Votyaks (Udmurts)	0.5	0.4	Poles	0.6	0.4
Maris	0.4	0.3	Maris	0.5	0.3
Chechens	0.3	0.2	Komi (Zyryans)	0.4	0.3
Moldovans	0.3	0.2	Chechens	0.4	0.3
Karels	0.2	0.2	Kazakhs	0.4	0.3
Yakuts	0.2	0.2	Moldovans	0.3	0.2
Buryats	0.2	0.2	Karels	0.3	0.2
Komi (Zyryans)	0.2	0.2	Yakuts	0.2	0.2
Armenians	0.2	0.2	Armenians	0.2	0.2
Greeks	0.2	0.1	Avars	0.2	0.2

Table A.3: The Correlation Between Mortality, Grain Production, and Procurement

	Dependent Variable: Mortality in year t+1							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grain	0.000	-0.003	0.000	0.001				
	[0.001]	[0.003]	[0.001]	[0.001]				
Grain x Famine			0.011**	-0.019***				
			[0.005]	[0.007]				
Predicted grain					0.000	0.000	0.000	0.001
					[0.000]	[0.000]	[0.000]	[0.000]
Predicted grain x Famine							0.014**	0.008
							[0.006]	[0.006]
Procurement		0.009		-0.002		0.003		0.001
		[0.006]		[0.004]		[0.004]		[0.002]
Procurement x Famine				0.068***				0.026**
				[0.016]				[0.010]
Urbanization	-0.012**	0.011	-0.011*	0.012	-0.012**	0.015	-0.010*	0.019*
	[0.006]	[0.009]	[0.006]	[0.008]	[0.006]	[0.010]	[0.005]	[0.010]
Urbanization x Famine	0.001	-0.002	0.014	-0.014	0.001	-0.004	0.020	0.011
	[0.015]	[0.014]	[0.021]	[0.014]	[0.015]	[0.014]	[0.018]	[0.018]
Observations	255	149	255	149	255	149	255	149
R-squared	0.612	0.594	0.662	0.776	0.612	0.587	0.688	0.765

Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Grain and procurement are measured in kilograms per person per day. Famine is an indicator that equals one in 1932 and zero otherwise. All regressions control for province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table A.4: Heterogeneous Effects of Politics on Collectivization, Grain, Procurement, and Retention

	Panel A: Dependent variable: Collectivization					Panel B: Dependent variable: Grain				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ukrainians x Famine	0.512 [0.648]	0.335 [0.453]	2.431** [1.116]	3.490 [6.371]	0.976* [0.521]	2.767 [1.837]	2.314 [1.611]	-0.285 [2.793]	3.848 [8.256]	1.345 [1.605]
Grain 1928 x Famine	0.110** [0.050]					-0.114 [0.133]				
Ukrainians x Grain 1928 x Famine	0.755 [0.470]					-1.875 [1.912]				
Bolshevik votes 1917 x Famine		-0.532*** [0.073]					0.468 [0.316]			
Ukrainians x Bolshevik votes 1917 x Famine		9.896*** [2.240]					-9.690 [10.993]			
Nationalist votes 1917 x Famine			0.132 [0.113]					-0.584 [0.410]		
Ukrainians x Nationalist votes 1917 x Famine			-1.118** [0.502]					1.674 [1.302]		
Urban communists x Famine				-0.008 [0.008]					0.007 [0.030]	
Ukrainians x Urban communists x Famine				0.555 [2.466]					1.709 [3.285]	
Rural communists x Famine				0.050** [0.021]					-0.096 [0.059]	
Ukrainians x Rural communists x Famine				0.181 [0.276]					-0.595 [0.599]	
Voting delegates 1930 x Famine					0.029 [0.342]					-0.634 [0.692]
Ukrainians x Voting delegates 1930 x Famine					10.572*** [3.397]					-12.732 [12.787]
Observations	221	221	221	221	221	272	272	272	272	272
R-squared	0.969	0.969	0.968	0.969	0.968	0.809	0.808	0.809	0.809	0.809
	Panel C: Dependent variable: Procurement share					Panel D: Dependent variable: Retention				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ukrainians x Famine	-0.129 [0.201]	0.186 [0.265]	1.490** [0.703]	-1.973 [1.733]	0.492** [0.217]	1.297 [1.257]	0.158 [1.136]	-2.409 [1.991]	2.696 [9.366]	-0.631 [1.051]
Grain 1928 x Famine	0.037* [0.019]					-0.033 [0.133]				
Ukrainians x Grain 1928 x Famine	0.858*** [0.261]					-2.386* [1.317]				
Bolshevik votes 1917 x Famine		-0.148* [0.077]					0.226 [0.330]			
Ukrainians x Bolshevik votes 1917 x Famine		4.631* [2.445]					-9.588 [8.033]			
Nationalist votes 1917 x Famine			0.160*** [0.053]					-0.344 [0.392]		
Ukrainians x Nationalist votes 1917 x Famine			-0.824*** [0.308]					1.678* [0.917]		
Urban communists x Famine				-0.001 [0.007]					0.018 [0.028]	
Ukrainians x Urban communists x Famine				-1.477** [0.668]					2.108 [3.605]	
Rural communists x Famine				0.003 [0.012]					-0.053 [0.051]	
Ukrainians x Rural communists x Famine				0.451*** [0.098]					-0.702 [0.451]	
Voting delegates 1930 x Famine					0.117 [0.149]					-0.782 [0.610]
Ukrainians x Voting delegates 1930 x Famine					6.661*** [2.054]					-11.897 [7.614]
Observations	149	149	149	149	149	149	149	149	149	149
R-squared	0.913	0.909	0.911	0.912	0.91	0.797	0.795	0.796	0.799	0.797

Notes: Observations are at the province and year level. Collectivization is the share of rural households in collective farms. Grain is the grain harvest measured in kilograms per person per day. Procurement share is the share of harvest procured by the government. Retention is grain production minus procurement measured in kilograms per person per day. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Census. Famine is an indicator that equals one in 1932 and zero otherwise. Grain 1928 is the 1928 grain harvest measured in kilograms per person per day. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. Nationalist votes 1917 is the share of votes for nationalistic parties in the 1917 elections. Urban (rural) communists is the average over 1922, 1927, and 1931 of the number of urban (rural) Communist Party members and candidates per 1,000 people. Voting delegates 1930 is the number of province-, county-, city-, and borough-level Party secretaries that participated and had a right to vote in the 1930 Party Congress per 100,000 people. All regressions control for Urbanization, Urbanization x Famine, Ukrainians x Urbanization x Famine, Predicted grain, Predicted grain x Famine, Ukrainians x Predicted grain x Famine, province, and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table A.5: A Reproduction of Table 1 with District Data

