

The Soviet Great Famine, 1932–33*

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

(Incomplete)

This paper investigates the causes of the Soviet Great Famine, 1932–33, and documents several new empirical facts. First, excess mortality was much higher in regions with a higher share of ethnic Ukrainians, even outside of the Soviet Republic of Ukraine. Second, this cannot be explained by differences in natural conditions, grain productivity, demographic structure or urbanization. Third, in regions with a higher share of ethnic Ukrainians, Soviet economic policies were implemented more zealously, which resulted in higher food procurement and famine mortality. Fourth, there is suggestive evidence that mortality was exacerbated by the presence of non-ethnic Ukrainian Communist Party bureaucrats. These and other results in the paper provide novel evidence for the presence of ethnic bias in famine-era Soviet policies and the contribution of ethnic bias to famine mortality.

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

*We express our deepest gratitude to Ekaterina Zhuravskaya, with whom we started this project. For helpful comments and insights, we thank Maxim Ananyev, Paul Catañeda Dower, Michael Ellman, Scott Gehlbach, Mikhail Golosov, Cormac Ó Gráda, Mark Harrison, Oleg Khlevniuk, Victor Malein, Joel Mokyr, Steven Nafziger, Stephen Wheatcroft and the participants of the Northwestern Economic History lunch, HECER (Helsinki Center of Economic Research) Labor and Public Economics seminar and the 2020 Summer Workshop in the Economic History and Historical Political Economy of Russia. For excellent research assistance, we thank Pavel Bacherikov, Nikita Karpov and Ivan Korolev. All mistakes are our own.

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

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

[§]Kellogg School of Management at Northwestern University, nancy.qian@kellogg.northwestern.edu, NBER, BREAD, CEPR

1 Introduction

During the past century, more people perished from famine than from both World Wars combined (Sen, 1981). In just two years, 1932 and 1933, an estimated 5.5 to 10.8 million individuals died in the Soviet Great Famine.¹ In terms of total deaths, this was the second worst famine in the 20th century.² An inextricable fact about the Soviet Great Famine is the variation in mortality rates across ethnic groups. Ethnic minorities such as Ukrainians and Kazakhs suffered higher excess mortality rates than the Russian majority. Ethnic Ukrainians, the largest ethnic minority in the Soviet Union, comprising 21% of the total 1926 population, contributed 30% to 45% of total famine deaths.³ Thus, understanding the causes of disproportionately high Ukrainian mortality is central to understanding the overall causes of the Soviet Famine.⁴

The causes of this famine, and particularly of the disproportionately high Ukrainian mortality, 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 persists about the extent of the state's responsibility versus bad weather and ethnic discrimination in Soviet policies. Existing explanations can be broadly categorized into two groups.

¹Conquest (1986) estimates total famine deaths to be 7 million. Davies and Wheatcroft (2004) gives an estimate of 5.5 to 6.5 million deaths. Ellman (2005) cites “‘about eight and a half million’ victims of famine and repression in 1930–33”. A leading Russian famine historian Victor Kondrashin gives a range between 5 and 7 million victims (Kondrashin, 2008). Russian historical demographers estimate 7.2 to 10.8 million famine victims (Polyakov and Zhiromskaya, eds, 2000). In 2008, Russian State Duma postulated 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 7 million people (State Duma, 2008).

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

³There are no reliable comprehensive data on the ethnic composition of famine victims. For the Soviet Republic of Ukraine (Ukrainian SSR), Meslé et al. (2013) estimates excess deaths to be 2.6 million, Rudnytskyi et al. (2015) — 3.9 million. The most cited total famine death toll is 7 million (footnote 1). 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. To be conservative, assume that within Ukrainian SSR famine victims were distributed among all ethnic groups proportionally to the group size, and that outside Ukrainian SSR none of the famine victims were ethnic Ukrainian (this paper refutes both assumptions). Then, ethnic Ukrainian deaths constitute 30% ($0.8 \cdot 2.6/7$) to 45% ($0.8 \cdot 3.9/7$) of the total famine deaths.

⁴We will discuss other ethnic groups after presenting the main results about Ukrainians.

The first blames Soviet economic policies, which together with bad weather reduced grain production in 1931 and 1932, 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 subject to Soviet agricultural policies (e.g., Davies and Wheatcroft, 2004; Kondrashin, 2008).⁵ The second ascribes an intent of Soviet policy to target Ukrainians in order to increase political and, perhaps, ethnic homogenization of the newly established U.S.S.R. (e.g., Graziosi, 2004; Mace, 2004).⁶ Some go as far as to argue that the famine was a genocide committed on the Ukrainian people.⁷

The main challenge in distinguishing between the two views was the limited availability of disaggregated data for rigorous empirical analysis. Earlier studies rely on narrative and descriptive evidence.

The primary goal of our study is to make progress in understanding the causes of the Soviet Great Famine by unraveling the causes of the disproportionately high ethnic Ukrainian mortality. We collect a vast amount of archival and published data to construct novel province- and district-level panels on mortality, policy outcomes and natural conditions. To the best of our knowledge, we are the first to construct such comprehensive and granular panel data for the Soviet Union. These data allow us to conduct rigorous empirical analysis and distinguish between the prevailing hypotheses by controlling for confounding factors and exploring a rich set of heterogeneous treatment effects.

We begin our analysis with an exploration of the province-level panel, which covers almost the entire inter-war period (1922–1940) and the Soviet republics of Belarus, Ukraine

⁵Also see Kotkin (2017) and Wheatcroft (2017).

⁶Also see, for example, Applebaum (2017), Ellman (2007), Rosefielde (2010) and Snyder (2010).

⁷This interpretation is given as early as Lemkin (2014 based on the speech from 1953) and Conquest (1986). Conquest 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, he 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 controversial verdict.

and almost all Russia. Our sample covers 84% of total Soviet population at the time and 97% of the populations of Belarus, Russia, and Ukraine. We examine the effect of the ethnic Ukrainian population share on famine mortality 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 to account for the urban favoritism of famine-era Soviet food procurement policies. We find that famine mortality rates increase with Ukrainian population share, which is consistent with the claim that ethnic Ukrainians suffered more than the other Soviet ethnicities (in practice, other ethnicities comprise mostly of ethnic Russians). This is true even if we exclude the Soviet Republic of Ukraine (Ukrainian SSR) from our sample, which implies that ethnic delineations are more important than administrative ones for understanding famine mortality.

Moreover, we find that higher mortality rates in areas with a higher share of ethnic Ukrainians were not due to an unhappy coincidence of bad weather in Ukrainian-populated areas. The results are robust to controlling for weather-driven grain production. In fact, temperature and rainfall during the famine were similar to other years when there was no famine. Thus, bad weather cannot explain excess Ukrainian famine mortality, and is unlikely to be an important driver to the overall famine. Similarly, we show that the baseline result is robust to controlling for a large number of proxies for political factors and economic policies (which we discuss below), as well as officially reported grain production, local geography or demographic structure and their interaction with the famine year variable. These results imply that Ukrainian famine mortality is not an artifact of pre-existing conditions or historical features of the region which may have made them a target for Soviet policy. In other words, it was not

the case that Ukrainians happened to have resided in regions with conditions that contributed to the famine.

The next step is to investigate whether Soviet policies drove higher mortality in areas with higher share of Ukrainians during the famine. To do this, we estimate heterogeneous treatment effects of the Ukrainian population share and proxies for political factors and economic policies (i.e., the triple interaction effects of Ukrainian population share, the famine year dummy variable and proxy variables). We find strong evidence of heterogeneous effects. The positive association between Ukrainian population share and famine mortality was higher in regions which 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 more delegates to the 1930 Party Congress to vote on forcing collectivization (a measure of commitment to Stalin and/or collectivization of the local Party elite). These results are consistent with claims that Soviet economic policies more zealously targeted Ukrainians. Interestingly, the estimates for voting delegates show that the number of both ethnic and non-ethnic Ukrainian delegates was associated with higher Ukrainian mortality. However, the magnitude of the effect was larger for non-ethnic Ukrainian delegates. These results are consistent with the belief that there were no formal ethnic divisions within the Communist Party, but also suggest that non-Ukrainian bureaucrats in Ukrainian regions may have been more relentless in enforcing Stalinist policies in the face of the famine.

To delve further into the mechanisms, we repeat these heterogeneous estimates with additional dependent variables which reflect various aspects of Soviet economic policies: collectivization, tractor horsepower (a proxy for mechanization), grain production and grain procurement. We find that places with higher Ukrainian population share experienced higher rates of collectivization, had less mechanization and suffered higher food procurement. This is consistent with the view that collectivization and the over-procurement of food from peasants were important contributors to famine mortality. The results are also consistent with

the view that while Stalinist policies contributing to mortality (collectivization, procurement) were implemented more forcefully on Ukrainians, potentially beneficial resources (mechanization) were diverted away from Ukrainians.

We also construct a district-level panel data set. These data include fewer variables and just two years, 1928 and 1933, but the increased granularity allows us to provide several additional pieces of evidence. 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. Thus, distance from the enforcers of Stalinist policies reduced Ukrainian famine mortality. Finally, we document the presence of a discrete increase in famine mortality when one crosses the border from the Russia to Ukraine, and that this increase dissipates if we control for district-level ethnic Ukrainian share. The results reinforce the importance of ethnicity (as opposed to administrative borders) for understanding Ukrainian famine mortality.

Finally, we conduct two supplementary exercises. First, we examine famine mortality for other ethnic groups in our data. We find evidence consistent with ethnic bias against other groups, but also that the bias may be driven by different policy mechanisms than for Ukrainians (see Section 5 for a more detailed discussion). Second, for comparison, we document that Ukrainian population share was uncorrelated with famine mortality in 1892 (the last large documented famine under the Tsarist regime). Thus, the Soviet-era findings are unlikely to be due to ethnic-specific characteristics which cause higher mortality during famines

(e.g., differences in genetics, social behavior 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 findings support the view that the disproportionately high Ukrainian famine mortality was a result of systematic bias in Soviet policy which targeted ethnic Ukrainians throughout the Soviet Union. Given the large share of famine deaths that comprised of Ukrainians, it follows that ethnic bias was an important contributor to the overall Soviet Famine.

In addition to the studies of the Soviet Famine discussed earlier, these findings are related to several literatures in economics. The first are studies of the causes of famine, which have examined contexts such as China (e.g., Meng et al., 2015), India (e.g., Sen, 1981), and Ireland (e.g., 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 factor for, or dimension of, inequitable distribution.

Our findings help to better understand the consequences of Soviet economic policies, which constituted one of the largest political-economic experiments in the 20th century, if not of all human history. We complement macro calibrations of Soviet industrialization policies by Allen (2003) and Cheremukhin et al. (2017).⁹ In examining the famine, we are most closely related to Naumenko (forthcoming), which documents a positive association between collectivization and famine mortality in a cross-section of districts of the Ukrainian SSR.

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 in our paper are directly in line with the theoretical predictions for mass killings and genocide from Esteban et al. (2015). The fact

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

⁹Note that Cheremukhin et al. (2017) does not take into account human costs of Stalin's great leap forward.

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

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 In examining Ukrainian-Russian friction, our study is related to the recent study by Korovkin and Makarin (2019), which documents the differential response of ethnic-Ukrainian- and ethnic-Russian-led firms in Ukraine to the current conflict between Russia and Ukraine.

This paper is organized as follows. Section 2 summarizes the historical background. Section 3 presents the main province-level analysis. Section 4 presents the district-level analysis. Section 5 presents additional results for other ethnicities and the Tsarist famine. Section 6 concludes.

2 Background

This section provides a brief discussion of the chronology of the famine and the policies on the eve of the famine.

2.1 The Chronology of the Famine

The first news of possible famine began to circulate during the harvest of 1931. According to the official estimates, production was 17% lower than the previous year.¹² News of famine traveled to Moscow, but instead of relaxing the policies that were believed to have caused it, the government intensified them: it increased grain procurement targets by 20%, from 22.1 million tons in 1930 to 26.6 million in 1931 (Wheatcroft, 2001). In the meantime, starving peasants often consumed seed stock. The lack of seed stock and weakened labor force con-

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.

¹²Davies and Wheatcroft (2004) Table 1 reports the official 1930 harvest estimate to be 83.5 million tons, and the official 1931 harvest estimates to be 69.5 million tons.

tributed to lowering production in 1932, when procurement initially remained high. When the famine became apparent, procurement was slightly lowered. The famine was exacerbated by the huge drop in the livestock that occurred during the peak of forced collectivization in 1930, a traditional buffer saved and consumed by peasants in times of extremely low harvest. Deaths from starvation began to increase at the end of 1932 and peaked in the winter and spring of 1933. National mortality rates returned to trend in 1934, although some places took longer to recover. Thus, the famine is typically defined to occur from 1932 to 1933. The Soviet government did not acknowledge the famine until late 1980s.

2.2 Basic Facts

Officially reported total per capita grain output in 1931 and 1932 was 1.2 kilograms per person per day, 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 in urban areas.

The Soviet government aimed to centralize food production and distribution. To do so, in late 1929, it began the collectivization of agriculture. The goal was to remove private property and to move peasants into large collective farms which were believed to be more productive than small individual farms. Peasants did not want to give up their property for free and resisted collectivization. They 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). Wealthier, more productive peasants, or those actively resisting collectivization were persecuted as kulaks. As a part of dekulakization campaign, about two million peasants were exiled to Siberia and other remote areas, and about half a million perished (Viola, 2007).

Collective farmers worked in teams in the field. Food was produced and stored by the collective, and later delivered to state procurement officers. Procured food was to be distributed to urban industrial population or exported. In principle, this meant setting production and

procurement targets for each region, leaving peasants with enough for subsistence. In practice, food was procured even if peasants were left with below subsistence amount of food. There are many documents showing that Stalin advocated for over-procurement as a method to discipline the peasants, whom he believed to intentionally understate their production capacity (Danilov et al., eds, 1999-2006; Davies and Wheatcroft, 2004). In principle, food left after procurement was to be allocated between collective farm members proportionally to their effort. However, in practice, payments per labor day were often close to zero. This undermined individual incentives. Initially, the state also aimed to remove all private holdings, including small personal plots for potatoes in the peasants' backyards.

Collectivization could have contributed to famine through several channels: reducing grain production because of poor incentives or facilitating higher procurement because of more state control over the harvest. We will study these two channels in our paper. Collectivization may have also reduced the traditional buffer savings of food, such as production of potatoes and cattle-breeding, or 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 these latter channels.

2.3 Ethnic Bias and Ukrainians

The accusation that the Soviet government targeted Ukrainians above and beyond Russians is driven by several sources. First, aggregate mortality rates in Ukraine were much higher than in Russia. Second, Ukraine, one of the most agriculturally productive regions, was among 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 form of nationalism undermined the Bolshevik ideal of building socialism in one country

which had been an official goal since 1925. Concerns about national opposition to the regime became stronger in the time of troubles.¹³

Historians offer little discussion of ethnic bias in early Soviet policies. Many historians note that there is no explicit evidence that Stalin “ordered” the famine.¹⁴ But there are examples of ethnic tensions and allegiances, as well as evidence that Stalin was well-aware of these and utilized them for central governance. For example, he was known to widely recruit members of ethnic minorities into his secret police (Gregory, 2009). In the context of the famine, a good example 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 intensified, they began asking Stalin to reduce procurement. In 1932, de-classified secret police reports to Stalin indicated the resignations of ethnic Ukrainian Party leaders at all levels explaining 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, neither of whom were ethnic Ukrainian, to implement his policies in the Republic of Ukraine and the North Caucasus, two key grain producing regions where the bulk of ethnic Ukrainians lived. We will later explore the extent to which such ethnic allegiances played a role.

¹³A 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, 1999-2006 Vol. 3, Document 226).

¹⁴E.g. Kotkin (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.

3 Main Results

3.1 Data

The data we use come from archival and published sources, Appendix C gives the exact source of every variable.¹⁵ We construct two main datasets: a more aggregated but long and detailed panel of provinces and a much less aggregated but short and less detailed panel of districts. Here we describe the panel of provinces; we discuss the district data later when they become relevant.

At the province level, we collect data on mortality, grain production and procurement, collectivization, tractors, livestock, weather, indicators of political development and other local characteristics. From 1922 to 1940, there were ongoing administrative reforms that led to multiple changes in administrative borders. We convert all our variables to 1932 provincial units because grain procurement figures are reported for these administrative units (in particular, the Soviet Republic of Ukraine stands for one province in our panel). We construct a panel of nineteen provinces spanning the republics of Belarus, Ukraine, and most of Russia, and the years of 1922 to 1940. Altogether, our provinces cover 84% of the 1926 Soviet population, and 88% of the 1928 Soviet grain production; Figure 1a maps the provinces in our sample, omitted regions are in white.¹⁶ An average province in our panel has 6.5 million people in 1926.

Our main variable of interest, population ethnic composition, comes from the 1926 Soviet census, which is commonly viewed as one of the best Soviet censuses and a reliable source

¹⁵We 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.

¹⁶The only Russian regions not covered by our panel are Far East, Yakutia, and the republics of North Caucasus. They were small in terms of population and grain output share. For these regions, and for the Soviet territories outside of Belarus, Russia, and Ukraine, there are no reliable mortality data until mid-1930s. In particular, our data do not include Kazakhstan with total estimated famine deaths between 1 and 2 million (Kondrashin, 2008).

of information (Andreev et al., 1998). We also use data on the distribution of population by mother tongue from the 1897 and 1926 population censuses. Note that the 1897 census is the only census conducted before the establishment of the communist regime in 1917. Appendix Table A.1 lists ethnic groups most often mentioned in the famine literature. Russians were the ethnic majority and constituted 53% of the 1926 Soviet population; Ukrainians were by far the largest ethnic minority, they constituted 21% of the Soviet population. The second-largest ethnic minority, Belorussians, were just 3% of the population. According to the 1926 population census, in Ukraine, ethnic Ukrainians comprised 23.2 out of 29.2 million citizens, and an additional 5.6 million ethnic Ukrainians lived outside of Ukraine. Appendix Figure A.1 shows the distribution of ethnic Russians, Ukrainians and of other ethnic groups on the map.

Figure 2a plots mortality rates, that is, the number of deaths divided by population, in our sample from 1900 to 1990.¹⁷ Figure 2b zooms in at the period from 1923 to 1940. Mortality declines over time, with spikes in 1910, 1932–1933 and 1947. These spikes correspond to the 1910 cholera epidemic, and the 1932–1933 and 1947 famines (Ellman, 2000; Davis, 2018). The largest spike in mortality is the 1932–1933 famine that we study. Average famine mortality rates in our sample are almost 60 per 1,000. Already mentioned Figure 1a maps excess mortality in 1933, the year with the highest famine mortality, for the provinces in our sample.¹⁸ There is a significant geographic variation: Ukraine and the southern provinces of Russia appear to suffer much more than the northern or eastern provinces.

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 are not known to have ever been systematically manipulated by the central government.

¹⁷No reliable mortality data are available for the periods of 1916–1922 (World War I, the Civil War, and the 1921–1922 famine) and 1941–45 (World War II).

¹⁸We calculate 1933 excess mortality as mortality in 1933 minus mortality in 1928. Note that we only use this calculation for illustration.

The government used various cross-check procedures to avoid manipulation at lower levels of government. One important exception is aggregate grain production. Grain production was viewed as one of the key and public indicators for Soviet economic health — i.e., the success of the new Bolshevik regime. As such, it was controversial and there exists much debate about whether reported aggregate grain production was higher or lower than actual production (e.g., Wheatcroft and Davies, 1994, 2004). For our study, it is important to note that all of the discussions of falsification have focused on aggregate (i.e., national) production and there have been no claims that misreporting was correlated with ethnic composition at disaggregated levels of government. Nevertheless, our analysis will carefully keep the possibility of mis-reporting in mind. Reported grain production will not be central to our analysis. When we use it as a control variable, we will also show that the results are similar if we control for grain production that is predicted by natural conditions.

3.2 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 previous 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 rate 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, X_{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.¹⁹ Our baseline defines the famine dummy to take a value of one in 1932, because 1933 was the year with the highest mortality rates when the famine became apparent in all regions. 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 that in a province comprised of 100% ethnic Ukrainians, famine mortality rates would have been higher than in 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 0.013 and the standard deviation in Ukrainian population share is 0.216. Thus, during the famine, increasing Ukrainian population share by one standard deviation would result in a 0.828 standard deviation increase in mortality. This is a large effect.

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

To observe the timing of differential Ukrainian mortality, we estimate an equation similar to the baseline, except that we interact Ukrainian population share with dummy variables for all years instead of only 1932. Each interaction coefficient with year t reflects the mortality difference in year $t + 1$ between regions with higher Ukrainian population share and regions with lower Ukrainian population share (relative to the mortality difference in the reference year, 1923). Figure 3a shows a striking pattern. Prior to the famine, from 1924 to 1931, there was no large difference in mortality rates across regions. However, regions with a

¹⁹The baseline uses a time-varying urbanization variable 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.

higher share of Ukrainians began to experience higher mortality in 1932 (approximately 10 more deaths per 1,000 individuals) and this difference peaked in 1933 (approximately 50 more deaths per 1,000 individuals). This pattern is consistent with historical evidence that there was a small famine after the harvest of 1931, which was greatly exacerbated after the harvest of 1932. Afterwards, from 1934 to 1940, regions with higher share of Ukrainians had mortality rates similar to 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, equal to one in 1931 and 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 uses Ukrainian share in the rural population because the famine was driven by agricultural policies targeted at the rural population. Columns (3), (4), (5) and (6) show that our results, in terms of standard deviations, are nearly identical if we alternatively use the urban share or total share of Ukrainians, or the share of people whose mother tongue is Ukrainian according to the 1926 or 1897 Population Censuses. This is not surprising since the five measures of Ukrainians population share — rural, total, urban, by mother tongue in 1926, by mother tongue in 1897 — 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 smaller than that in the rural areas or the province total, making the estimated level effect on mortality larger. Henceforth, we will use the 1926 rural Ukrainian population share as the explanatory variable.

In column (7), we exclude the the Ukrainian SSR, where 78% of all Ukrainians in our

sample reside. The Ukrainian interaction coefficient is 0.072 and statistically significant at the 1% level, although, the standardized effect is slightly smaller, 0.704, than in the full sample, 0.828. This means that the baseline results are not driven by a comparison of differences between the Ukrainian SSR and other republics. Instead, they reflect systematic higher mortality rates for regions with a higher share of *ethnic* Ukrainians.

In column (8), we control for the population gender ratio and the share of individuals aged ten and younger (as reported by the 1926 population census), each interacted with the famine indicator. These controls are motivated by the observation that young children were more likely to perish during the famine, and the possibility that men and women may have experienced different famine mortality. The Ukrainian interaction coefficient is 0.050 and significant at the 1% level. Thus, higher famine mortality in areas with more ethnic Ukrainians is not driven by the difference in the demographic composition between Ukrainian-populated regions and other regions.²⁰

In column (9), we control for officially reported per capita grain production and its interaction with the famine indicator. The Ukrainian interaction coefficient is 0.047, nearly identical to the baseline, and statistically significant at the 1% level. This means that even after controlling for how much grain was produced in 1932, Ukrainians still died more during the famine. This implies that the Soviet government procured a higher share from regions with a higher share of Ukrainians. Later, we will examine procurement directly.

3.3 Weather

Next, we examine the role of natural conditions — weather — in causing high Ukrainian famine mortality. We use monthly temperature and precipitation data from Matsuura and Willmott (2014) together with province-level grain production for years prior to the estab-

²⁰Our results are also robust to a large number of other demographic controls: e.g., share of the elderly, age-by-gender controls, etc. The estimates are available upon request.

lishment of the Soviet Union, 1901 to 1915, to predict weather-driven production during our sample of interest.²¹ We control for predicted grain and its interaction with the famine dummy in column (10). The estimates for Ukrainians do not change with these additional controls. Thus, higher mortality in the regions with a larger share of ethnic Ukrainians cannot be explained by abnormal weather in these regions in 1932.²² Moreover, in column (11) we control for province latitude interacted with the famine indicator, longitude interacted with the famine indicator, and latitude interacted with longitude and with the famine indicator. The Ukrainian coefficient is 0.045 and significant at the 1% level. Taken literally, this means that for two provinces located at the same latitude and longitude, the one with more ethnic Ukrainians experienced significantly higher famine mortality.

3.4 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 because of the widespread belief that Soviet agricultural policy targeted the most productive areas. 1928 was the beginning of the first Five Year Plan and therefore a common reference for Soviet planning (e.g. see Wheatcroft, 2001). Column (3) further splits 1928 grain production into wheat, rye, and other grain since they may have varied in importance for state procurement. Wheat was the most domestically and internationally traded staple crop. Rye, on the other hand, was the main grain that peasants sowed for self consumption. Column (4) examines proxies for the population's commitment and/or loyalty to the Bolsheviks with vote shares from the 1917 Constituency Assembly election. This was 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

²¹See the Appendix A for more discussion of the prediction.

²²Appendix Figure A.4 plots demeaned average seasonal temperature and precipitation over time and shows that weather in an average province during the famine was similar to other years, when there were no famines.

population turned out.²³ Column (5) controls for the urban and rural 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. Urban and rural communists were both parts of the state bureaucracy, but played very different roles during the famine. We will discuss this more in the next section. Column (6) controls for the number of Party secretaries (at the province, district, city and, if the city was large, the borough level) who attended the 1930 Party Congress to vote formally for the policy of comprehensive 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 (7), we control for grain production in 1928 and all of the political factors together, each interacted with the famine dummy variable. Our main interaction coefficient for Ukrainians is very robust and always similar to the baseline in column (1). These results mean that higher mortality in areas with higher share of ethnic Ukrainians is not due to a coincidence of the presence of political factors or higher 1928 grain production and the Ukrainian population.

We note that all interactions with the one exception of 1928 wheat production are statistically insignificant. The positive and statistically significant wheat coefficient fit the historical narrative that Soviet agricultural policies targeted wheat-producing areas more, which led to higher procurement and correspondingly higher mortality. We postpone the discussion of the other political variables to the section on heterogeneous results, where we demonstrate that these average effects mask meaningful heterogeneity.

²³We follow Castañeda Dower and Markevich (2020) and use disaggregated district-level data on votes for the Bolsheviks from Protasov et al. (2014). See the Data Appendix for details.

3.5 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 3.4 for the Ukrainian population share. Table 3 estimates the fully saturated triple interaction specification of these policy variables on mortality. The double interactions capture the effects of the policy proxy on excess famine mortality rates in a province with no Ukrainians in rural areas. The triple interaction captures the effect of the policy on excess famine mortality rates in a province that is 100% Ukrainian relative to a province with no Ukrainians (in rural areas). The interaction of Ukrainian population share and the famine dummy variables is not meaningful in this table since it captures excess mortality rates for Ukrainians in provinces where the political variables of interest have 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 political variables.

Column (1) shows that regions which were agriculturally productive in 1928 suffered higher mortality rates during the famine, 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.

Columns (2) conducts a similar analysis with Bolshevik vote share in the 1917 election. In places with no Ukrainians, higher Bolshevik vote share is negatively correlated with 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. If Stalin viewed Ukrainians as disloyal and Bolshevik vote share proxies for Stalinist administrative capacity, these results are consistent with the Stalinist practice of rewarding

loyal populations and penalizing disloyal ethnicities, which was openly realized in the late 1930s and mid 1940s (Polyan, 2001).

Column (3) 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 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 population and mitigating the consequences of famine when they spilled over to urban areas.²⁴ Rural communists increase mortality only in areas with many Ukrainians. Similarly, urban communists moderate 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 (4) examines the effect of the presence of Party Secretaries, who participated in the 1930 Party Congress and voted for collectivization, in a province. The triple coefficient is positive and statistically significant at the 1% level, and the double coefficient is close to zero and 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 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.²⁵ This allows us to distinguish the effects of having delegates who were Ukrainians themselves versus having delegates who were of other ethnicities. In column (5), we add the triple interactions of each of the two new

²⁴For example, many famine refugees went to cities to beg for food. Urban government also set up orphanages for abandoned children (Davies and Wheatcroft, 2004; Kondrashin, 2008).

²⁵See the Data Appendix.

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 for non-Ukrainian delegates are larger in magnitude than for Ukrainian delegates. The difference between the two coefficients is almost statistically significant at the 10% level (p-value for equality of the coefficients is 0.13). We interpret these results as suggestive evidence supporting the idea that there were ethnic differences in how Party leaders, who had ostensibly similar ideologic commitment prior to the famine, responded to the crisis, with non-Ukrainians more zealously implementing the policies that led to Ukrainian mortality and ethnic Ukrainian leaders relenting more to the reality of famine. These results are also consistent with the view that Stalin often sent “outsiders” to govern to counteract local loyalties and nationalism.²⁶

Figure 4 shows the timing of the effects of these policies. 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.

3.6 Soviet Agricultural Policies

To dig deeper into the mechanism causing ethnic bias in famine mortality, we explore heterogeneity treatment effects for the implementation of Soviet economic policies, which affected agricultural production and the amount of food retained in the countryside. For that, we examine collectivization, tractors supply, grain production and procurement share as dependent variables. The triple interaction estimation is similar to our earlier analysis of mortality. We estimate the same set of specifications as before with one exception. For brevity, we go directly to the specification that distinguishes ethnic Ukrainian and non-Ukrainian delegates. As before, we are mostly interested in the double and triple interactions of the political variables we have examined earlier; the interaction of Ukrainian population share and the famine

²⁶See the Background Section.

dummy is not meaningful in these additional specifications since it captures excess mortality rates in Ukrainian provinces, where the political variable of interest has a value of zero.