	Dependent variable: Mortality					
	Baseline (1)	Province- Year FE (2)	Total Ukrainians (3)	Urban Ukrainians (4)	Grain suitability (5)	Omit Ukraine SSR (6)
Ukrainians x Famine	0.048*** [0.002]	0.042*** [0.005]	0.050*** [0.002]	0.063*** [0.003]	0.046*** [0.002]	0.047*** [0.005]
<i>Normalized</i>	<i>0.611</i>	<i>0.543</i>	<i>0.618</i>	<i>0.622</i>	<i>0.593</i>	<i>0.388</i>
Urbanization	0.002 [0.006]	0.002 [0.006]	-0.003 [0.006]	-0.008 [0.007]	0.001 [0.006]	-0.003 [0.005]
Urbanization x Famine	-0.010*** [0.003]	-0.011*** [0.003]	-0.005* [0.003]	-0.004 [0.003]	-0.009*** [0.003]	0.001 [0.003]
Grain suitability x Famine					Y	
Observations	3155	3155	3157	1927	3155	2376
R-squared	0.732	0.764	0.735	0.748	0.734	0.641
Provinces	16	16	16	16	16	15
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y		Y	Y	Y	Y
Province-Year FE		Y				
Mortality 1933						
Mean	0.039	0.039	0.039	0.042	0.039	0.031
Std. Dev.	0.029	0.029	0.029	0.03	0.029	0.021
Ukrainians						
Mean	0.257	0.257	0.245	0.228	0.257	0.066
Std. Dev.	0.378	0.378	0.362	0.295	0.378	0.171

Notes: Observations are at the district and year level. Ukrainians is the share of ethnic Ukrainians in the rural population (Columns 1--2, 5--6), total population (Column 3), or urban population (Column 4) according to the 1926 Census. Famine is an indicator that equals one in 1933 and zero in 1928. Estimates in Column 5 control for FAO GAEZ Grain suitability x Famine. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table A.6: Heterogeneous Effects of Politics on Other Ethnic Groups

	Dependent variable: Mortality in year t+1											
	Kazakhs						Belorussians					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ethnic Group Share x Famine	-4686.277*** [1351.411]	66.863*** [19.750]	-1758.032*** [526.997]	-402.496*** [105.527]	321.132 [340.605]	3467.282*** [1021.708]	-2.643 [1.658]	-6.987*** [2.239]	-5.385** [2.099]	-3.504 [5.969]	-9.449*** [2.258]	-0.031 [2.544]
Grain 1928 x Famine	0.008 [0.007]						0.017** [0.007]					
Ethnic Group Share x Grain 1928 x Famine	3035.170*** [861.256]						-4.230*** [1.488]					
Bolshevik votes 1917 x Famine		0.002 [0.015]		0.022 [0.021]				-0.013 [0.022]		0.027 [0.023]		
Ethnic Group Share x Bolshevik votes 1917 x Famine		3230.403*** [1186.330]		3692.440*** [960.958]				3.911* [2.064]		-2.159 [10.850]		
Nationalist votes 1917 x Famine			0.023* [0.014]	0.029** [0.013]					0.035* [0.018]	0.042 [0.025]		
Ethnic Group Share x Nationalist votes 1917 x Famine			-1311.012*** [386.783]	-361.872*** [87.284]					-6.563** [2.727]	-7.647 [12.879]		
Urban communists x Famine					0.000 [0.002]							-0.001 [0.002]
Ethnic Group Share x Urban communists x Famine					18.056 [16.249]							-2.492 [1.580]
Rural communists x Famine					-0.008 [0.006]							0.006 [0.004]
Ethnic Group Share x Rural communists x Famine					-175.488 [169.465]							-5.002*** [1.319]
Voting delegates 1930 x Famine						0.079*** [0.027]						0.115** [0.051]
Ethnic Group Share x Voting delegates 1930 x Famine						14375.248*** [4239.746]						-8.028 [9.128]
Observations	255	255	255	255	255	255	255	255	255	255	255	255
R-squared	0.803	0.794	0.821	0.827	0.820	0.828	0.782	0.726	0.763	0.769	0.834	0.772

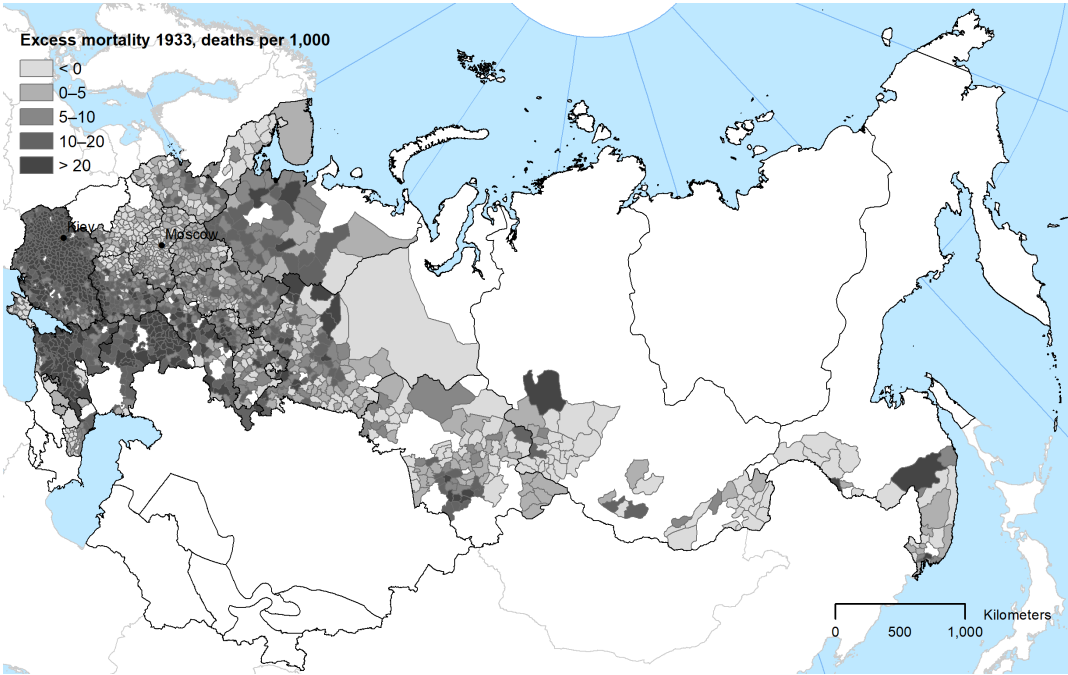
Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Ethnic group share is the share of Kazakhs (Columns 1–6) or Belorussians (Columns 7–12) in the rural population according to the 1926 Population Census. Famine is an indicator that equals one in 1932 and zero otherwise. Grain 1928 is 1928 grain harvest measured in kilograms per person per day. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. Nationalist votes 1917 is the share of votes for nationalistic parties in the 1917 elections. Urban (rural) communists is the average over 1922, 1927, and 1931 of the number of urban (rural) Communist Party members and candidates per 1,000 people. Voting delegates 1930 is the number of province-, county-, city-, and borough-level Party secretaries that participated and had a right to vote in the 1930 Party Congress per 100,000 people. All estimates control for predicted grain, predicted grain x famine. Ethnic group share x predicted grain x famine, urbanization, urbanization x famine, ethnic group share x urbanization x famine, province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Figure A.1: Excess Mortality 1933