Table 4 presents heterogeneous treatment estimates for collectivization (the share of households in collective farms, columns 1 to 4) and the amount of tractor horsepower per capita (columns 5 to 8), as dependent variables. Collectivization was a major policy tool to reform the production of agricultural goods in the USSR. It was meant to boost food production and government food procurement. One of the main methods for boosting production was increased mechanization. Thus, we examine tractor horsepower as an indicator for mechanization. In columns (1)-(4), the estimates for collectivization vary slightly in precision, but the signs of the estimates mirror those for mortality, which is consistent with collectivization being a key contributor to famine mortality. Similarly, in columns (5)-(8), the signs are opposite of the estimates for mortality and collectivization. This implies that the same factors that caused Ukrainian-regions to suffer higher mortality and more intense collectivization, also caused these regions to receive less mechanization.²⁷

In Table 5, the estimates present the examine per capita grain production and grain procurement share (the share of the harvest procured by the government) as dependent variables. Given concerns in reported grain data, we interpret these results very cautiously as suggestive evidence. The triple interaction effects for grain production are mostly statistically insignificant, which suggests that differential grain production does not play an obvious role in explaining excess Ukrainian mortality. The triple interaction effects for procurement share are more precisely estimated. The signs are similar to the signs for the triple interaction estimates for collectivization and mortality shown earlier. Thus, this suggestive evidence is consistent

²⁷ An alternative explanation of mechanization results is that the Soviet government allocated tractors in such a way that would compensate for a drop in the number of horses caused by collectivization policy (which had ethnic bias). We explore this possibility in the Appendix Table A.4. In Section A, we replicate the specifications from Table 4 Section B with the lagged number of work horses as an additional control. With a single exception, the coefficients on interactions of political factors with the share of Ukrainians and the famine dummy lose their statistical significance but all of them keep negative sign. When we use the number work horses as a dependent variable (Section B), we get mostly imprecise but always negative estimates of coefficients on triple interactions.

with the notion that excess procurement played an important role in causing famine mortality in Ukrainian regions.

4 District-Level Analysis

The district-level panel comprises of two years: 1928 and 1933. All data are manually collected from the former Soviet archives. See the Data Appendix for more details.

There is substantial variation in famine mortality across districts, even those within the same province. Figure 1b shows a map of excess mortality in 1933 for each district where data are available. We define excess mortality as the difference between 1933 and 1928 mortality rates. The increased granularity allows us to provide several additional pieces of evidence.

First, these data allow us to examine the claim that there was a strong border effect and that the famine was notably more severe on the Ukrainian SSR side of the border between Russia and Ukraine.²⁸ Figure 5a plots 1933 excess mortality against the distance to the border between Russia and Ukraine. 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 ethnic Ukrainians. This can be seen in Figure 5b, 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 6 column (1) first replicates the baseline specification from the province-level estimate.²⁹ 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

²⁸The government introduced a ban on migration from the Republic of Ukraine and from the North Caucasus region in January 1933 (Danilov et al., eds, 1999-2006, Vol. 3).

²⁹Note that we use urbanization from 1926 and 1933 because urbanization is not available for 1928.

increase of the local share of ethnic Ukrainians results in a 0.543 standard deviation increase in mortality.

Column (2) includes province-year fixed 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 region leaders, 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 to the railroad and to large 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 or other regions to escape the famine. Similarly, distance to large cities in principle could have opposing effects. Soviet policy aimed to secure the grain supply for industrial populations, which 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.

We gradually introduce the triple interactions of distance to the nearest railway station and distance to the nearest large city (a city with population of 20,000 or more) with Ukrainian population share and the famine dummy variable in columns (3) and (4). In column (5), 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 to 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 the 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 replicate most 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.

5 Additional Results

In this section, we return to using the province-level panel to explore potential ethnic bias in famine mortality with respect to other ethnicities as well as in the context of the last documented large-scale famine of the Tsarist regime.

5.1 Other Ethnic Minorities

Conventional wisdom about the Soviet Famine postulates that Belorussians (3.2%), Kazakhs (2.7%), Germans (0.8%) and Poles (0.5%) suffered higher mortality than Russians, while ethnic Jews (1.8%), who mostly resided in urban areas, suffered lower mortality.³⁰ Table 7 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. The omitted ethnic group is Russians. Regions with higher shares of Ukrainians, Belorussians and Germans suffer higher famine mortality relative to ethnic Russians. The magnitude is the largest for the interaction with ethnic German share, although very imprecisely estimated. The areas with other ethnic groups suffered similar mortality as territories with ethnic Russians.

Column (2) controls for weather-induced grain production and its interaction with the

³⁰Appendix Table A.1 lists the ethnicities most often mentioned in the famine literature.

famine dummy. Column (3) controls for the proxies for political factors and economic policies that we used in our earlier analysis, each interacted with the famine dummy variable. Column (4) includes the two sets of controls from columns (2) and (3) simultaneously. The estimates for Ukrainians and Belorussians 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 Germans remains positive but becomes imprecisely estimated, while the coefficient for Kazakhs becomes positive and highly statistically significant and the coefficient for Poles becomes large, negative and statistically significant as we add controls.³¹

In the Appendix, we examine the mechanisms driving mortality rates in regions with a high share of Belorussian, Germans and Kazakhs. The estimates suggest that the policy mechanisms may be 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.³² It is beyond the scope of this paper to delve deeper into the Belorussian or German famine experience. This is an important topic for future study.

5.2 The 1892 Famine

To understand whether ethnic bias in Ukrainian famine mortality is an outcome of Soviet policy, or whether some ethnic groups always suffer higher mortality during famines (e.g., because of differences in social capital, networks or culture), we examine mortality rates during the Tsarist famine of 1892 using province-level mortality data from 1885 to 1913.³³ Table 8 replicates the baseline as closely as possible with the 1885–1913 data. For comparison, we present results in pairs for the Tsarist famine (Columns 1 to 4) and the Soviet famine

³¹Note that the Polish population was very small and resided almost entirely in two provinces.

³²The government did not record mortality for Kazakhstan until 1935.

³³Volha Charnysh kindly shared 1885–1896 mortality data with us, Charnysh and McElroy (2020).

(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 we studied in the previous subsection (Ukrainians, Belorussians, Kazakhs, Germans, Poles, Jews and others, omitting Russians for comparison) by including a 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 variable of interest, Ukrainian population share interacted with the famine dummy, is statistically zero in all specifications for the Tsarist famine. In contrast, it is positive, large in magnitude and statistically significant at the 1% for the Soviet Famine. This implies that our main results are unlikely to be driven by time-invariant characteristics of ethnic Ukrainians. Instead, the results are consistent with the view that during the Soviet famine higher mortality in Ukrainian regions was a result of Soviet-specific policies, and the earlier evidence that Ukrainians experienced policies which led to famine, such as collectivization and procurement, more intensely.

5.3 Robustness to Historical Institutions and Cultural Differences

Another way to investigate whether our main results on excess Ukrainian mortality is driven by differences in cultural practices or historical institution is to directly control for these factors. One important historical institution in this context is the repartition commune. Living in one required cooperative behavior, and these communes were less widespread among Ukrainians than among Russians. If the values of cooperation were transmitted intergenerationally, this difference could contribute to the difference in mortality between the two ethnicities. We collect data on the share of households in repartition communes from the 1905 land census. In addition, we also collect data on other potentially important variables such as the shares of Catholics and Orthodox Christians (the two major religion groups in Ukraine) from the 1897 census, the share of peasant land and the land Gini estimated from the 1905 land census. Appendix Table A.7 shows that our results are robust if we add interactions of these variables with the famine dummy into the baseline specification.

6 Conclusion

The Soviet Great Famine has been one of the largest and most controversial economic disasters in recent history. Within just two years, between 5.5 and 10.8 million people died throughout the nation and the ethnic Ukrainian population, the second largest ethnic group in the Soviet Union, had been decimated. Without systematic data to rule out confounding factors or direct documentary evidence on the intent 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 suggest 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, Germans 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, or, more generally, the Soviet Union and post-Soviet Eastern Europe. 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 research in this 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.
- 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.

- Castañeda Dower, Paul and Andrei Markevich**, “Democratic support for the Bolshevik Revolution: An empirical investigation of 1917 Constituent Assembly elections,” *Working paper*, 2020.
- 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 Sovetskoy 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 Industrialization 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.

- Kondrashin, Viktor**, *Golod 1932-1933 godov: Tragediya Sovetskoy derevni*, Moscow: Rosspen, 2008.
- Korovkin, Vasily and Alexey Makarin**, “Conflict and Inter-Group Trade: Evidence from the 2014 Russia-Ukraine Crisis,” *Working paper*, 2019.
- 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.*, Routledge, 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.
- Meng, Xin, Nancy Qian, and Pierre Yared**, “The Institutional Causes of China’s Great Famine, 1959–1961,” *The Review of Economic Studies*, 2015, 82 (4), 1568–1611.
- 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.
- Naumenko, Natalya**, “The Political Economy of Famine: the Ukrainian Famine of 1933,” *Journal of Economic History*, forthcoming.
- 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.
- Padró i Miquel, Gerard**, “The Control of Politicians in Divided Societies: The Politics of Fear,” *The Review of Economic Studies*, 10 2007, 74 (4), 1259–1274.
- Polyakov, Yu.A. and V.B. Zhiromskaya, eds**, *Naselenie Rossii v XX veke: istoricheskie ocherky. 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**, *Vserossiiskoe uchreditelnoe sobranie. Entsiklopedia. (All-Russian Constitutional Assembly. Encyclopedia)*, Moscow: Rosspen, 2014.
- Rosefelde, Steven**, *Red Holocaust*, London and New York: Routledge, 2010.
- Rudnytskyi, Omelian, Natalia 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.
- 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										
	Famine =			Mother		Mother Tongue		Omit		Reported	
	Baseline	1931, 32	Total Ukrainians	Urban Ukrainians	1926	1897	Ukrainian	SSR	Ukraine	Grain	Predicted Grain
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ukrainians x Famine	0.051*** [0.007]	0.031*** [0.011]	0.056*** [0.007]	0.093*** [0.013]	0.0591*** [0.0104]	0.0592*** [0.0086]	0.072*** [0.006]	0.050*** [0.007]	0.047*** [0.004]	0.051*** [0.008]	0.045*** [0.008]
<i>Normalized</i>											
Urbanization	-0.010** [0.004]	-0.011*** [0.004]	-0.010** [0.004]	-0.010** [0.004]	-0.0100** [0.0040]	-0.0097** [0.0040]	-0.010** [0.004]	-0.010** [0.004]	-0.008** [0.004]	-0.010** [0.004]	-0.010** [0.004]
Urbanization x Famine	0.004 [0.009]	0.010 [0.006]	0.005 [0.009]	0.003 [0.009]	0.0073 [0.0092]	0.0035 [0.0086]	0.003 [0.009]	-0.009 [0.013]	0.011 [0.015]	0.004 [0.009]	0.008 [0.012]
Age structure, gender ratio x Famine	N	N	N	N	N	N	N	Y	N	N	N
Grain, x Famine	N	N	N	N	N	N	N	N	Y	Y	N
Latitude, Longitude x Famine	N	N	N	N	N	N	N	N	N	N	Y
Observations	337	337	337	337	337	337	319	337	337	337	337
R-squared	0.781	0.735	0.781	0.775	0.767	0.786	0.759	0.787	0.802	0.782	0.790
Mortality 1933											
Mean	0.031	0.031	0.031	0.031	0.031	0.031	0.029	0.031	0.031	0.031	0.031
Std. Dev.	0.013	0.013	0.013	0.013	0.013	0.013	0.012	0.013	0.013	0.013	0.013
Ukrainians											
Mean	0.104	0.104	0.095	0.055	0.074	0.085	0.061	0.104	0.104	0.104	0.104
Std. Dev.	0.216	0.216	0.197	0.116	0.178	0.188	0.113	0.216	0.216	0.216	0.216

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, 7-11), total population (Column 3), urban population (Column 4), or the share of people whose mother tongue is Ukrainian (Column 5) according to the 1926 Census; in Column 6, Ukrainians is the share of people whose mother tongue is Ukrainian according to the 1897 Census. Famine is an indicator that equals one in 1932 (Columns 1, 3-11) or in 1931 and 1932 (Column 2), and zero otherwise. Estimates in Column 8 control for Reported grain and Reported grain x Famine. Estimates in Column 9 control for Predicted grain and Predicted grain x Famine. Reported and predicted grain are measured in kilograms per person per day. Estimates in Column 10 control for the province centroid Latitude x Famine, Longitude x Famine, and Latitude x Longitude x Famine. Estimates in Column 11 control for the share of people aged 10 and younger x Famine, and for the male/female ratio x Famine (according to the 1926 Census). All estimates control for province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Normalized coefficients express the effect of increasing Ukrainians 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)
Ukrainians x Famine	0.051*** [0.007]	0.049*** [0.005]	0.059*** [0.009]	0.049*** [0.007]	0.048*** [0.005]	0.050*** [0.006]	0.048*** [0.004]
Grain 1928 x Famine		0.006 [0.005]					0.003 [0.004]
Wheat 1928 x Famine			0.012* [0.007]				
Rye 1928 x Famine			0.023 [0.015]				
Other grain 1928 x Famine			-0.021 [0.017]				
Bolshevik votes 1917 x Famine				-0.012 [0.008]			-0.001 [0.011]
Urban communists x Famine					-0.001 [0.001]		0.000 [0.001]
Rural communists x Famine					0.002 [0.002]		0.002 [0.001]
Voting delegates 1930 x Famine						0.032 [0.023]	0.022 [0.016]
Observations	337	337	337	337	337	337	337
R-squared	0.781	0.796	0.813	0.785	0.795	0.789	0.801

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 Census. Famine is an indicator that equals one in 1932 and zero otherwise. Grain 1928 is the 1928 grain harvest; Wheat 1928 is the 1928 wheat harvest; Rye 1928 is the 1928 rye harvest; Other grain 1928 is the 1928 harvest of all grains except wheat and rye. Grain, Wheat, Rye, and Other grain are measured in kilograms per person per day. Bolshevik votes 1917 is the share of votes for Bolsheviks 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: Heterogeneous Effects on Mortality

	Dependent variable: Mortality in year t+1				
	(1)	(2)	(3)	(4)	(5)
Ukrainians x Famine	-0.172*** [0.040]	0.001 [0.041]	0.062 [0.070]	-0.040* [0.022]	-0.035 [0.032]
Grain 1928 x Famine	0.000 [0.002]				
Ukrainians x Grain 1928 x Famine	0.297*** [0.062]				
Bolshevik votes 1917 x Famine		-0.023* [0.012]			
Ukrainians x Bolshevik votes 1917 x Famine		1.034* [0.551]			
Urban communists x Famine			0.000 [0.001]		
Ukrainians x Urban communists x Famine			-0.132*** [0.046]		
Rural communists x Famine			0.000 [0.001]		
Ukrainians x Rural communists x Famine			0.107*** [0.025]		
Voting delegates 1930 x Famine				-0.005 [0.011]	
Ukrainians x Voting delegates 1930 x Famine				2.393*** [0.397]	
Voting Ukrainian delegates 1930 x Famine					0.332 [0.202]
Ukrainians x Voting Ukrainian delegates 1930 x Famine					1.797*** [0.517]
Voting non-Ukrainian delegates 1930 x Famine					-0.024 [0.020]
Ukrainians x Voting non-Ukrainian delegates 1930 x Famine					2.225*** [0.329]
Observations	337	337	337	337	337
R-squared	0.848	0.818	0.844	0.848	0.852

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 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. 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 4: Heterogeneous Effects on Collectivization and Mechanization