(a) Province-level



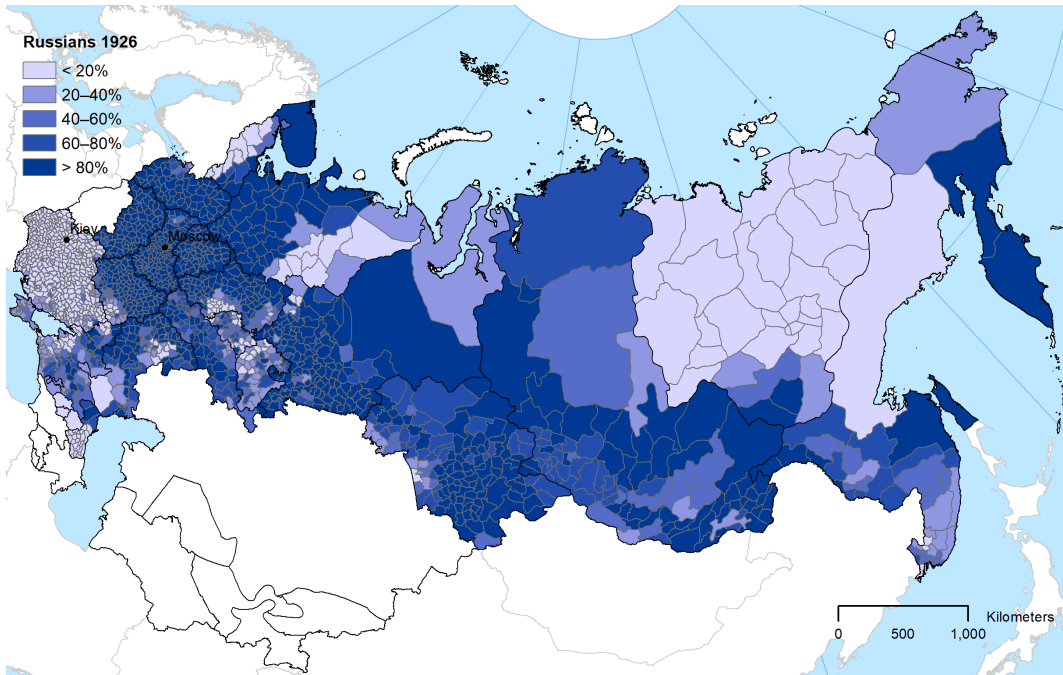
(b) District-level



Note: Excess mortality 1933 is 1933 mortality minus 1928 mortality. Source: see AppendixD.

Figure A.2: Rural Ethnic Composition 1926

(a) Russians

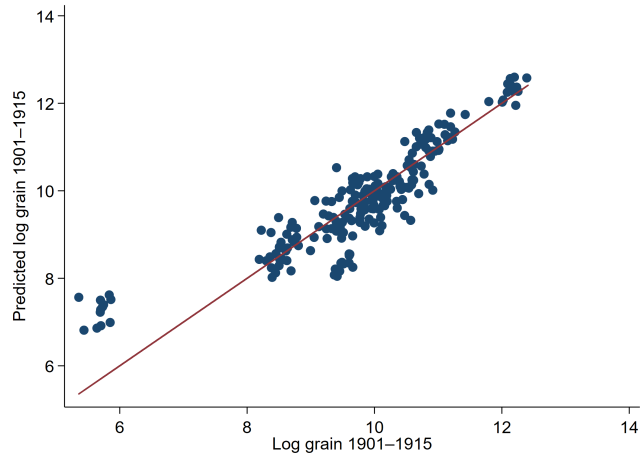


(b) Ukrainians

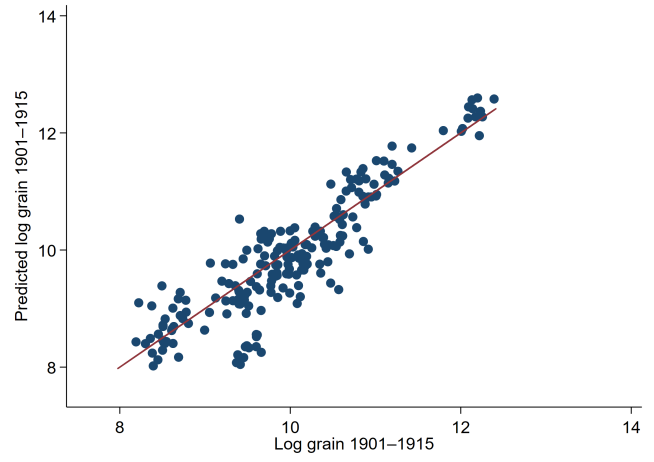


Figure A.3: Reported and Predicted Grain

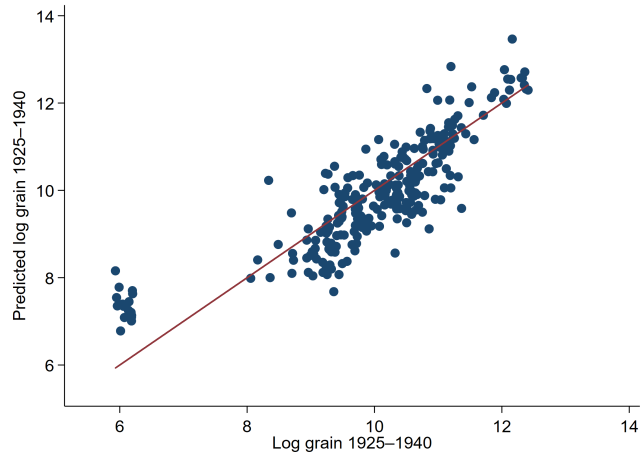
(a) In-Sample Fit



(b) In-Sample Fit Without Karelia



(c) Out-Of-Sample Fit



(d) Out-Of-Sample Fit Without Karelia

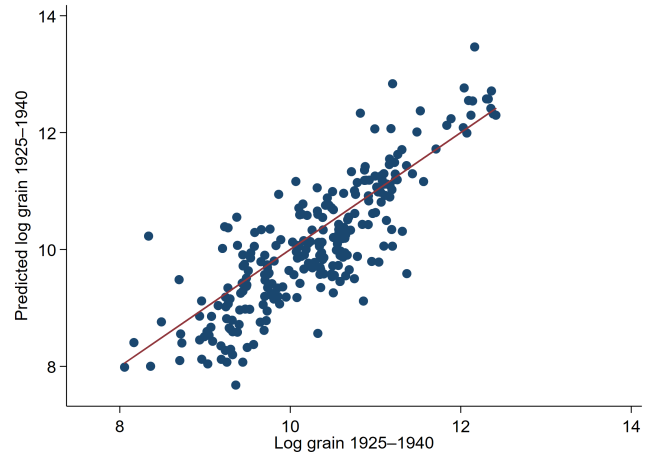
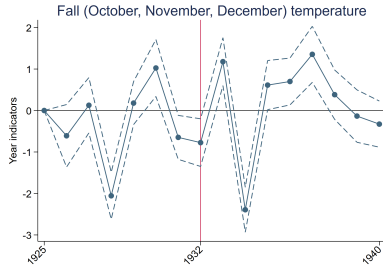
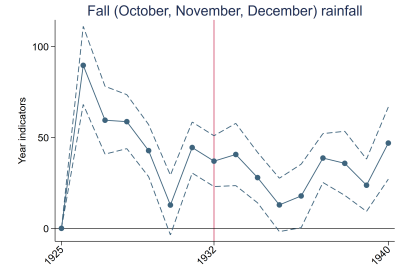


Figure A.4: Seasonal Temperature and Rainfall

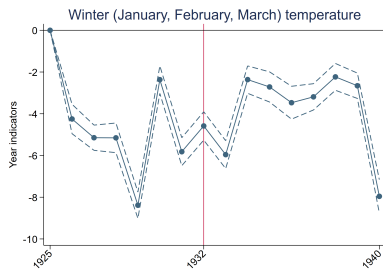
(a) Fall (October, November, December) temperature