	Dependent variable:							
	A. Collectivization				B. Tractors' horse power			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	-0.792 [0.598]	0.070 [0.416]	1.156 [1.019]	0.572 [0.712]	0.132** [0.061]	-0.024 [0.055]	-0.077 [0.087]	0.045 [0.038]
Grain 1928 x Famine	0.062 [0.051]				-0.003 [0.003]			
Ukrainians x Grain 1928 x Famine	1.329* [0.679]				-0.265** [0.110]			
Bolshevik votes 1917 x Famine		-0.487*** [0.084]				0.020* [0.011]		
Ukrainians x Bolshevik votes 1917 x Famine		7.847*** [2.947]				-0.969 [0.609]		
Urban communists x Famine			-0.001 [0.005]				0.001 [0.001]	
Ukrainians x Urban communists x Famine			-0.196 [0.499]				0.114 [0.073]	
Rural communists x Famine			0.052*** [0.017]				-0.001 [0.001]	
Ukrainians x Rural communists x Famine			0.306 [0.204]				-0.088** [0.043]	
Voting Ukrainian delegates 1930 x Famine				0.275 [2.732]				-0.199 [0.225]
Ukrainians x Voting Ukrainian delegates 1930 x Famine				7.529 [8.183]				-2.125** [0.935]
Voting non-Ukrainian delegates 1930 x Famine				-0.413 [0.281]				0.008 [0.020]
Ukrainians x Voting non-Ukrainian delegates 1930 x Famine				11.864** [4.849]				-2.216*** [0.822]
Observations	228	228	228	228	247	247	247	247
R-squared	0.968	0.969	0.969	0.968	0.790	0.787	0.790	0.790

Notes: Observations are at the province and year level. Collectivization is the share of rural households in collective farms. Tractors' horse power is the reported tractors' horse power per capita. 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. 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 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 5: Heterogeneous Effects on Grain Production and Procurement

	Dependent Variable:							
	A. Grain				B. Procurement share			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	-0.835 [1.688]	-0.373 [1.257]	-0.492 [2.737]	1.614 [1.915]	-0.241 [0.321]	0.410 [0.292]	1.790** [0.754]	0.373 [0.500]
Grain 1928 x Famine	-0.224** [0.101]				0.063*** [0.023]			
Ukrainians x Grain 1928 x Famine	1.033 [2.334]				1.056** [0.469]			
Bolshevik votes 1917 x Famine		0.065 [0.275]				-0.164* [0.090]		
Ukrainians x Bolshevik votes 1917 x Famine		3.556 [8.201]				5.526** [2.320]		
Urban communists x Famine			0.062** [0.029]				-0.008 [0.008]	
Ukrainians x Urban communists x Famine			-0.489 [1.539]				-0.016 [0.307]	
Rural communists x Famine			-0.053 [0.051]				0.016 [0.014]	
Ukrainians x Rural communists x Famine			0.524 [0.750]				0.330** [0.142]	
Voting Ukrainian delegates 1930 x Famine				11.454* [6.631]				-3.189** [1.388]
Ukrainians x Voting Ukrainian delegates 1930 x Famine				-30.508 [24.539]				15.296*** [5.356]
Voting non-Ukrainian delegates 1930 x Famine				-1.780** [0.765]				0.253* [0.148]
Ukrainians x Voting non-Ukrainian delegates 1930 x Famine				-6.729 [16.584]				12.329*** [3.375]
Observations	361	361	361	361	186	186	186	186
R-squared	0.763	0.762	0.763	0.763	0.882	0.876	0.880	0.880

Notes: Observations are at the province and year level. Grain is the grain harvest measured in kilograms per person per day. Procurement share is the share of harvest procured by the government. 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. 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 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 6: Ethnic Ukrainians and Famine Mortality, District-level Estimates

	Dependent variable: Mortality				
	(1)	(2)	(3)	(4)	(5)
Ukrainians x Famine	0.048*** [0.002]	0.039*** [0.005]	0.046*** [0.006]	0.083*** [0.011]	0.079*** [0.012]
Log distance to a railroad + 1 x Famine			0.003 [0.003]		0.004 [0.003]
Ukrainians x Log distance to a railroad + 1 x Famine			-0.005 [0.017]		-0.002 [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				-0.010*** [0.002]	-0.008*** [0.003]
District FE	Y	Y	Y	Y	Y
Year FE	Y	N	N	N	N
Province-Year FE	N	Y	Y	Y	Y
Observations	3,513	3,513	3,206	3,513	3,206
R-squared	0.727	0.763	0.784	0.777	0.788
Provinces	18	18	18	18	18

Notes: Observations are at the district and year level, the sample has two years: 1928 and 1933. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Census. Famine is an indicator that equals one in 1933 and zero in 1928. Log distance to a railroad + 1 is the logarithm of the distance from the district's administrative center to the nearest railroad station. Log distance to a 20K city is the logarithm of the distance from the district's centroid to the nearest city with at least 20,000 inhabitants according to the 1926 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 7: Other Ethnic Groups and Famine Mortality

	Dependent variable: Mortality in year t+1			
	(1)	(2)	(3)	(4)
Ukrainians x Famine	0.067*** [0.005]	0.101*** [0.004]	0.053** [0.022]	0.106*** [0.005]
Belorussians x Famine	0.022*** [0.006]	0.048*** [0.004]	0.024* [0.014]	0.055*** [0.005]
Jews x Famine	-2.964* [1.708]	-0.447 [0.830]	-3.907 [2.566]	-0.080 [1.016]
Germans x Famine	0.307* [0.181]	0.104 [0.071]	0.258 [0.194]	0.047 [0.071]
Poles x Famine	1.114 [1.521]	-2.170*** [0.815]	1.447 [2.308]	-2.790*** [0.913]
Kazakhs x Famine	-0.511 [1.574]	1.946*** [0.599]	-0.252 [1.715]	2.351*** [0.620]
Other x Famine	-0.004 [0.006]	0.005 [0.003]	-0.029 [0.024]	0.009 [0.007]
Predicted Grain, x Famine	N	Y	N	Y
Political Variables x Famine	N	N	Y	Y
Observations	337	337	337	337
R-squared	0.843	0.867	0.851	0.868

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 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 Census — the omitted category is Russians. 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. Estimates in Columns (2) and (4) control for Predicted grain and Predicted grain x Famine. Estimates in Columns (3) and (4) control for Political variables x Famine: Grain 1928 x Famine, Bolshevik 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. 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. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

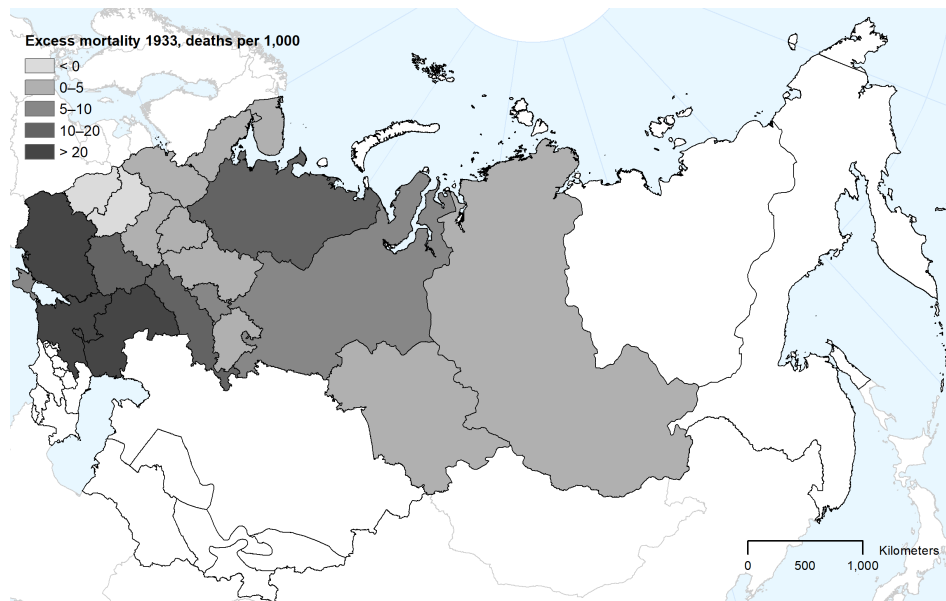
Table 8: Ethnic Composition and the 1892 Famine Mortality

	Dependent variable: Mortality in year t+1							
	1892				1933			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	0.002 [0.003]	0.002 [0.003]	-0.001 [0.003]	-0.003 [0.003]	0.067*** [0.005]	0.071*** [0.005]	0.067*** [0.005]	0.067*** [0.004]
Belorussians x Famine	0.001 [0.004]	0.002 [0.005]	-0.001 [0.004]	-0.001 [0.003]	0.022*** [0.006]	0.025*** [0.006]	0.022*** [0.007]	0.022*** [0.006]
Jews x Famine	-0.081** [0.040]	-0.097* [0.058]	-0.069* [0.036]	-0.081** [0.037]	-2.964* [1.708]	-2.980* [1.764]	-3.000* [1.782]	-3.319 [2.130]
Germans x Famine	0.104* [0.053]	0.104* [0.059]	0.075* [0.045]	0.051 [0.048]	0.307* [0.181]	0.323 [0.206]	0.310 [0.188]	0.339 [0.220]
Poles x Famine	-0.005 [0.039]	0.014 [0.066]	-0.011 [0.033]	0.021 [0.035]	1.114 [1.521]	0.977 [1.466]	1.148 [1.592]	1.461 [1.943]
Kazakhs x Famine	0.040*** [0.005]	0.046*** [0.014]	0.032*** [0.005]	0.020*** [0.008]	-0.511 [1.574]	-0.446 [1.504]	-0.531 [1.616]	-0.643 [1.730]
Other x Famine	0.000 [0.004]	0.000 [0.004]	0.003 [0.004]	0.002 [0.003]	-0.004 [0.006]	-0.003 [0.005]	-0.005 [0.007]	-0.003 [0.007]
Grain, x Famine	N	Y	N	N	N	Y	N	N
Yield pc, x Famine	N	N	Y	Y	N	N	Y	Y
Log distance to the capital x Famine				0.005*** [0.002]				-0.001 [0.001]
Observations	1,297	1,297	1,297	1,297	337	337	337	337
R-squared	0.870	0.870	0.875	0.877	0.843	0.846	0.843	0.844
Provinces	50	50	50	50	19	19	19	19

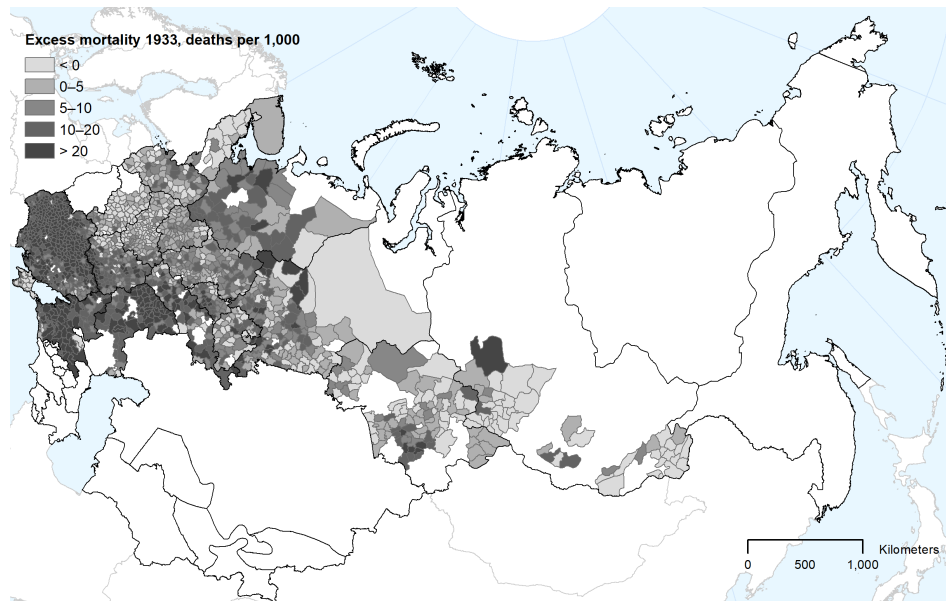
Notes: Observations are at the province and year level (In Columns 1–4, 50 European provinces of the Russian Empire; in Columns 5–8, 19 provinces of the Soviet Union). Mortality is the number of deaths divided by population. Ukrainians, Belorussians, Jews, Germans, Poles, and Kazakhs are the share of each ethnic group in the rural population according to the 1897 Census (Columns 1–4) or the 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 the 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: Excess Mortality 1933

(a) Province-level

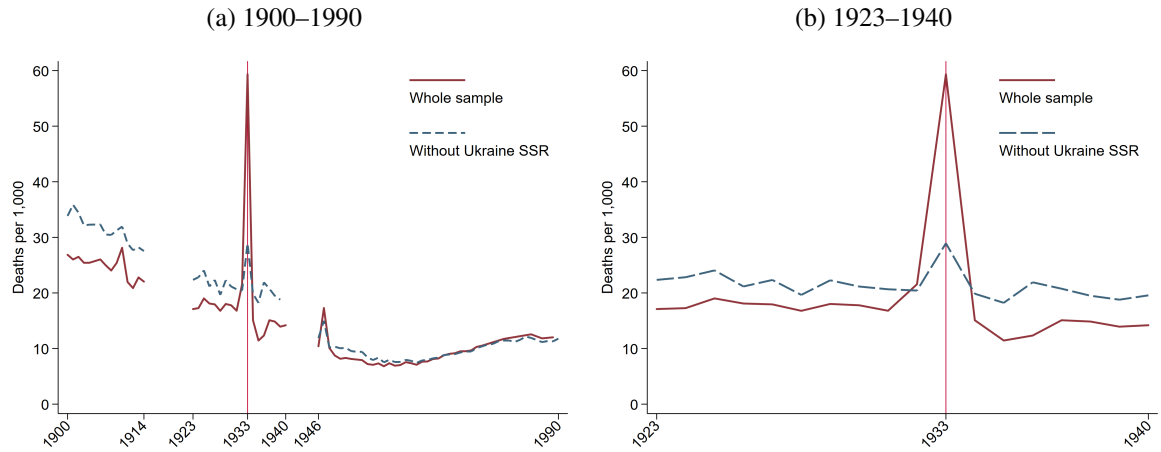


(b) District-level



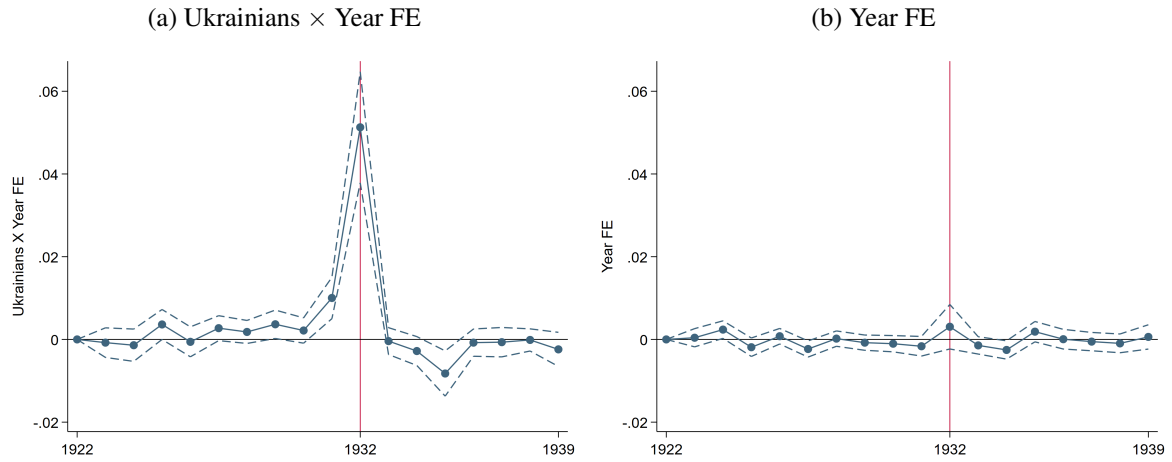
Notes: Excess mortality 1933 is mortality in 1933 minus mortality in 1928. *Source:* See Appendix C.