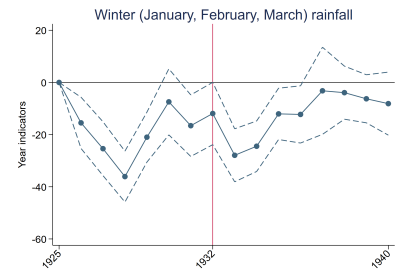
(b) Fall (October, November, December) rainfall



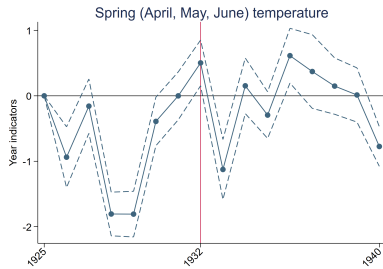
(c) Winter (January, February, March) temperature



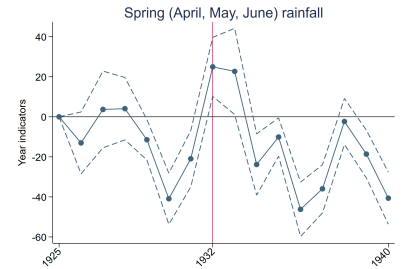
(d) Winter (January, February, March) rainfall



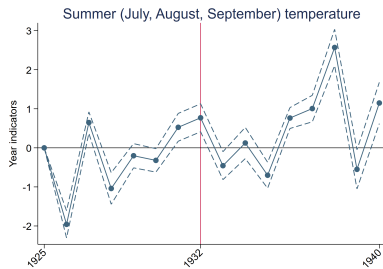
(e) Spring (April, May, June) temperature