Figure 2: Mortality



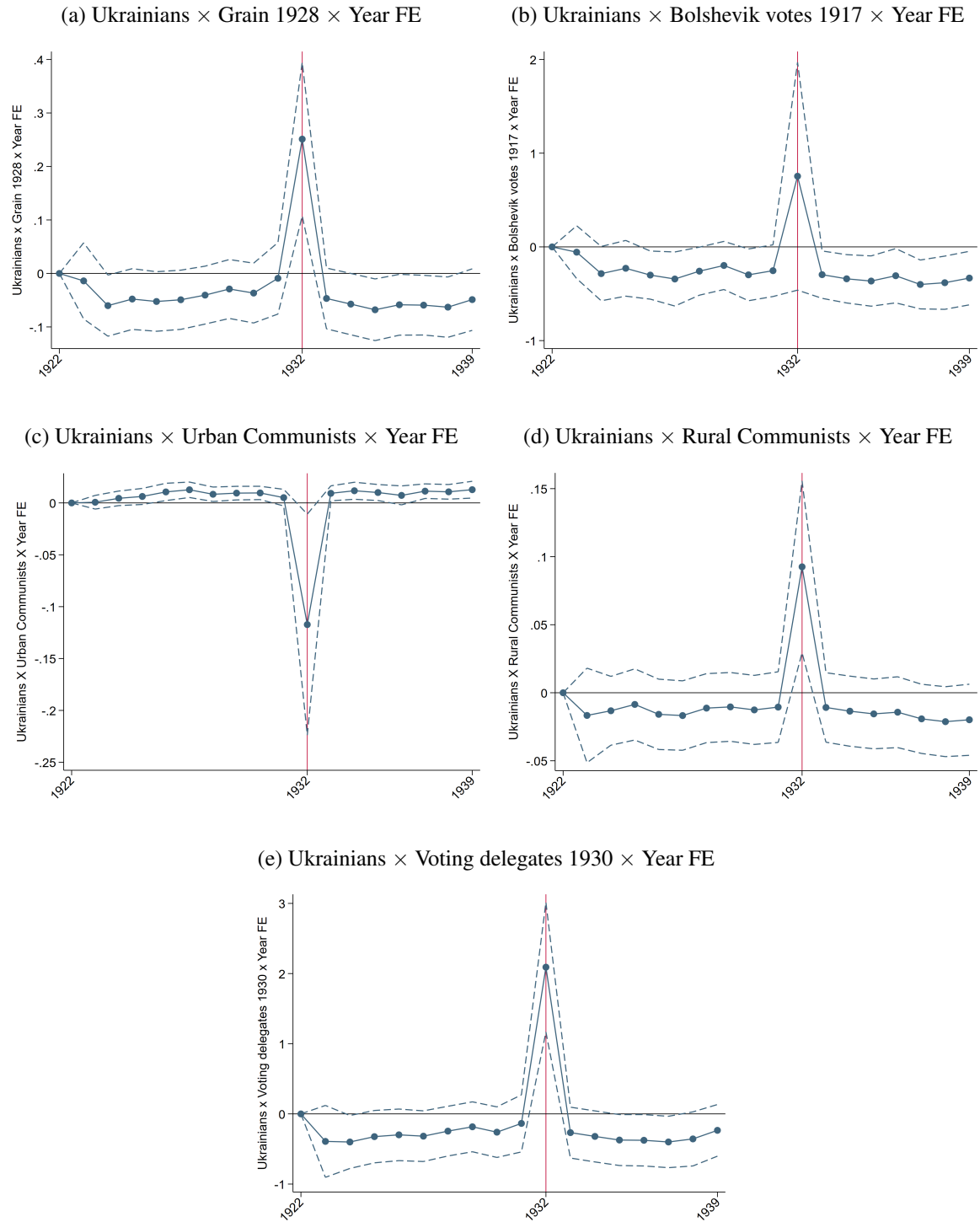
Notes: Mortality is the total number of deaths per 1,000 individuals. Source: See Appendix C.

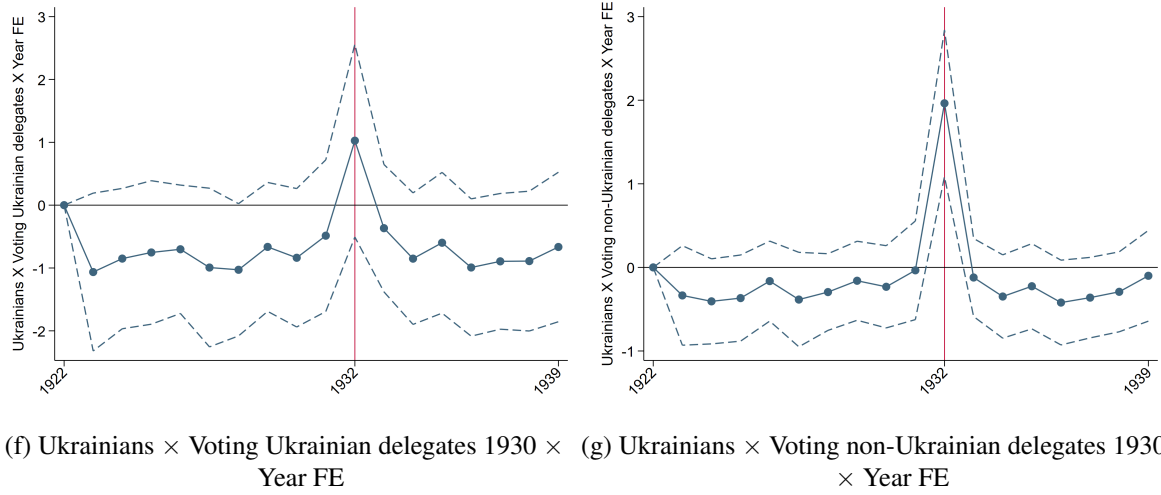
Figure 3: Ethnic Ukrainians and Famine Mortality



Notes: The figures show regression coefficients with their 95% 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 1922 year indicator is omitted for comparison. The coefficients plotted in figures (a) and (b) are estimated from one regression. Standard errors are robust to heteroskedasticity. Source: See Appendix C.

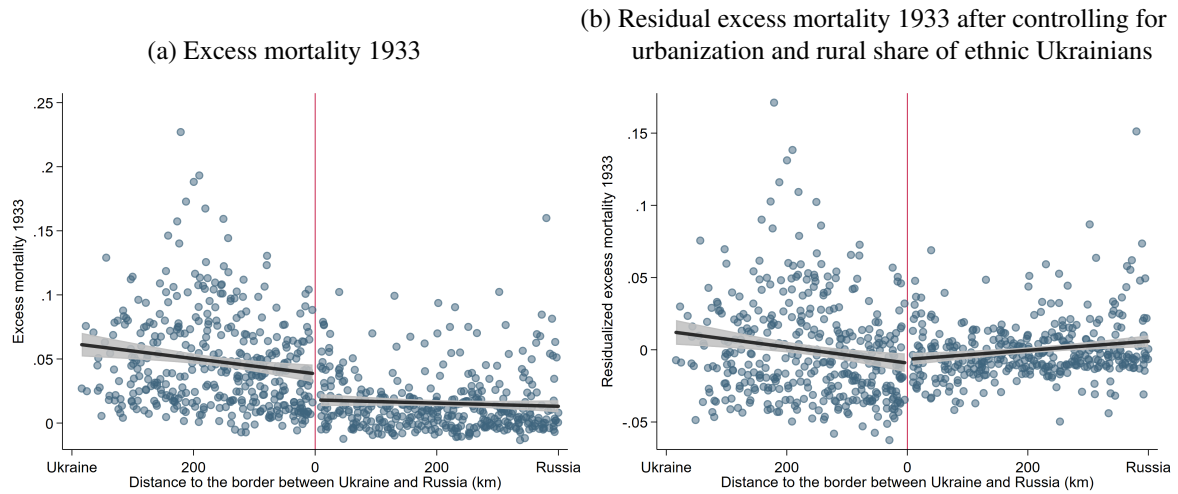
Figure 4: Heterogenous Effects





Notes: Figures (a), (b) and (e) plot coefficients and their 95% confidence intervals estimated from separate regressions. Figures (c) and (d) are estimated from one regression. Figures (f) and (g) are estimated from one regression. The dependent variable in each regression is mortality in year $t + 1$. The right hand side variables include the triple interaction stated in the sub-figure heading, all of the lower order interaction terms, urbanization, urbanization \times the famine indicator, urbanization \times the rural share of ethnic Ukrainians \times the famine indicator, predicted grain, predicted grain \times the famine indicator, predicted grain \times the rural share of ethnic Ukrainians \times the famine indicator, year FE, and province FE. 1922 is the omitted reference year. Standard errors are robust to heteroskedasticity. *Source:* See Appendix C.

Figure 5: Excess Mortality 1933 at the border between Ukraine and Russia



Notes: Excess mortality 1933 is mortality in 1933 minus mortality in 1928. Distance to the border is measured in kilometers. *Source:* See Appendix C.

Appendix

A Predicted Grain

To estimate grain production function we regress 1901–1915 log grain on log province area, log FAO GAEZ grain suitability index, their interaction, temperature and precipitation figures for four seasons, their pairwise interactions and square terms (without a constant). The seasons are: fall (October, November, and December of the previous calendar year), winter (January, February, March), spring (April, May, June), summer (July, August, September). Appendix Table A.2 shows the estimated grain production function. We then use this production function to predict grain harvest from 1922 to 1940. The predicted grain and actual grain are closely correlated; the two exceptions are Karelia and Eastern Siberia provinces, both are likely a result of errors in our matching procedure. In-sample R-squared is 0.90; out-of-sample R-squared is 0.77 (Appendix Figure A.3). This is consistent with the lack of major technological changes in the Soviet agriculture before the 1930s argued by historians (Allen, 2003).

B Other Ethnicities

In Appendix Table A.6, we estimate the heterogeneous treatment effects with the rural share of Belorussians and Kazakhs 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 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.

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

Ethnicity	Population,		Population,	
	millions	% of population	millions	% of population
	1926 census		1939 census	
Panel A. All USSR				
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
Panel B. Our Sample: Belarus SSR, Russia SSR, Ukraine SSR				
Russians	75.3	59.3	94.8	65.0
Ukrainians	30.1	23.7	27.1	18.6
Belorussians	4.7	3.7	5.2	3.6
Kazakhs	0.1	0.1	0.4	0.3
Jews	2.5	2.0	2.9	2.0
Germans	1.1	0.9	1.3	0.9
Poles	0.7	0.6	0.6	0.4

Table A.2: Grain Production Function

	Dependent variable: Log grain
Log area	0.352*** [0.067]
Log grain suitability	-4.643*** [0.640]
Log area x Log grain suitability	0.278*** [0.023]
Fall temperature	0.015 [0.037]
Winter temperature	0.027 [0.043]
Spring temperature	-0.169** [0.079]
Summer temperature	-0.978*** [0.194]
Fall precipitation	-0.006 [0.006]
Winter precipitation	-0.005 [0.007]
Spring precipitation	0.010 [0.007]
Summer precipitation	-0.025*** [0.009]
Fall temperature x Fall precipitation	0.000* [0.000]
Winter temperature x Winter precipitation	0.001* [0.000]
Spring temperature x Spring precipitation	0.000 [0.000]
Summer temperature x Summer precipitation	0.001*** [0.000]
Fall temperature^2	0.004** [0.002]
Winter temperature^2	0.000 [0.002]
Spring temperature^2	-0.001 [0.003]
Summer temperature^2	0.028*** [0.005]
Fall precipitation^2	0.000 [0.000]
Winter precipitation^2	0.000 [0.000]
Spring precipitation^2	-0.000** [0.000]
Summer precipitation^2	0.000** [0.000]
Observations	220
R-squared	0.998

Notes: Observations are at the province and year level. Log grain is the logarithm of grain harvest. Log area is the logarithm of province area. Log grain suitability is the logarithm of the province's FAO GAEZ grain suitability index for rain-fed low-input agriculture. Fall is October, November, December of the previous calendar year; Winter is January, February, March; Spring is April, May, June; Summer is July, August, September. Appendix C shows the exact source of every variable.

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

	Dependent Variable: Mortality in year t+1							
	Reported grain				Predicted grain			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grain	0.000 [0.000]	0.000 [0.000]	0.001 [0.001]	0.002 [0.001]	0.000 [0.000]	0.000 [0.000]	0.001 [0.001]	0.001 [0.001]
Grain x Famine		0.012** [0.005]	0.012** [0.005]	-0.013** [0.007]		0.008 [0.006]	0.008 [0.006]	0.002 [0.006]
Procurement				-0.002 [0.003]				0.001 [0.002]
Procurement x Famine				0.053*** [0.016]				0.027*** [0.010]
Urbanization	-0.011** [0.004]	-0.010** [0.004]	0.017* [0.010]	0.011 [0.007]	-0.012*** [0.004]	-0.011** [0.004]	0.015 [0.010]	0.011 [0.009]
Urbanization x Famine	0.002 [0.014]	0.014 [0.019]	0.007 [0.018]	-0.007 [0.012]	0.002 [0.014]	0.010 [0.014]	0.003 [0.013]	0.003 [0.013]
Observations	337	337	184	184	337	337	184	184
R-squared	0.618	0.671	0.651	0.712	0.618	0.657	0.628	0.701

Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Grain is reported grain harvest (Columns 1-4), or predicted grain harvest (Columns 5-8). 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 on Rural Horse Power, Controlling for the Lagged Number of Work Horses

	A. Dependent variable: Tractors' horse power				B. Dependent variable: Work horses			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	0.065 [0.067]	-0.053 [0.048]	-0.077 [0.084]	0.024 [0.036]	0.109* [0.064]	0.049 [0.080]	0.034 [0.199]	0.249** [0.123]
Grain 1928 x Famine	-0.002 [0.003]				-0.024*** [0.004]			
Ukrainians x Grain 1928 x Famine	-0.199 [0.122]				-0.071 [0.071]			
Bolshevik votes 1917 x Famine		0.014 [0.009]				0.029 [0.024]		
Ukrainians x Bolshevik votes 1917 x Famine		-0.685 [0.562]				-0.518 [0.673]		
Urban communists x Famine			0.000 [0.001]				0.005*** [0.002]	
Ukrainians x Urban communists x Famine			0.092 [0.080]				0.070 [0.069]	
Rural communists x Famine			0.000 [0.001]				-0.003 [0.004]	
Ukrainians x Rural communists x Famine			-0.068 [0.047]				-0.037 [0.026]	
Voting Ukrainian delegates 1930 x Famine				-0.323 [0.265]				0.690* [0.389]
Ukrainians x Voting Ukrainian delegates 1930 x Famine				-1.602 [1.011]				-3.900*** [1.288]
Voting non-Ukrainian delegates 1930 x Famine				0.027 [0.024]				-0.112*** [0.042]
Ukrainians x Voting non-Ukrainian delegates 1930 x Famine				-1.763** [0.872]				-2.235*** [0.678]
Observations	247	247	247	247	266	266	266	266
R-squared	0.807	0.805	0.807	0.807	0.972	0.97	0.972	0.971

Notes: Observations are at the province and year level. Tractors' horse power is the reported tractors' horse power per capita. Work horses is the number of work horses per capita. 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. 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 regressions control for Urbanization, Urbanization x Famine, Ukrainians x Urbanization x Famine, Predicted grain, Predicted grain x Famine, Ukrainians x Predicted grain x Famine, lagged number of work horses, province, and year fixed effects. Standard errors robust to heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Table A.5: Baseline Estimates with District-Level Data (Analogous to Table 1)

	Dependent variable: Mortality								
	Baseline	Province- Year FE	Total Ukrainians	Urban Ukrainians	Mother Tongue Ukrainian	Grain suitability	Omit Ukraine SSR		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ukrainians x Famine	0.048*** [0.002]	0.039*** [0.005]	0.051*** [0.002]	0.064*** [0.003]	0.050*** [0.002]	0.046*** [0.002]	0.043*** [0.004]	0.041*** [0.003]	0.048*** [0.002]
<i>Normalized</i>	<i>0.617</i>	<i>0.499</i>	<i>0.625</i>	<i>0.633</i>	<i>0.627</i>	<i>0.592</i>	<i>0.366</i>	<i>0.526</i>	<i>0.619</i>
Urbanization	0.002 [0.006]	0.005 [0.006]	-0.002 [0.006]	-0.008 [0.007]	0.002 [0.006]	0.002 [0.006]	-0.003 [0.005]	0.004 [0.006]	0.002 [0.006]
Urbanization x Famine	-0.008*** [0.003]	-0.009*** [0.003]	-0.004 [0.003]	-0.004 [0.003]	-0.007** [0.003]	-0.007** [0.003]	0.002 [0.003]	-0.009*** [0.003]	-0.008*** [0.003]
Grain suitability x Famine	N	N	N	N	N	Y	N	N	N
Latitude, Longitude x Famine	N	N	N	N	N	N	N	Y	N
Gender ratio x Famine	N	N	N	N	N	N	N	N	Y
Observations	3,513	3,513	3,515	2,052	3,505	3,513	2,734	3,513	3,513
R-squared	0.727	0.763	0.731	0.747	0.732	0.730	0.628	0.730	0.727
Provinces	18	18	18	18	18	18	17	18	18
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	N	Y	Y	Y	Y	Y	Y	Y
Province-Year FE	N	Y	N	N	N	N	N	N	N
Mortality 1933									
Mean	0.038	0.038	0.038	0.041	0.038	0.038	0.030	0.038	0.038
Std. Dev.	0.028	0.028	0.028	0.029	0.028	0.028	0.019	0.028	0.028
Ukrainians									
Mean	0.237	0.237	0.227	0.215	0.219	0.237	0.069	0.237	0.237
Std. Dev.	0.363	0.363	0.348	0.290	0.358	0.363	0.165	0.363	0.363

Notes: Observations are at the district and year level, the sample has two years: 1928 and 1933. Ukrainians is the share of ethnic Ukrainians in the rural population (Columns 1-2, 6-9), total population (Column 3), urban population (Column 4), or the share of people whose mother tongue is Ukrainian (Column 5) according to the 1926 Census. Famine is an indicator that equals one in 1933 and zero in 1928. Estimates in Column 6 control for FAO GAEZ Grain suitability index x Famine. Estimates in Column 8 control for the district centroid Latitude x Famine, Longitude x Famine, and Latitude x Longitude x Famine. Estimates in Column 9 control for the male/female ratio x Famine (according to the 1926 Census). Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

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

	Dependent variable: Mortality in year t+1										
	Belorussians					Kazakhs			Germans		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ethnic Group Share x Famine	-0.628 [0.450]	-0.971 [1.815]	-1.791 [1.331]	-0.727 [0.632]	30.924 [215.626]	-10.696 [7.439]	579.803*** [123.663]	-65.380** [28.703]	7.828*** [1.191]	4.812 [4.115]	4.573*** [0.957]
Grain 1928 x Famine	0.005 [0.006]				-0.002 [0.008]				0.003 [0.002]		
Ethnic Group Share x Grain 1928 x Famine	-2.001*** [0.620]				-15.065 [119.769]				-2.251*** [0.316]		
Bolshevik votes 1917 x Famine		-0.019 [0.018]				-0.013 [0.011]				-0.011 [0.009]	
Ethnic Group Share x Bolshevik votes 1917 x Famine		2.151 [2.924]				84.079* [42.768]				-11.208 [10.397]	
Urban communists x Famine			-0.001 [0.003]				0.000 [0.002]				-0.002*** [0.001]
Ethnic Group Share x Urban communists x Famine			0.447 [1.282]				257.372*** [67.619]				1.170*** [0.328]
Rural communists x Famine			0.002 [0.002]				0.002 [0.003]				0.001 [0.002]
Ethnic Group Share x Rural communists x Famine			-1.147*** [0.399]				-578.715*** [132.390]				-0.394* [0.217]
Voting delegates 1930 x Famine				0.062** [0.029]				0.049* [0.027]			
Ethnic Group Share x Voting delegates 1930 x Famine				-7.028** [2.960]				-1007.760** [412.142]			
Observations	337	337	337	337	337	337	337	337	337	337	337
R-squared	0.765	0.734	0.768	0.763	0.757	0.762	0.777	0.770	0.827	0.768	0.829

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 Belorussians (Columns 1–4), Kazakhs (Columns 5–8), (Columns 9–12) 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 day. Bolshevik votes 1917 is the share of votes for Bolsheviks in the 1917 elections. Urban (rural) communists is the average over 1922, 1927, and 1931 of the number of urban (rural) Communist Party 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 people. All estimates control for Predicted grain, Predicted grain x Famine, Ethnic group share x Famine, Urbanization, Urbanization x Famine, Ethnic group share x Urbanization x province and year fixed effects. Standard errors robust for heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

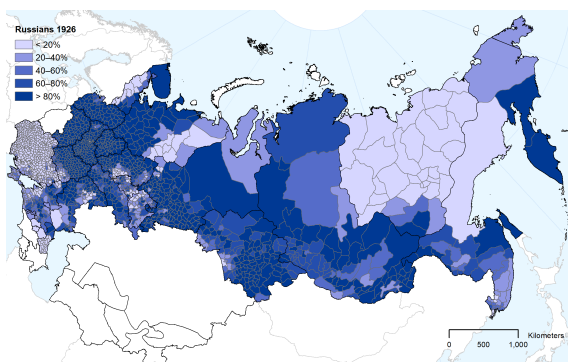
Table A.7: Robustness Checks for Table

	Dependent Variable: Mortality in year t+1							
	Baseline				Baseline with info on land 1905			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians x Famine	0.051*** [0.007]	0.052*** [0.006]	0.051*** [0.007]	0.052*** [0.006]	0.051*** [0.007]	0.059*** [0.005]	0.061*** [0.005]	0.051*** [0.008]
Urbanization	-0.010** [0.004]	-0.010** [0.004]	-0.010** [0.004]	-0.010** [0.004]	-0.007 [0.005]	-0.007 [0.004]	-0.007 [0.004]	-0.007 [0.005]
Urbanization x Famine	0.004 [0.009]	0.003 [0.008]	0.004 [0.009]	0.004 [0.009]	0.002 [0.009]	-0.004 [0.009]	-0.004 [0.009]	0.002 [0.009]
Share of catholics 1897 x Famine		-0.082*** [0.023]		-0.088*** [0.032]				
Share of orthodox christians 1897 x Famine			-0.002 [0.011]	-0.005 [0.013]				
Share of peasant land in repartition commune 1905 x Famine						0.021** [0.009]		
Share of peasant households in repartition commune 1905 x Famine							0.024** [0.009]	
Peasant and private land gini 1905 x Famine								0.000 [0.008]
Observations	337	337	337	337	286	286	286	286
R-squared	0.781	0.786	0.781	0.787	0.783	0.795	0.798	0.783

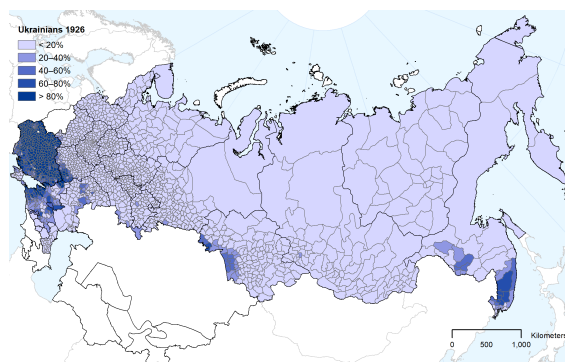
Notes: Observations are at the province and year level. Mortality is the number of deaths divided by population. Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Census. Famine is an indicator that equals one in 1932 and zero otherwise. Estimates in Columns 2 and 4 control for the share of Catholics in the population according to the 1897 Census interacted with the famine indicator. Estimates in Columns 3 and 4 control for the share of Orthodox Christians in the population according to the 1897 Census interacted with the famine indicator. Estimates in Column 5 replicate the baseline on a restricted sample of provinces for which data from 1905 Land Census is available. Estimates in Column 6 control for the share of peasant land in repartition commune according to the 1905 Land Census interacted with the famine indicator. Estimates in Column 7 control for the share of peasant households that belonged to a repartition commune according to the 1905 Land Census interacted with the famine indicator. Estimates in Column 8 control for land gini coefficient from the 1905 Land Census interacted with the famine indicator. All estimates control for province and year fixed effects. Standard errors robust to heteroskedasticity are in brackets. Appendix C shows the exact source of every variable.

Figure A.1: Rural Ethnic Composition 1926

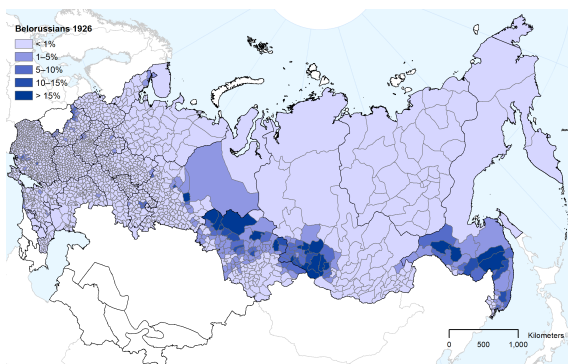
(a) Russians



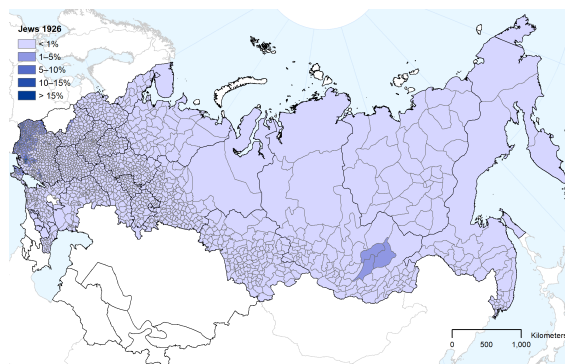
(b) Ukrainians



(c) Belorussians



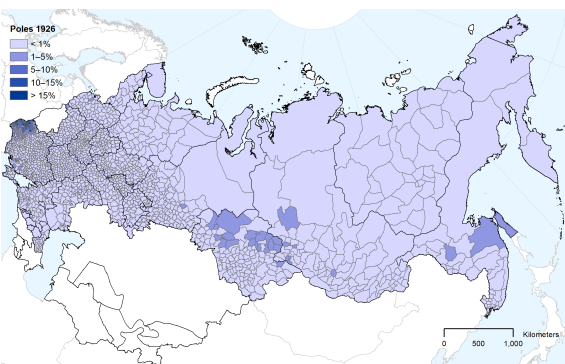
(d) Jews



(e) Germans



(f) Poles

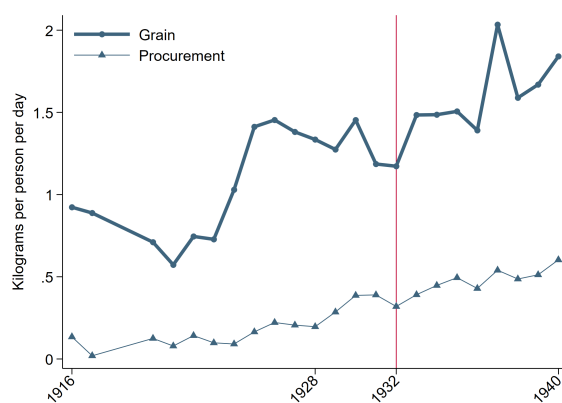




(g) Kazakhs

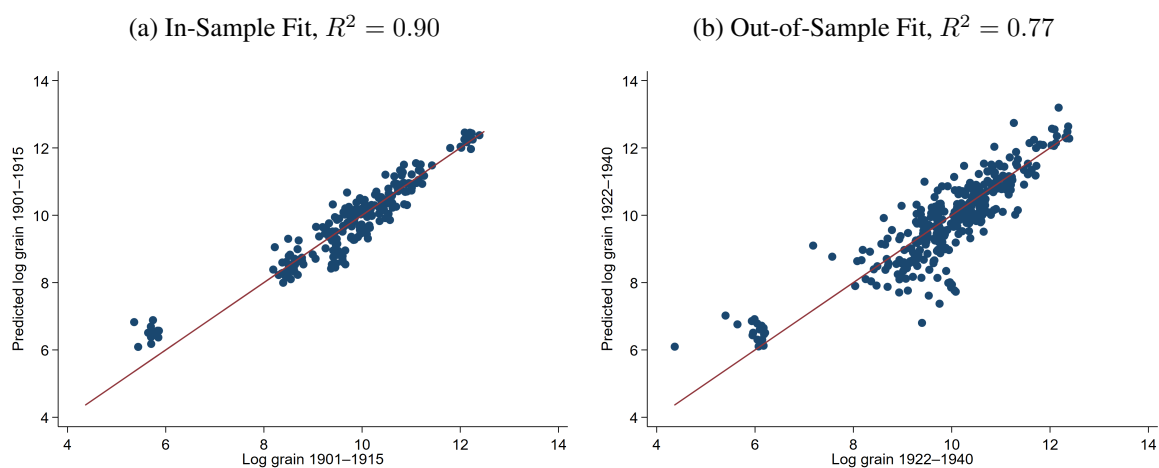
Notes: The share of (a) ethnic Russians, (b) Ukrainians, (c) Belorussians, (d) Jews, (e) Germans, (f) Poles, or (g) Kazakhs in the rural population. Notice that the scale changes from figures (a)–(b) to figures (c)–(g). *Source:* The 1926 Population Census, see Appendix C.