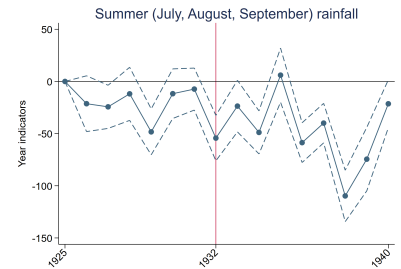
(f) Spring (April, May, June) rainfall



(g) Summer (July, August, September) temperature

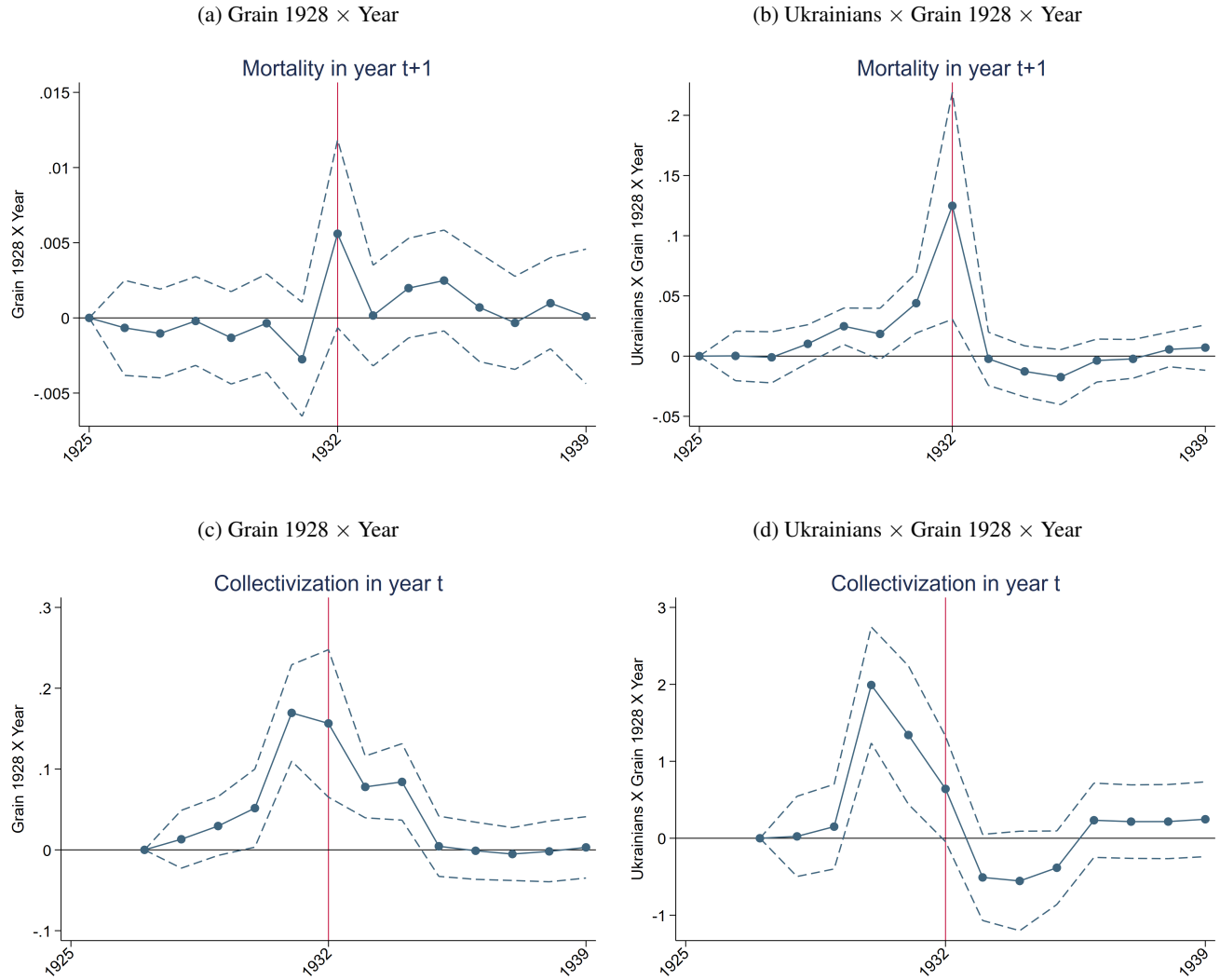


(h) Summer (July, August, September) rainfall



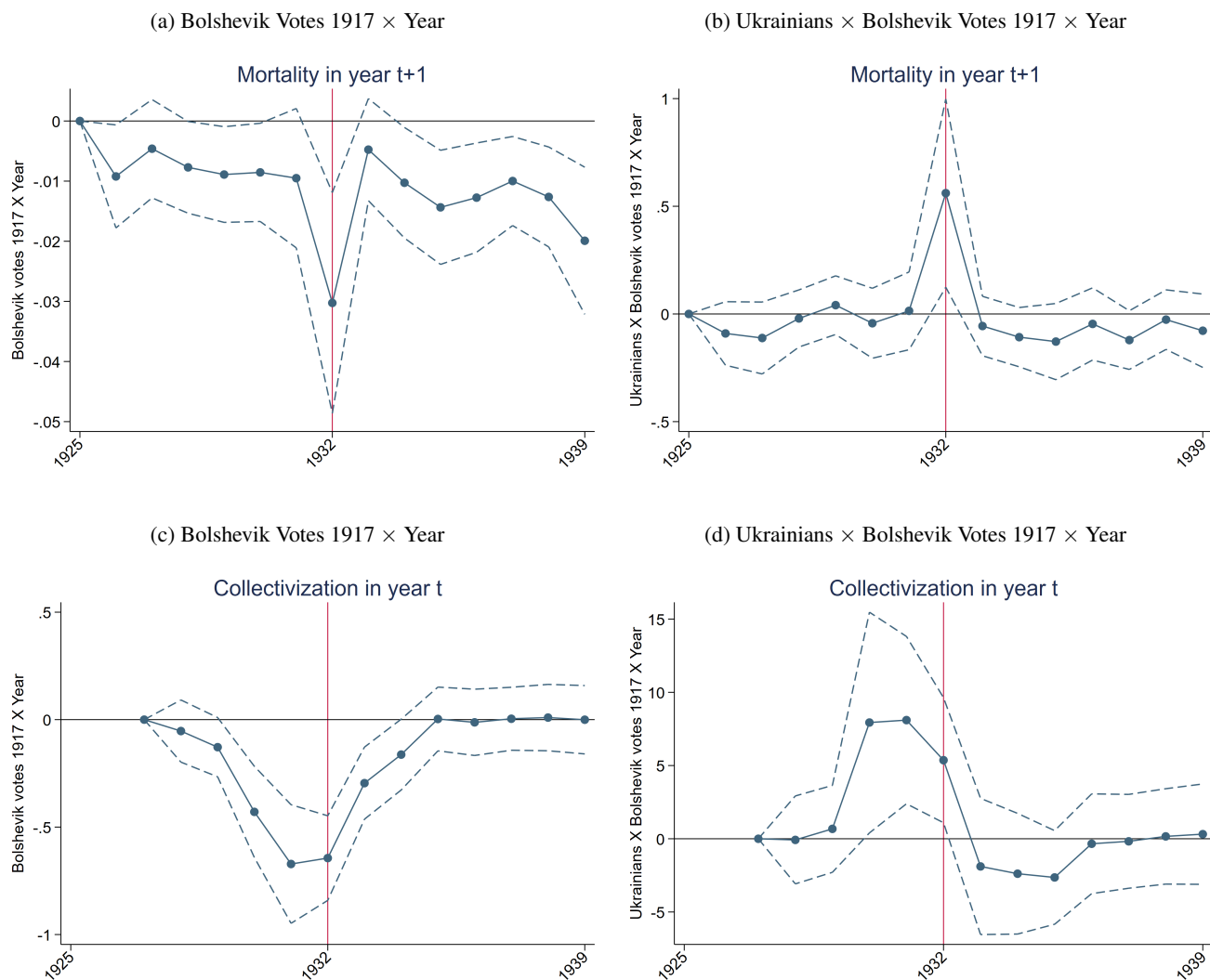
Note: Regression coefficients with their 90% confidence intervals from regressing each season's temperature (rainfall) on year indicators and province fixed effects. The 1925 year indicator is omitted for comparison. Temperature is measured in degrees Celsius. Rainfall is measured in millimeters. Standard errors are robust for heteroskedasticity. Each set of coefficients is from a separate regression.

Figure A.5: Heterogenous Effects of Grain 1928 on Mortality and Collectivization



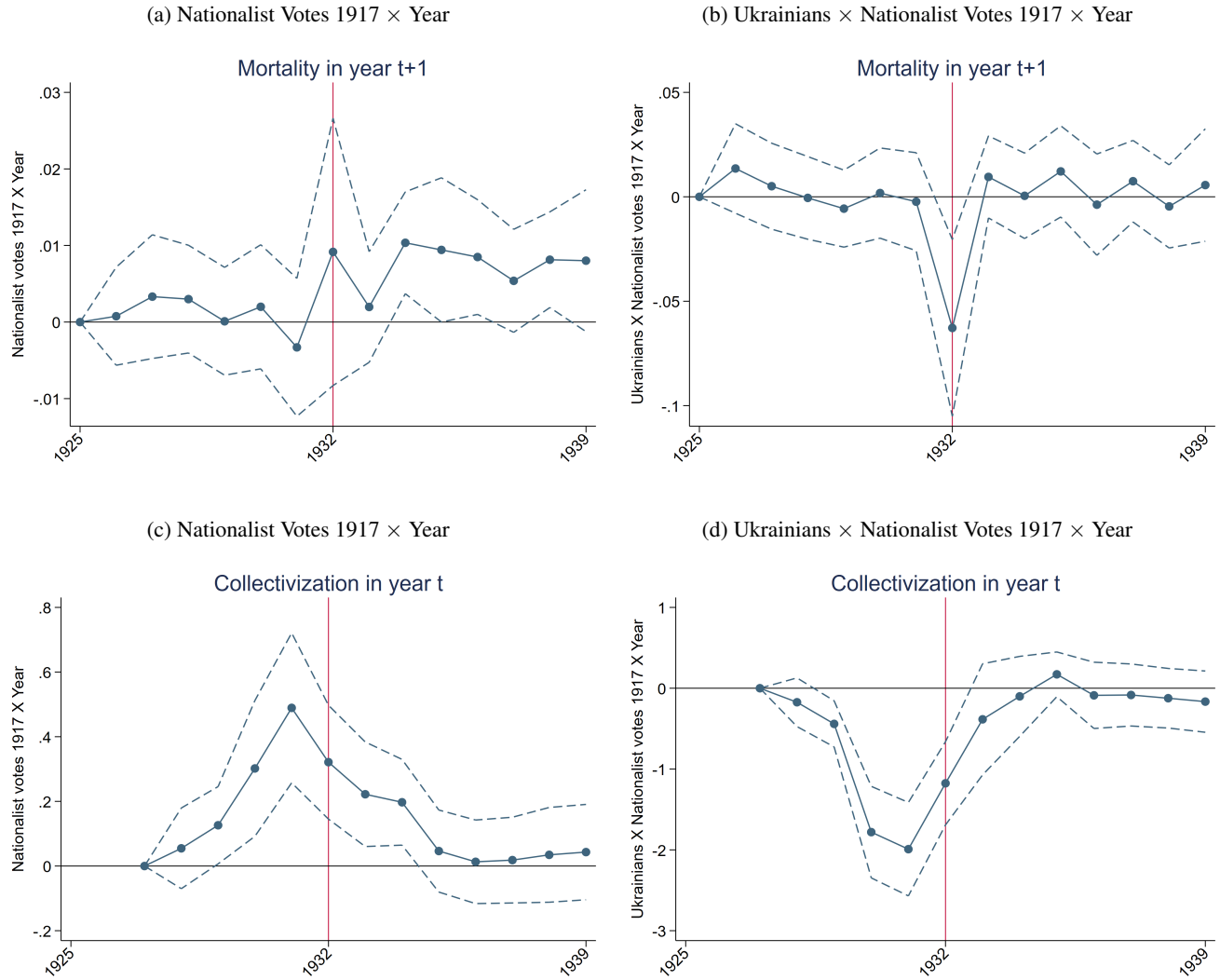
Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t+1$ (Figures a and b) or Collectivization in year t (Figures c and d) on the rural share of ethnic Ukrainians interacted with year indicators, Grain 1928 interacted with year indicators, rural share of ethnic Ukrainians interacted with Grain 1928 and with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year fixed effects, and province fixed effects. Grain 1928 is 1928 grain harvest measured in kilograms per person per day. The 1925 (Figures a and b) or 1927 (Figures c and d) year indicator is omitted for comparison. Two sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

Figure A.6: Heterogenous Effects of Bolshevik Votes 1917 on Mortality and Collectivization