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



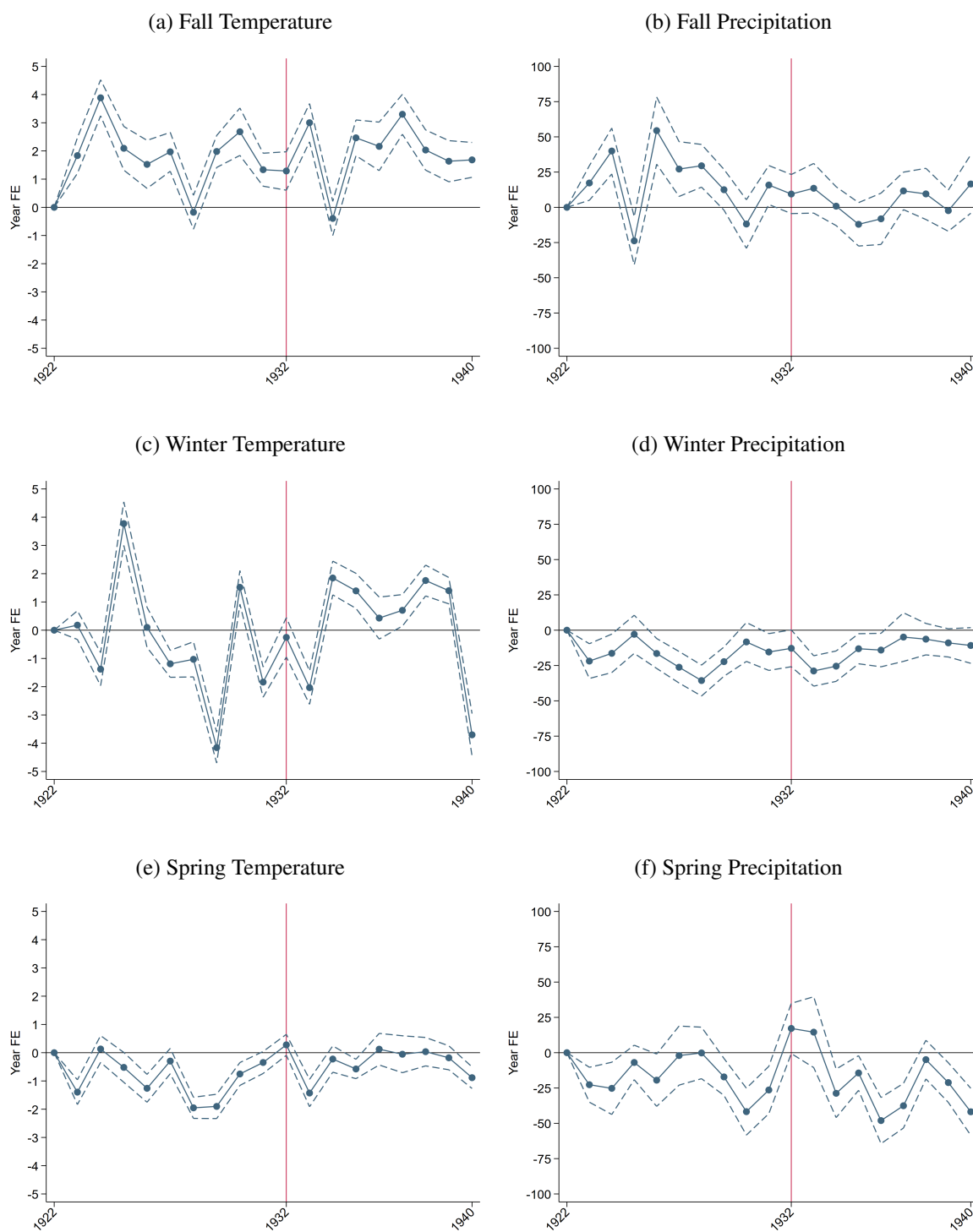
Source: See Data Appendix, Section C.4

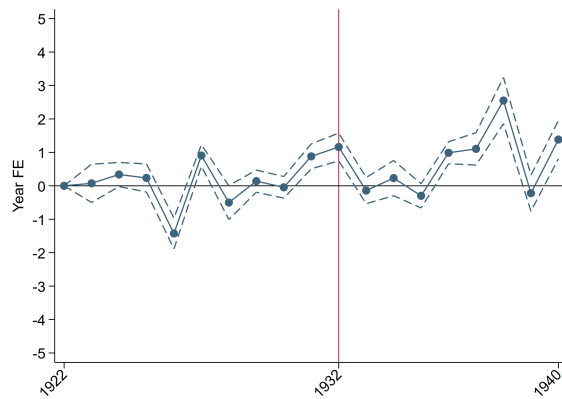
Figure A.3: Reported and Predicted Grain



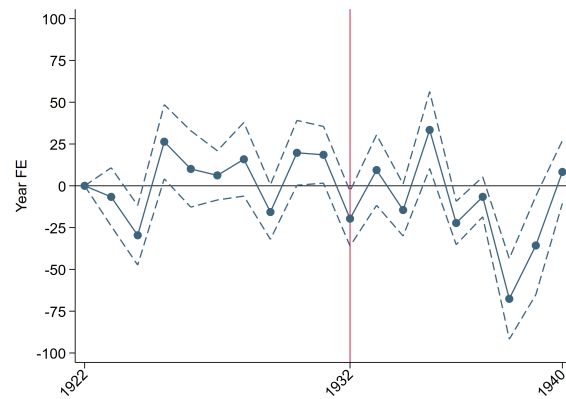
Notes: The figures show logs of reported and predicted grain with a 45-degree line; (a) for 1901–1915, a sample on which grain production function is estimated (in-sample fit), and (b) for 1922–1940 (out-of-sample fit); see Appendix section A for details., Appendix C shows the exact source of every variable.

Figure A.4: Seasonal Temperature and Precipitation





(g) Summer Temperature



(h) Summer Precipitation

Notes: Regression coefficients with their 95% confidence intervals from regressing each season's temperature (precipitation) on year indicators and province fixed effects. The 1922 year indicator is omitted for comparison. Temperature is measured in degrees Celsius. Precipitation is measured in millimeters. Fall is October, November, December; Winter is January, February, March; Spring is April, May, June; Summer is July, August, September. Standard errors are robust for heteroskedasticity. Each set of coefficients is from a separate regression. *Source:* Matsuura and Willmott (2014), see Appendix C.

C Data Appendix (Not for Publication)

C.1 Province-level data

Province-level panel dataset spans the years of 1922 to 1940 and covers 19 provinces of the republics of Belarus, Russia, and Ukraine. These provinces correspond to the 1934 administrative division. Belarus and Ukraine are a single province each. Our dataset covers 84% of the 1926 population of the Soviet Union and 88% of the 1928 grain production. Omitted are the territories for which no reliable mortality data are available: Far Eastern Province, Yakut Autonomous SSR, and the North Caucasus ethnic territories: Chechen Autonomous Province, Cherkess Autonomous Province, Dagestan Autonomous SSR, Ingush Autonomous Province, Kabardino-Balkarian Autonomous Province, Karachay Autonomous Province, North Ossetian Autonomous Province. Figure 1a shows our provinces on the map (omitted territories are in white).

Total and urban population

- 1897: the 1897 Population Census from Demoscope.ru.
- 1898–1903: is interpolated between 1897 and 1904.
- 1904–1913: *Yezhegodnik Rossii* 1904–1916.
- 1920: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskii yezhegodnik 1924 god (Vypusk pervyy)* [Statistical Yearbook 1924 (First Issue)]”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part I, Table 1.B.
- 1922: total population is interpolated between 1920 and 1923; urban population is interpolated between 1920 and 1925.
- 1923: total population is calculated using the total number of deaths and deaths per 10,000 from Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskii yezhegodnik 1924 god (Vypusk pervyy)* [Statistical Yearbook 1924 (First Issue)]”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part I, Table 5; urban population is interpolated between 1920 and 1925.
- 1924: total population is calculated using the total number of deaths and deaths per 10,000 from Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskii yezhegodnik 1924 god (Vypusk pervyy)* [Statistical Yearbook 1924 (First Issue)]”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part I, Table 8; urban population is interpolated between 1920 and 1925.

- 1925: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskiy yezhegodnik 1924 god (Vypusk pervyy) [Statistical Yearbook 1924 (First Issue)]*”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part I, Table 1.B.
- 1926: is interpolated between 1925 and 1927.
- 1927: December 1926 Population Census.
- 1928–1932: is interpolated between 1927 and 1933.
- 1933: RGAE 1562/329/19 p. 1–12.
- 1934–1936: is interpolated between 1933 and 1937.
- 1937: the 1937 Population Census from Zhiromskaya, V.B. and Kiselev, I.N. and Polyakov, Yu.A. (1996) “*Polveka pod grifom “sekretno”: Vsesoyuznaya perepis naseleniya 1937 goda [Classified for half a century: All-Union population census of 1937]*”, Moscow: Nauka.
- 1938: is interpolated between 1937 and 1939.
- 1939: the 1939 Population Census corrected for the centralized additions (*pripiski*) from Demoscope.ru.
- 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: is interpolated between 1950 and 1959.
- 1959: the 1959 Population Census from Demoscope.ru.
- 1960: is interpolated between 1959 and 1961.
- 1961–1966: *Narodnoye khozyaystvo RSFSR, Narodnoye khozyaystvo SSSR* 1960, 1961, 1962, 1963, 1964, 1965.
- 1967: is interpolated between 1966 and 1968.
- 1968: *Narodnoye khozyaystvo RSFSR, Narodnoye khozyaystvo SSSR* 1967.
- 1969: is interpolated between 1968 and 1970.

- 1970: the 1970 Population Census from Demoscope.ru.
- 1971–1976: *Narodnoye khozyaystvo RSFSR, Narodnoye khozyaystvo SSSR* 1970, 1971, 1972, 1973, 1974, 1975.
- 1977: is interpolated between 1976 and 1978.
- 1978: *Narodnoye khozyaystvo RSFSR, Narodnoye khozyaystvo SSSR* 1977.
- 1979: the 1979 Population Census from Demoscope.ru.
- 1980–1986: *Narodnoye khozyaystvo RSFSR, Narodnoye khozyaystvo SSSR* 1979, 1980, 1981, 1982, 1983, 1984, 1985.
- 1987–1988: is interpolated between 1986 and 1989.
- 1989: the 1989 Population Census from Demoscope.ru.
- 1990: used 1989 value.

Except for 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 procedure is legitimate because reported data are more disaggregated than our provinces. 1933 is used as reported. 1900–1914 and 1946–1990 data are only used for Figure 2a.

Deaths

- 1900–1914: *Yezhegodnik Rossii* 1904–1916.
- 1923: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskii yezhegodnik 1924 god (Vypusk pervyy) [Statistical Yearbook 1924 (First Issue)]*”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part I, Table 5.
- 1924: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskii yezhegodnik 1924 god (Vypusk pervyy) [Statistical Yearbook 1924 (First Issue)]*”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part I, Table 8.
- 1925: Tsentralnoye Statisticheskoye Upravleniye SSSR [Central Statistical Office of the USSR] (1928) “*Yestestvennoye dvizheniye naseleniya Soyuzo SSR 1923–1925 [Natural movement of the population of the USSR]*”, Volume I, Issue 1, Table 1.
- 1926: Yestestvennoye dvizheniye naseleniya Soyuzo SSR v 1926 g, Izdaniye TsSU SSSR (1929), Table 1
- 1927–1932: Belarus, Ukraine – RGAE 1562/329/256; Russia – Demoscope.ru.

- 1933–1940: Demoscope.ru.
- 1946–1990: Demoscope.ru, *Narodnoye khozyaystvo UkrSSR, Narodnoye khozyaystvo BelSSR* (1986 not reported, appears to be still classified, probably because of Chernobyl).

Except for 1933, we calculated deaths 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 2a.

Mortality

Mortality is the total number of deaths divided by population (what demographers call crude death rate).

Ethnic composition

Ethnic composition comes from the 1897 and the 1926 Population Censuses. The 1897 Census reports population by mother tongue. We use the share of people whose mother tongue is Belorussian, German, Jewish (*Yevreyskiy*), Kazakh (*Kyrgyz-Kaysatsky*), Polish, Russian (*Velikorusskiy*), and Ukrainian (*Malorusskiy*). The 1926 Census reports population by self-proclaimed ethnicity and by mother tongue, we use both. Data is calculated in our province borders using hand-created district-level 1897 and 1926 maps.

Age structure

Region (*okrug*)-level population by 1-year age groups from the 1926 Population Census is reported by Demoscope.ru. We calculated the share of people aged 10 and younger using hand-created region (*okrug*)-level map. This procedure is legitimate because regions (*okruga*) are smaller than our provinces.

Gender ratio

Male to female ratio is from the 1926 Population Census. We calculated it in our province borders using hand-created district (*volost*)-level 1926 map. This procedure is legitimate because districts (*volosty*) are smaller than our provinces.

Grain harvest, sown area, and yield

- 1901–1914: Obukhov V.M. (1927) “*Dvizheniye urozhayev zernovykh kultur v Yevropeyskoy Rossii v period 1883–1915 g.g. [Movement of grain crops in European Russia in the period 1883–1915]*” and *Yezhegodnik Rossii* 1904–1916.

- 1922: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1924) “*Sbornik statisticheskikh svedeniy po Soyuzu SSR 1918–1923. Za pyat let raboty Tsentralnogo Statisticheskogo Upravleniya* [A collection of statistical information on the USSR 1918–1923. Five years of work of the Central Statistical Office.]”, Volume XVIII of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part VI, Tables 7 and 8.
- 1923: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1924) “*Statisticheskii yezhegodnik 1922 i 1923 g. (Vypusk pervyy)* [Statistical Yearbook 1922 and 1923 (First Issue)]”, Volume VIII, Issue 5 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part III, Tables 3 and 4.
- 1924: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskii yezhegodnik 1924 god (Vypusk pervyy)* [Statistical Yearbook 1924 (First Issue)]”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part III, Tables 6 and 7.
- 1925–1927: Statisticheskoye izdatelstvo TsSU SSSR [Statistical Publishing House of the Central Statistical Office of the USSR] (1929) “*Selskoye khozyaystvo SSSR 1925–1928. Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii* [Agriculture of the USSR 1925–1928. A collection of statistical information for the XVI All-Union Party Congress]”, Part III.
- 1928: RGAE 1562/329/1409.
- 1929–1930: Gosudarstvennoye sotsialno-ekonomicheskoye izdatelstvo [State Socio-Economic Publishing House] (1932) “*Narodnoye khozyaystvo SSSR. Statisticheskii spravochnik 1932* [The national economy of the USSR. Statistical Handbook 1932]”, Part II.3.A, Tables 30 and 33.
- 1931: Gosudarstvennoye izdatelstvo kolkhoznoy i sovkhoznoy literatury “Selkhozgiz” [State publishing house of collective and state farm literature “Selkhozgiz”] (1936) “*Selskoye khozyaystvo SSSR. Yezhegodnik 1935* [Agriculture of the USSR. Yearbook 1935]”, p. 269, Tables 106 and 107.
- 1932–1940: RGAE 1562/329/1409.