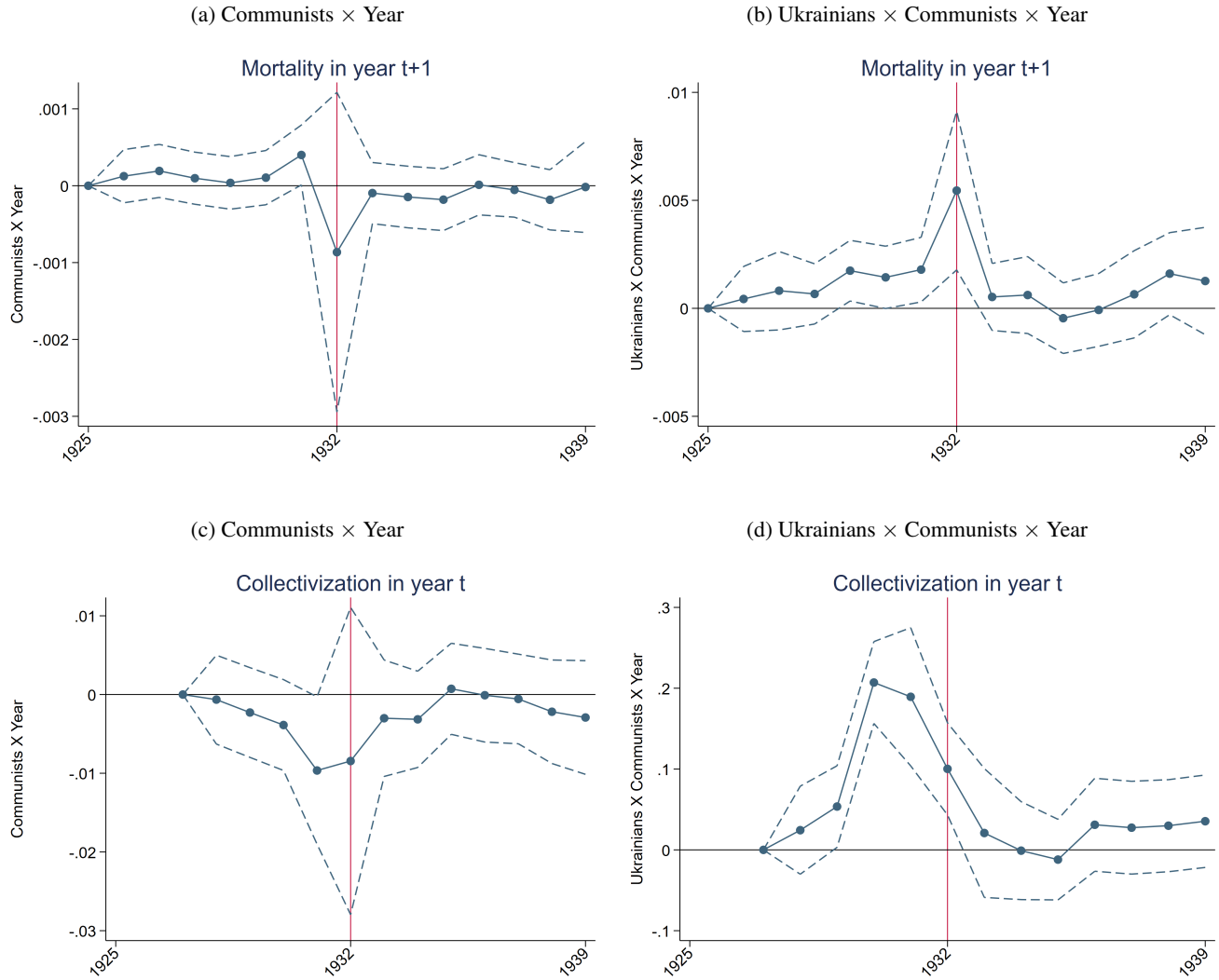
Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t + 1$ (Figures a and b) or Collectivization in year t (Figures c and d) on the rural share of ethnic Ukrainians interacted with year indicators, Bolshevik votes 1917 interacted with year indicators, rural share of ethnic Ukrainians interacted with Bolshevik Votes 1917 and with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year fixed effects, and province fixed effects. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. The 1925 (Figures a and b) or 1927 (Figures c and d) year indicator is omitted for comparison. Two sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

Figure A.7: Heterogeneous Effects of Nationalist Votes 1917 on Mortality and Collectivization



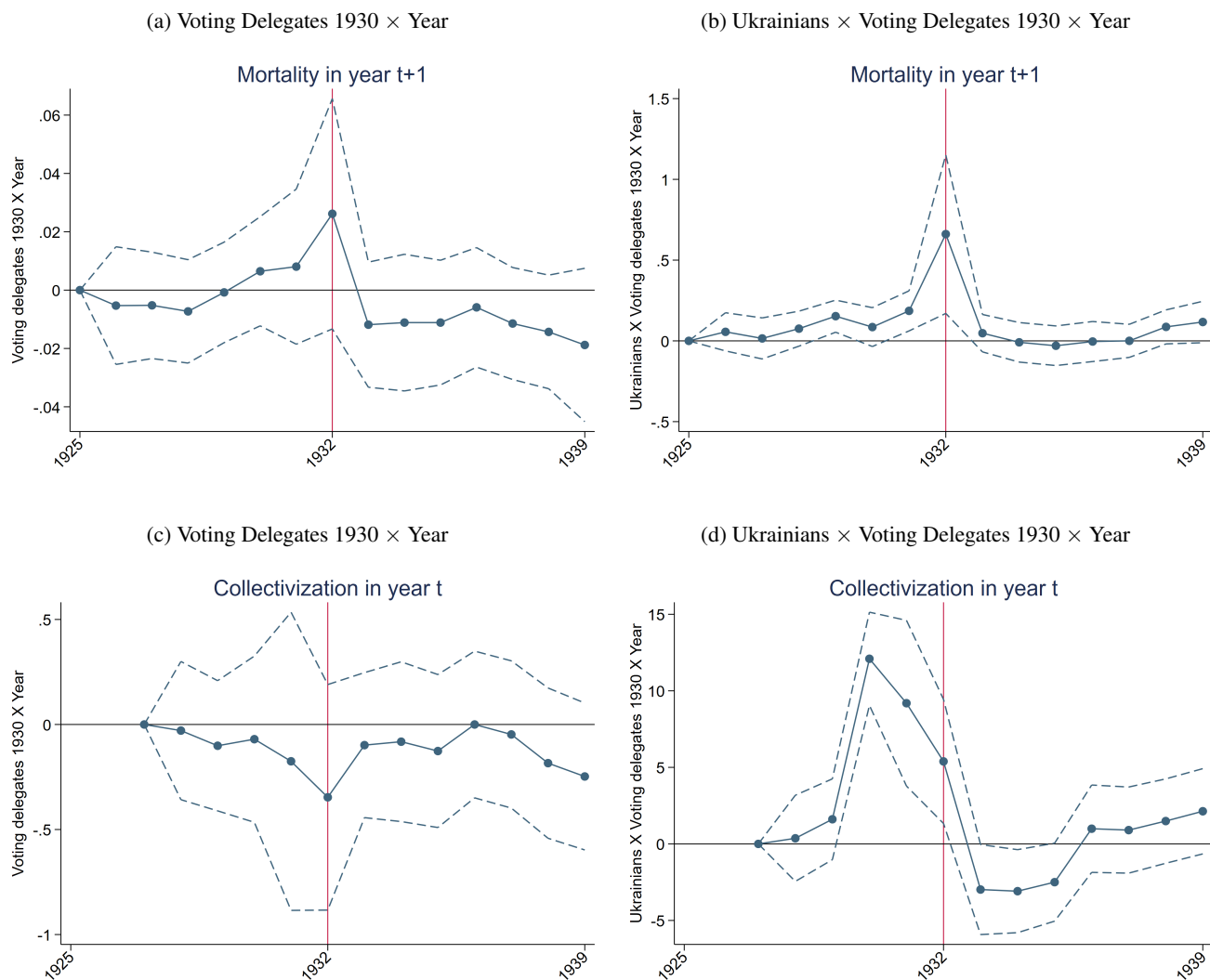
Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t + 1$ (Figures a and b) or Collectivization in year t (Figures c and d) on the rural share of ethnic Ukrainians interacted with year indicators, Nationalist votes 1917 interacted with year indicators, rural share of ethnic Ukrainians interacted with Nationalist votes 1917 and with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year fixed effects, and province fixed effects. Nationalist votes 1917 is the share of votes for nationalistic parties in the 1917 elections. The 1925 (Figures a and b) or 1927 (Figures c and d) year indicator is omitted for comparison. Two sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