We use the 1901–1914 grain to estimate grain production function. We calculate grain data in administrative borders corresponding to our provinces using hand-created ArcGIS maps (each year is reported using a different administrative division). The years 1922, 1924–1927 are reported for larger units than our provinces. The data is calculated in our province borders in proportion to the 1913 district (*uezd*) sown area.

Wheat, rye, and other grain 1928

Harvest decomposition by cultures is from RGAE 1562/329/1409.

Procurement

- 1924: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) “*Yezhegodnik khlebnoy trgovli N1 [Yearbook of grain trade N 1]*”, Table 6.
- 1925: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) “*Yezhegodnik khlebnoy trgovli N1 [Yearbook of grain trade N 1]*”, Table 14.
- 1926: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) “*Yezhegodnik khlebnoy trgovli N1 [Yearbook of grain trade N 1]*”, Table 22.
- 1927: Statisticheskoye izdatelstvo TsSU SSSR [Statistical Publishing House of the Central Statistical Office of the USSR] (1929) “*Selskoye khozyaystvo SSSR 1925–1928. Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii [Agriculture of the USSR 1925–1928. A collection of statistical information for the XVI All-Union Party Congress]*”, Part V.
- 1928: calculated from the 1928 grain harvest and procurement as a share of harvest from RGAE 4372/30/871 p. 30.
- 1929: Narodnyy Komissariat Snabzheniya SSSR [People’s Commissariat of Supply of the USSR] (1932) “*Yezhegodnik khlebooborota N4 [Yearbook of grain turnover N 4]*”, Tables 3 and 10.
- 1930: Narodnyy Komissariat Snabzheniya SSSR [People’s Commissariat of Supply of the USSR] (1932) “*Yezhegodnik khlebooborota N4 [Yearbook of grain turnover N 4]*”, Table 29 and Table 36
- 1931: Komitet po zagotovkam S.-Kh produktov pri SNK SSSR [Committee for Procurement of Agricultural Products under the Council of People’s Commissars of the USSR] (1934) “*Yezhegodnik khlebooborota za 1931-32, 1932-33 i predvaritelnyye itogi zagotovok 1933 g. [Yearbook of grain turnover for 1931-32, 1932-33 and preliminary results of procurement in 1933]*”, Table 21.
- 1932: Komitet po zagotovkam S.-Kh produktov pri SNK SSSR [Committee for Procurement of Agricultural Products under the Council of People’s Commissars of the USSR] (1934) “*Yezhegodnik khlebooborota za 1931-32, 1932-33 i predvaritelnyye itogi zagotovok 1933 g. [Yearbook of grain turnover for 1931-32, 1932-33 and preliminary results of procurement in 1933]*”, Table 33.
- 1933: Komitet po zagotovkam S.-Kh produktov pri SNK SSSR [Committee for Procurement of Agricultural Products under the Council of People’s Commissars of the USSR] (1934) “*Yezhegodnik khlebooborota za 1931-32, 1932-33 i predvaritelnyye*

itogi zagotovok 1933 g. [Yearbook of grain turnover for 1931-32, 1932-33 and preliminary results of procurement in 1933]”, 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.

Collectivization

- 1927: Statizdat TSSU SSSR [Statistical publishing house of the Central Statistical Office of the USSR] (1929) “*Kollektivizatsiya Sovetskoy derevni. Predvaritelnyye itogi sploshnykh obsledovaniy 1928 i 1929 gg. [Collectivization of the Soviet countryside. Preliminary results of comprehensive surveys in 1928 and 1929]*”, Table 10.
- 1928: RGAE 1562/82/271.
- 1929: Gosplan SSSR i RSFSR. Ekonomiko-statisticheskiy sektor [State Planning Committee of the USSR and the RSFSR. Economic and statistical sector] (1931) “*Kolkhozy v 1929 g. Itogi sploshnogo obsledovaniya kolkhozov [Collective farms in 1929. Results of a comprehensive survey of collective farms]*”.
- 1930: Gosplan SSSR. Upravleniye Narodnokhozyaystvennogo Ucheta [State Planning Committee of the USSR. Department of National Economic Accounting] (1931) “*Kolkhozy v 1930 g. Itogi raportov kolkhozov k XVI s’yezdu VKP(b) [Collective farms in 1930. Resume of the collective farms’ reports to the XVI Congress of the CPSU(b)]*”.
- 1931: Izd. Kolkhoztstva SSSR i RSFSR [Publishing House of the Collective Farm Center of the USSR and the RSFSR] (1931) “*Kolkhoznoye stroitelstvo v SSSR [Collective farms building in the USSR]*”, p. 15 and Davies and Wheatcroft (2004), Table 27.
- 1932: RGAE 1562/82/271.
- 1933: “*Plan. Zhurnal Gosplana i TsUNKhU SSSR [Plan. Journal of the State Planning Committee and TsUNKhU USSR]*”, 2-1933.
- 1934–1936: RGAE 1562/82/271.
- 1937: interpolated between 1936 and 1938.
- 1938: Gosplanizdat (1939) “*Selskoye khozyaystvo Soyuzov SSR 1939 (Statisticheskiy spravochnik) [Agriculture of the USSR 1939 (Statistical handbook)]*”, Part IV.

Collectivization is the share of rural households in collective farms.

Bolshevik votes 1917

Bolshevik vote share is from Protasov et al. (2014). Data is calculated in our province borders using district (*uezd*)-level 1917 map from Castañeda Dower and Markevich (2020).

Urban and Rural Communists

Urban and rural communists is the average number of Communist Party members and candidates over 1922, 1927, and 1931.

- 1922: Izdatelskoye otdeleniye TsK RKP [Publishing Department of the Central Committee of the RCP] (1922) “*Vserossiyskaya perepis chlenov RKP 1922 goda [All-Russian census of the members of the RCP in 1922]*”, Issue 3, Table 6.
- 1927: Statisticheskii otdel TsK VKP(b) [Statistical Department of the Central Committee of the CPSU(b)] (1927) “*Vsesoyuznaya partiynaya perepis 1927 goda. Chislennyi sostav VKP(b) na 10 yanvarya 1927 g. [All-Union Party Census of 1927. The composition of the CPSU(b) on January 10, 1927]*”, Issue 1.
- 1931: Tsentralnyy Komitet VKP(b). Organizatsionno-instruktorskiy otdel [Central Committee of the CPSU(b). Organizational and instructor department] (1932) “*Sostav VKP(b) v tsifrakh. Dinamika osnovnykh pokazateley rosta parti za 1930 i pervoye polugodiye 1931 g. [Composition of the CPSU(b) in numbers. Dynamics of the main indicators of the growth of the party for 1930 and the first half of 1931]*”

We calculated 1922 and 1927 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 are used as reported.

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.

Province Latitude and Longitude

The latitude and longitude of the province centroid, calculated using ArcGIS.

Tractors’ horse power

- 1927–1928: the number of collective farms’ tractors times 13 (the average tractor horse power in 1929) from Vsesoyuznyy Sovet Kolkhozov [All-Union Council of Collective

Farms] (1929) “Kolkhozy SSSR (Statisticheskiiy spravochnik) [Collective farms of the USSR (Statistical handbook)]

- 1929: horse power of tractors belonging to collective farms and to machine-tractor stations from Gosplan SSSR i RSFSR. Ekonomiko-statisticheskiiy sektor [State Planning Committee of the USSR and the RSFSR. Economic and statistical sector] (1931) “Kolkhozy v 1929 g. Itogi sploshnogo obsledovaniya kolkhozov [Collective farms in 1929. Results of a comprehensive survey of collective farms]”, Tables 1 and 2.
- 1930: horse power of tractors belonging to collective farms is from Gosplan SSSR. Upravleniye Narodnokhozyaystvennogo Ucheta [State Planning Committee of the USSR. Department of National Economic Accounting] (1931) “Kolkhozy v 1930 g. Itogi reportov kolkhozov k XVI s’yezdu VKP(b) [Collective farms in 1930. Resume of the collective farms’ reports to the XVI Congress of the CPSU(b)]”; horse power of tractors belonging to machine-tractor stations is from Tsentralnoye Upravleniye Narodnokhozyaystvennogo Ucheta Gosplana SSSR [The Central Statistical Administration of Gosplan] (1935) “Sotsialisticheskoye stroitelstvo SSSR (Statisticheskiiy yezhegodnik), 1935 g. [Socialist construction of the USSR (Statistical Yearbook), 1935]”, Part II.6, Table 3.
- 1931–1934: Tsentralnoye Upravleniye Narodnokhozyaystvennogo Ucheta Gosplana SSSR [The Central Statistical Administration of Gosplan] (1935) “Sotsialisticheskoye stroitelstvo SSSR (Statisticheskiiy yezhegodnik), 1935 g. [Socialist construction of the USSR (Statistical Yearbook), 1935]”, Part II.6, Table 3.
- 1935–1936: RGAE 1562/79/275 p. 26–30.
- 1937: RGAE 1562/81/276a.
- 1937: RGAE 1562/81/269.
- 1937: RGAE 1562/83/222.

In 1929–1930, 87% of tractors belonged to collective farms. In 1931 a shift occurred – the majority of tractors moved to machine-tractor stations (MTS) that served collective farms but formally were a state property. Therefore, we use collective farms’ and machine-tractor stations’ tractors in 1927–1930, and use tractors belonging to machine-tractors stations from 1931 onward. We calculated tractors 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.

Grain suitability

Each province’s average FAO GAEZ wheat suitability index for rain-fed low-input agriculture.

Weather

Land surface temperature and precipitation are from Matsuura and Willmott (2014). For each province, we calculated the province’s average monthly temperature and precipitation using ArcGIS.

C.2 District-level data

District-level dataset spans two years, 1928 and 1933, and covers some 3,500 districts of the republics of Belarus, Russia, and Ukraine. These districts correspond to the 1934 administrative division. Omitted are territories for which no reliable 1933 mortality data are available. Figure 1b shows our districts on the map (omitted territories are in white).

Mortality

- 1928: GARF 374/23/7, 13, 31–32, 67, 72–91, 132, 158.
- 1933: RGAE 1562/329/18–19.

Ethnic composition

Ethnic composition comes from the 1926 Population Censuses. This census reports population by self-proclaimed ethnicity and by mother tongue, we use both. Data is calculated in our district borders using hand-created district (*volost*)-level 1926 map.

Urbanization

- 1928: used value from December 1926 Population Census. This census reports district (*volost*)-level rural population and, separately, the population of each urban settlement. To calculate rural and urban population in 1934 administrative borders, we hand-created district (*volost*)-level 1926 map and located all urban settlements on the map.
- 1933: RGAE 1562/329/18–19.

Distance to a railroad

Distance from each district’s administrative center to the nearest railroad station is from Tsentralnyy Ispolnitelnyy Komitet Soyuz SSR, Vserossiyskiy Tsentralnyy Ispolnitelnyy Komitet [Central Executive Committee of the USSR, All-Russian Central Executive Committee] (1931) “*Administrativno-territorial’noye deleniye Soyuz SSR. Rayony i goroda SSSR [Administrative-territorial division of the USSR. Districts and cities of the USSR]*”. 1931 districts are matched to 1934 districts by name.

Distance to a city with at least 20,000 inhabitants

We located cities with at least 20,000 inhabitants according to the 1926 Census on the map, and calculated distance from each district centroid to the nearest of these cities.

Grain suitability

District's average FAO GAEZ wheat suitability index for rain-fed low-input agriculture.

Gender ratio

Gender ratio is a ratio of males to females according to the 1926 Population Census. To calculate data in 1934 administrative borders, we hand-created district (*volost*)-level 1926 map.

District Latitude and Longitude

The latitude and longitude of the district centroid, calculated using ArcGIS.

C.3 Data on the 1892 famine

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

Population

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

Deaths

- 1885–1896: kindly shared by Volha Charnysh from an ongoing project (Charnysh and McElroy, 2020).
- 1899–1914: *Yezhegodnik Rossii* 1904–1916.

Ethnic composition

1897 Population Census.

Grain, sown area, yield

Obukhov V.M. (1927) “*Dvizheniye urozhayev zernovykh kultur v Yevropeyskoy Rossii v period 1883–1915 g.g. [Movement of grain crops in European Russia in the period 1883–1915]*”.

C.4 Sources of Figure A.2

Total Grain Production

- 1916–1917, 1920–1922: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1924) “*Sbornik statisticheskikh svedeniy po Soyuzu SSR 1918–1923. Za pyat let raboty Tsentralnogo Statisticheskogo Upravleniya [A collection of statistical information on the USSR 1918–1923. Five years of work of the Central Statistical Office.]*”, Volume XVIII of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office].
- 1923: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1924) “*Statisticheskiy yezhegodnik 1922 i 1923 g. (Vypusk pervyy) [Statistical Yearbook 1922 and 1923 (First Issue)]*”, Volume VIII, Issue 5 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office].
- 1924–1927: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) “*Yezhegodnik khlebnoy trgovli N1 [Yearbook of grain trade N 1]*”.
- 1928: RGAE 1562/329/1409.
- 1929–1931: Gosudarstvennoye izdatelstvo kolkhoznoy i sovkhoznoy literatury “Selkhozgiz” [State publishing house of collective and state farm literature “Selkhozgiz”] (1936) “*Selskoye khozyaystvo SSSR. Yezhegodnik 1935 [Agriculture of the USSR. Yearbook 1935]*”.
- 1932–1940: RGAE 1562/329/1409; 1940 figure excludes western regions occupied by the Soviet Union in 1939.

Total Grain Procurement

RGAE 8040/8/360.

Total Population

- 1916: used 1914 figure from Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1921) “*Statisticheskiy yezhegodnik 1918–1920 gg. (Vypusk pervyy) [Statistical Yearbook 1918–1920 (First Issue)]*”, Volume VIII, Issue 1 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office].

- 1920: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1921) “*Statisticheskiy yezhegodnik 1918–1920 gg. (Vypusk pervyy) [Statistical Yearbook 1918–1920 (First Issue)]*”, Volume VIII, Issue 1 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office].
- 1921–1922: interpolated between 1920 and 1923.
- 1923: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1924) “*Sbornik statisticheskikh svedeniy po Soyuzu SSR 1918–1923. Za pyat let raboty Tsentralnogo Statisticheskogo Upravleniya [A collection of statistical information on the USSR 1918–1923. Five years of work of the Central Statistical Office.]*”, Volume XVIII of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office].
- 1924: interpolated between 1923 and 1925.
- 1925: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1926) “*Statisticheskiy yezhegodnik 1924 god (Vypusk pervyy) [Statistical Yearbook 1924 (First Issue)]*”, Volume VIII, Issue 7 of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office].
- 1926: interpolated between 1925 and 1927.
- 1927: December 1926 Population Census.
- 1928–1933: RGAE 4372/30/107.
- 1934–1936: interpolated between 1933 and 1937.
- 1937: the 1937 Population Census from Zhiromskaya, V.B. and Kiselev, I.N. and Polyakov, Yu.A. (1996) “*Polveka pod grifom “sekretno”: Vsesoyuznaya perepis naseleniya 1937 goda [Classified for half a century: All-Union population census of 1937]*”, Moscow: Nauka.
- 1938: interpolated between 1937 and 1939.
- 1939: the 1939 Population Census corrected for the centralized additions (*pripiski*) from Demoscope.ru.
- 1940: used 1939 value.