Figure A.8: Heterogeneous Effects of Communists on Mortality and Collectivization



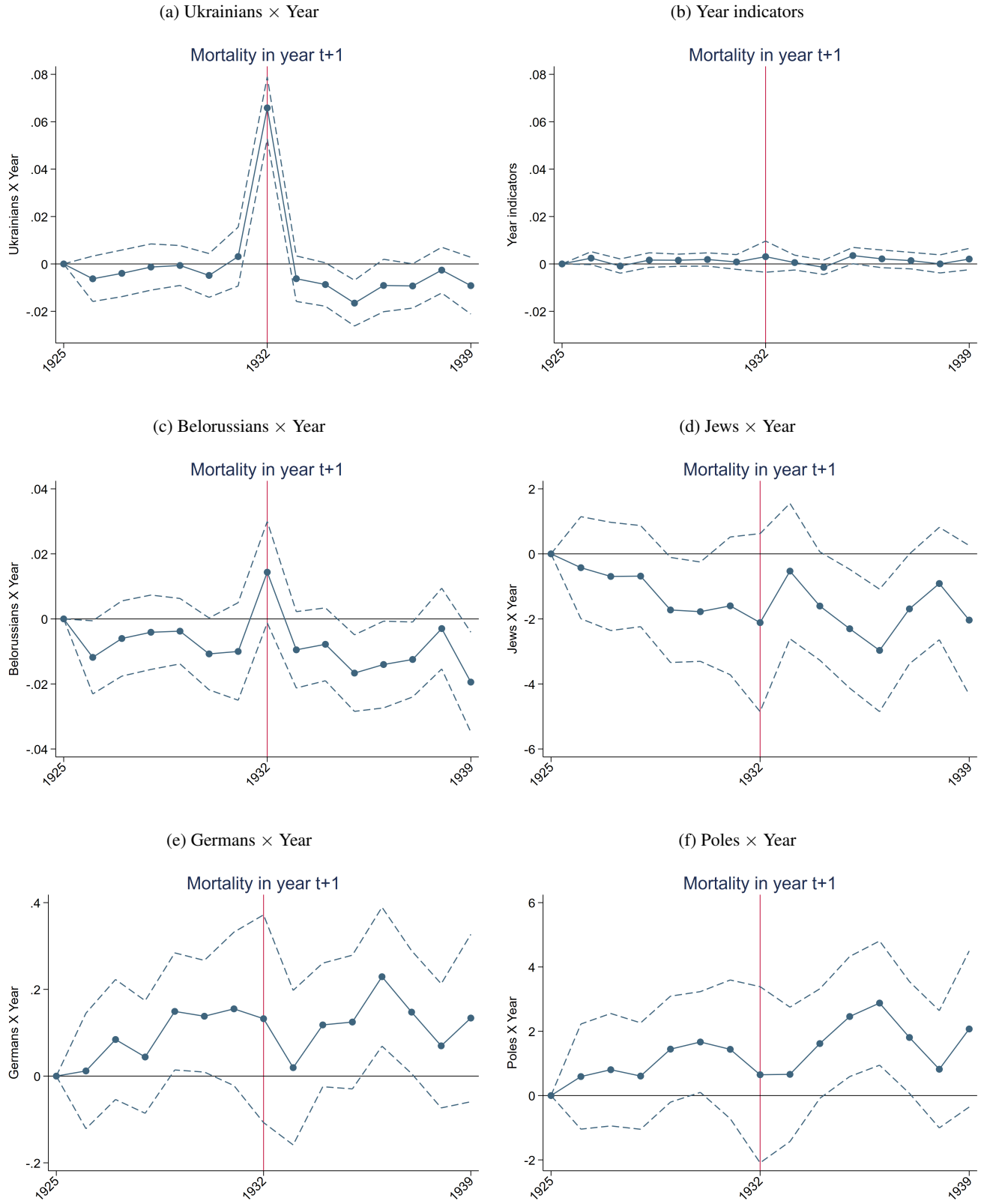
Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t+1$ (Figures a and b) or Collectivization in year t (Figures c and d) on the rural share of ethnic Ukrainians interacted with year indicators, Communists interacted with year indicators, rural share of ethnic Ukrainians interacted with Communists and with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year fixed effects, and province fixed effects. Communists is the 1922–1931 average number of Communist Party members and candidates per 1,000 people. The 1925 (Figures a and b) or 1927 (Figures c and d) year indicator is omitted for comparison. Two sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

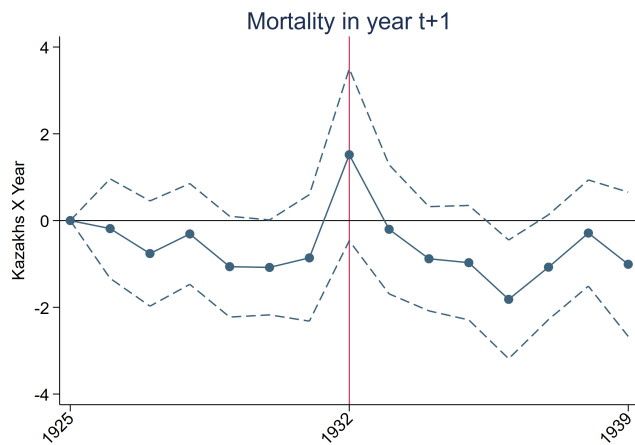
Figure A.9: Heterogeneous Effects of Voting delegates 1930 on Mortality and Collectivization



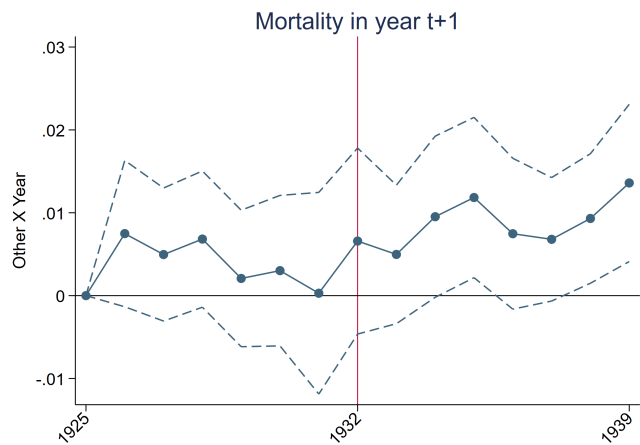
Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t + 1$ (Figures a and b) or Collectivization in year t (Figures c and d) on the rural share of ethnic Ukrainians interacted with year indicators, Voting Delegates 1930 interacted with year indicators, rural share of ethnic Ukrainians interacted with Voting Delegates 1930 and with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year fixed effects, and province fixed effects. Voting Delegates 1930 is the number of province-, county-, city-, and borough-level Party secretaries that participated and had a right to vote in the 1930 Party Congress per 100,000 people. The 1925 (Figures a and b) or 1927 (Figures c and d) year indicator is omitted for comparison. Two sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

Figure A.10: Other Ethnic Groups and Famine Mortality





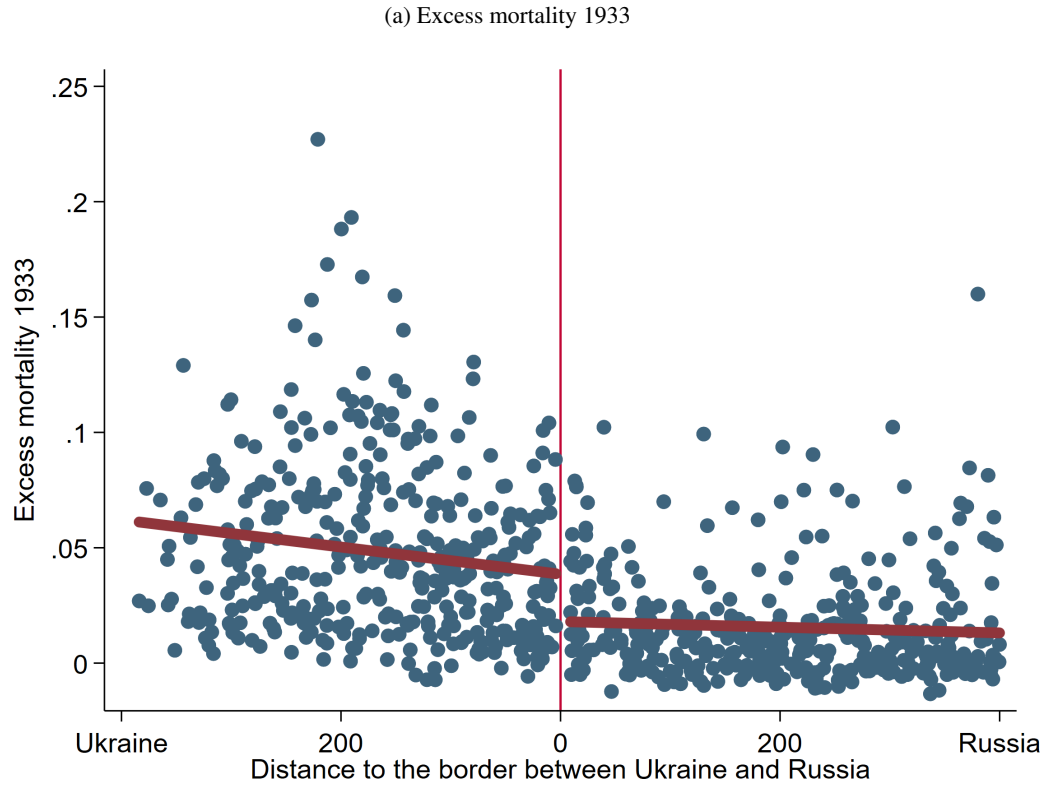
(g) Kazakhs \times Year



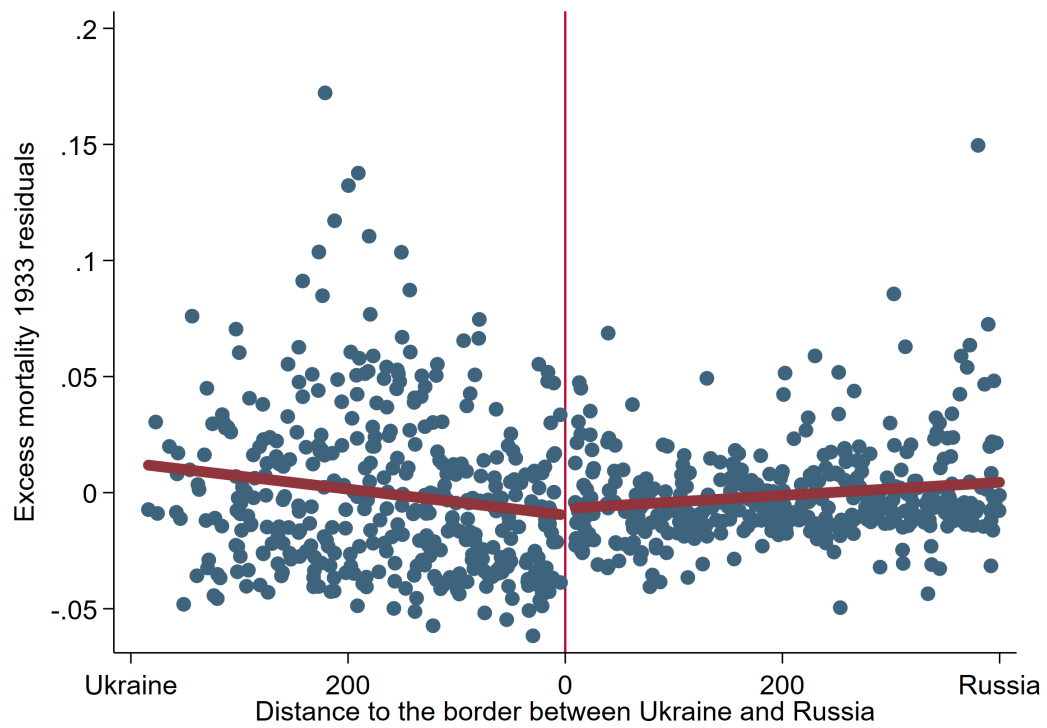
(h) Other \times Year

Note: The figures show regression coefficients with their 90% confidence intervals from regressing mortality in year $t + 1$ on the rural share of ethnic Ukrainians, Belorussians, Jews, Germans, Poles, Kazakhs, and a synthetic group “Other ethnicities” interacted with year indicators, urbanization, urbanization interacted with the famine indicator (that equals to one in 1932 and zero otherwise), year indicators (year fixed effects), and province fixed effects. Other ethnicities is the rural share on non-Ukrainians, non-Belorussians, non-Jews, non-Germans, non-Poles, non-Kazakhs, and non-Russians (omitted category is Russians). The 1925 year indicator is omitted for comparison. Eight sets of coefficients from one regression. Standard errors are robust to heteroskedasticity.

Figure A.10: Excess Mortality 1933 at the border between Ukraine and Russia



(b) Residual excess mortality 1933 after controlling for urbanization and rural share of ethnic Ukrainians



Note: Excess mortality 1933 is 1933 mortality minus 1928 mortality. Distance to the border is measured in kilometers. Source: see AppendixD.